UNIVERSITY OF OSLO
Department of Informatics

Adaptive Beamforming for Medical Ultrasound Imaging

Johan-Fredrik Synnevåg

November 2008

Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
Nr. 835

ISSN 1501-7710

Cover: Inger Sandved Anfinsen.
Printed in Norway: AiT e-dit AS, Oslo, 2009.

Produced in co-operation with Unipub AS.
The thesis is produced by Unipub AS merely in connection with the
thesis defence. Kindly direct all inquiries regarding the thesis to the copyright
holder or the unit which grants the doctorate.

Unipub AS is owned by
The University Foundation for Student Life (SiO)

Preface

This dissertation has been submitted to the Faculty of Mathematics and Natural Sciences at the University of Oslo (UiO), in partial fulfillment of the requirements for the degree *Philosophiae Doctor* (Ph. D.). The work has been carried out at the group for Digital Signal Processing and Image Analysis (DSB) at the Department of Informatics (IFI), under the supervision of Professor Sverre Holm and Associate Professor Andreas Austeng. The work was financed by IFI, UiO.

Acknowledgements

I would like to thank my supervisors, Professor Sverre Holm and Associate Professor Andreas Austeng, for their contributions to this work. Sverre introduced me to the field of signal processing. His broad knowledge of the field and guidance through this work has been of great value. I would like to thank him for all his input to this work. He was also my supervisor for the cand. scient. degree. His belief in my skills at the time was of great importance when deciding to go back to the university to do a ph.d.

Andreas has been an important discussion partner throughout this project. I am grateful for all the encouragement and support I have received. Also, I would like to thank him for his contributions to get financing for the project.

I would also like to thank my family for their support during the last four years – especially my wife, Siw. Also, special thanks to my parents for always supporting me.

Finally, thanks to the whole of the DSB group for making an inspiring and social working environment.

Johan-Fredrik Synnevåg
Bergen, Norway, November 2008

Abstract

Adaptive beamforming methods is a class of high resolution techniques that can improve the performance of imaging systems. Adaptive beamformers offer increased resolution over conventional methods as they take the recorded wavefield into account to find optimal beampatterns. These techniques have been exploited in fields like radar, sonar and seismology. In medical ultrasound, delay-and-sum beamforming is still the method of choice. This is a simple and robust method which suits the real-time requirements of such imaging systems. However, it offers limited resolution and sidelobe suppression compared to adaptive methods.

The main goal of this project has been to study adaptive beamforming techniques for medical ultrasound imaging. Active systems, such as ultrasound, present specific challenges for these methods. Four papers concerning different aspects of adaptive beamforming for medical imaging have been included in this dissertation. We have also included a paper concerning adaptive beamforming in a single snapshot context, and a paper on blind source separation.

The first three papers concern the "minimum variance" beamformer. In the first paper we present how the method can be applied to medical ultrasound. We demonstrate increased resolution and suppression of sidelobes, both on simulated and experimental RF data. We also evaluate the robustness of the method, and show that increased robustness can be achieved by simple means.

In the second paper, we investigate how the estimate of the spatial covariance matrix, used by the minimum variance beamformer, affects the statistics of speckle patterns. We show that an implementation based on a single snapshot of the wavefield gives very different speckle statistics compared to delay-and-sum. By averaging in depth, similar speckle as delay-and-sum is achieved, while the resolution of the method is retained.

We exploit the high-resolution properties of the minimum variance beamformer in the third paper, and show that it can be used to decrease transducer size, increase frame rates or give higher penetration without sacrificing image quality compared to delay-and-sum.

The minimum variance beamformer is significantly more complex than conventional methods, which is one of the reasons why it is not used in

medical ultrasound systems. In the fourth paper we present a simplified adaptive beamformer, which better suits the real-time requirements of medical ultrasound. The method requires only a fraction of the number of computations compared to a full adaptive beamformer, but still gives significant improvements compared to delay-and-sum.

In the fifth paper we present a unifying framework to analyze the minimum variance beamformer in a single snapshot context. We show that implementations based on single snapshots may suffer from signal cancellation, a well-known phenomenon when sources are correlated. The framework allows us to construct an optimization which completely eliminates signal cancellation.

The last paper concerns the field of blind source separation. Blind source separation is a class of methods that can extract source information when we only have observations of mixtures of the sources available. We present a new approach to separation of convolutive mixtures by preprocessing the data using an array processing technique. We apply the method to the so-called "cocktail-party problem".

List of publications

This dissertation includes the following six papers, referred to in the text by their Roman numerals (I-VI).

I J.-F. Synnevåg, A. Austeng, and S. Holm, "Adaptive Beamforming Applied to Medical Ultrasound Imaging", in *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 54, no. 8, pp. 1606-1613, August 2007

II J.-F. Synnevåg, C. I. C. Nilsen, and S. Holm, "Speckle Statistics in Adaptive Beamforming", in *Proc. IEEE Ultrasonics Symposium, pp. 1545-1548, October 2007*

III J.-F. Synnevåg, A. Austeng, and S. Holm, "Benefits of Minimum Variance Beamforming in Medical Ultrasound Imaging", to appear in *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*

IV J.-F. Synnevåg, A. Austeng, and S. Holm, "A Low Complexity Data-Dependent Beamformer", submitted to *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*

V J.-F. Synnevåg, and A. F. C. Jensen, "On Single Snapshot Minimum Variance Beamforming", submitted to *Signal Processing*

VI J.-F. Synnevåg, and T. Dahl, "Blind Source Separation of Convolutive Mixtures Using Spatially Resampled Observations", in *Proc. 14th European Signal Processing Conference*, September 2006

Related publications

The following papers are related to the included papers, and are referred to in the text by their Roman numerals.

VII J.-F. Synnevåg, A. Austeng, and S. Holm, "Minimum Variance Adaptive Beamforming Applied to Medical Ultrasound Imaging", in *Proc. IEEE Ultrasonics Symposium*, vol. 2, pp. 1199-1202, September 2005

VIII J.-F. Synnevåg, A. Austeng, and S. Holm, "High Frame-Rate and High-Resolution Medical Imaging Using Adaptive Beamforming", in *Proc. IEEE Ultrasonics Symposium*, pp. 2164-2167, October 2006

IX A. Austeng, T. Bjastad, J.-F. Synnevåg, S.-E. Masoy, H. Torp, and S. Holm, "Sensitivity of Minimum Variance Beamforming to Tissue Aberrations", to appear in *Proc. IEEE Ultrasonics Symposium*, November 2008

X J.-F. Synnevåg, S. Holm, and A. Austeng, "A Low Complexity Data-Dependent Beamformer", to appear in *Proc. IEEE Ultrasonics Symposium*, November 2008

XI S. Holm, J.-F. Synnevåg, and A. Austeng, "Capon Beamforming for Active Ultrasound Imaging Systems", to appear in *Proc. IEEE 13th DSP Workshop*, January 2009

Contents

Introduction

This dissertation consists of an introduction and six papers. The purpose of the introduction is to give relevant background information to the included papers, review the state-of-the-art, and discuss the contributions of this dissertation. We have also reviewed some related methods that we have not applied directly in our work. These references may, however, be relevant for further research.

Papers I-IV concerns adaptive beamforming in medical ultrasound imaging. The term *adaptive beamforming* is somewhat confusing in this field. In the array processing literature, adaptive beamforming refers to methods that use the recorded data to find optimal weighting functions for sensor arrays. In medical ultrasound, adaptive beamforming also refers to techniques for phase aberration correction. We stress that we are not concerned with phase aberration correction in our work. In Paper IX we have investigated the sensitivity of the minimum variance beamformer to tissue aberrations, but we have made no attempt to correct for them.

The main theme of the dissertation is adaptive beamforming for medical ultrasound imaging, but we have also included two papers that are not directly targeted at ultrasound imaging. Paper V concerns adaptive beamforming in general. Paper VI concerns *blind source separation* (BSS), which is another class of adaptive methods. Also, we believe the methods presented in Paper I and IV can be applied to other imaging systems as well.

This chapter is organized as follows: In Sections 1 and 2 we give brief introductions to propagating waves and beamforming. In Section 3 we review the minimum variance beamformer, and describe a few other adaptive methods. We give a short introduction to medical ultrasound in Section 4 and explain how we have applied the minimum variance beamformer to this imaging technique. In Section 5 we review related work in medical ultrasound. We give a brief introduction to BSS and *independent component analysis* in Section 6, which is relevant for Paper VI. We also relate BSS to adaptive beamforming. In Section 7 we give a summary of the included papers. We discuss our contributions in Section 8 and draw conclusions in Section 9.

1 Propagating waves

In array signal processing we are concerned with extracting information from signals propagating as waves. The propagation is described by the wave equation for the appropriate medium and boundary conditions. The wave equation describing propagation in a homogeneous medium is given by:

$$\frac{\delta^2 s}{\delta x^2} + \frac{\delta^2 s}{\delta y^2} + \frac{\delta^2 s}{\delta z^2} = \frac{1}{c}\frac{\delta^2 s}{\delta t^2}, \tag{1}$$

where $s(x, y, z, t)$ represents a scalar field, (x, y, z) represents a spatial location, and t is time. In acoustics, $s(x, y, z, t)$ represents sound pressure in time and space, and c is the propagation speed of the medium. We present two solutions to the equation which are important for this work: the plane wave and the spherical wave. The (monochromatic) plane wave solution is given by:

$$s(\mathbf{x}, t) = Ae^{j(\omega t - \mathbf{kx})}, \tag{2}$$

where A is the amplitude of the propagating signal, $\omega/2\pi$ is the frequency, $\mathbf{x} = (x, y, z)$, $\mathbf{k}$ is the wavenumber vector (k_x, k_y, k_z), and $k = |\mathbf{k}| = \omega/c$ is the wavenumber. It is called a plane wave because the wavefronts form planes in three-dimensional space.

By using an array of sensors we can observe a wavefield in both time and space. In this thesis we only consider linear arrays, such that we only sample the wavefield in one spatial dimension. A measurement of a single propagating wave can be described as:

$$\begin{aligned} x_m(t) &= f(\bar{x}_m, 0, 0, t) && (3)\\ &= Ae^{j(\omega t - k_x \bar{x}_m)} && (4) \end{aligned}$$

where $x_m(t)$ is the measured signal from the mth sensor, and $\bar{x}_m$ is the x-component of sensor location. We can describe the measurements in vector form as:

$$\begin{aligned} \mathbf{X}(t) &= \begin{bmatrix} x_0(t) \\ x_1(t) \\ \vdots \\ x_{M-1}(t) \end{bmatrix} && (5)\\ &= Ae^{j\omega t} \cdot \mathbf{a}(k_x), && (6) \end{aligned}$$

where $\mathbf{a}$ is the *steering vector* given by

$$\mathbf{a}(k_x) = \begin{bmatrix} e^{-jk_x\bar{x}_0} \\ e^{-jk_x\bar{x}_1} \\ \vdots \\ e^{-jk_x\bar{x}_{M-1}} \end{bmatrix}. \tag{7}$$

Another solution to the wave equation is the spherical wave, which is given by:

$$s(r,t) = \frac{A}{r} e^{j(\omega t - kr)}, \tag{8}$$

where r is the radius in spherical coordinates, and k is the wavenumber. It can be interpreted as a wave traveling outwards from (or inwards to) the origin in all directions. This solution is important in medical ultrasound because we consider every point in the image to be a source of spherical radiation. If we observe a spherical wave using an array of sensors with finite extent, and the extent is small compared to the distance to the source, the observations can be described by a plane wave. The source of the spherical wave is then in the *far-field* of the array.

Both the plane- and the spherical wave solution describes the propagation of a monochromatic wave. But because the wave equation is linear, and an arbitrary physical signal can be described as a sum of complex exponentials, we can use these solutions to describe how arbitrary signals propagate.

2 Beamforming

Methods to focus an array of sensors towards a specific direction or point in space are known as *beamforming*. A simple, yet powerful, method is known as delay-and-sum (DAS) beamforming. The sensor outputs are delayed and summed such that signal components coming from a specific direction are reinforced with respect to noise and signals coming from other directions. The general definition of a delay-and-sum (DAS) beamformer is [1]:

$$z(t) = \sum_{m=0}^{M-1} w_m x_m(t - \Delta_m) \tag{9}$$

where $z(t)$ is the output, M is the number of channels, w_m is (complex) weight m, $x_m(t)$ is the output of channel m and Δ_m is the delay applied to channel m. Delays are applied to bring all signals originating from the focal point in phase, and the weights are used to control rejection of off-axis interference. The purpose of beamforming is twofold:

- Increase the signal-to-noise ratio (SNR) compared to a single sensor.
- Steer or focus the sensor array towards a signal coming from a specific direction or point.

We can show that provided all signals are in phase and the noise is spatially white, the increase in SNR by using an array of sensors is equal to the number of sensors, M. This is referred to as the *array gain*. The second objective of beamforming is to steer or focus the array in a particular direction or towards a specific point in space, with the goal of estimating a signal coming from a direction (or point). Focusing requires near-field imaging. If the wavefield consists of many signals, the capabilities to reject other signals determines how well the desired signal is estimated. The rejection capabilities is given by the beampattern, which is defined below.

2.1 Beampattern

We assume a monochromatic plane wave, $f(\mathbf{x}, t) = Ae^{j(\omega t - \mathbf{k}^\circ \mathbf{x})}$ being observed by a linear array of M sensors. The output of a delay-and-sum beamformer can be described as:

$$z(t) = \sum_{m=0}^{M-1} w_m f(\bar{x}_m, t - \Delta_m) \tag{10}$$

where the delays are of the form $\Delta_m = -k_x \bar{x}_m / \omega$, giving:

$$z(t) \quad = \quad \sum_{m=0}^{M-1} w_m A e^{j(\omega t - k_x^\circ \bar{x}_m + k_x \bar{x}_m)} \tag{11}$$

$$= Ae^{j\omega t}\sum_{m=0}^{M-1} w_m e^{j((k_x-k_x^\circ)\bar{x}_m)} \quad (12)$$

$$= Ae^{j\omega t}W(k_x - k_x^\circ), \quad (13)$$

where

$$W(k) = \sum_{m=0}^{M-1} w_m e^{jk\bar{x}_m}. \quad (14)$$

The function, $W(k_x - k_x^\circ)$, is called the *beampattern* [1]. We see that the beampattern describes how a monochromatic signal, $Ae^{j\omega t}$, propagating in a direction given by k_x° is attenuated by a delay-and-sum beamformer steered towards k_x. Hence, we can use the beampattern to analyze how any frequency component of a broadband propagating signal is attenuated due to the array processing. The beampattern will normally have its maximum for $k_x^\circ = k_x$, which means that the array has maximum sensitivity in the steering direction. We will see in Papers III and IV that this is not always the case for adaptive beamformers.

If we include phase delays in the weights, $\bar{w}_m = w_m e^{-jk_x\bar{x}_m}$, we can express the beampattern as:

$$W(k_x^\circ) = \mathbf{w}^H\mathbf{a}(k_x^\circ), \quad (15)$$

and the output of the delay-and-sum beamformer for narrowband signals as:

$$z(t) = \mathbf{w}^H\mathbf{X}(t) \quad (16)$$

where

$$\mathbf{w} = \begin{bmatrix} \bar{w}_0 \\ \bar{w}_1 \\ \vdots \\ \bar{w}_{M-1} \end{bmatrix}, \quad (17)$$

and

$$\mathbf{X}(t) = \begin{bmatrix} x_0(t) \\ x_1(t) \\ \vdots \\ x_{M-1}(t) \end{bmatrix}. \quad (18)$$

As we have assumed a wavefield consisting of a plane wave, (15) is the far-field beampattern. In Papers III and IV we have used the far-field beampattern to illustrate the difference between deterministic and adaptive beamformers, and to illustrate the effect of different window functions. We should note that this is only an approximation when describing the performance of a near-field beamformer. It is possibly to derive the near-field beampattern by the same exercise as for the far-field case. By considering a wavefield consisting of a

spherical wave, $f(r,t) = A/r \cdot e^{j(\omega t - kr)}$, we can get the expression for the near-field beampattern [1]:

$$W(k, \mathbf{x}, \mathbf{x}^{\circ}) = \sum_{m=0}^{M-1} w_m \frac{r^{\circ}}{r_m^{\circ}} e^{jk((r^{\circ}-r)-(r_m^{\circ}-r_m))}, \tag{19}$$

where $\mathbf{x}$ is the focal point of the beamformer, $\mathbf{x}^{\circ}$ is the location of the source, r and r_m is the distance from the focal point to the center of the array and sensor m, respectively, and r° and r_m° is the distance from the actual source location to the center of the array and sensor m, respectively .

3 Minimum Variance Beamforming

Early works on adaptive beamforming include the papers of Bryn [2], Capon [3], Widrow [4] and Applebaum [5]. Bryn was concerned with optimal signal detection from three-dimensional arrays. Capon developed a method to improve localization of earthquakes using seismic arrays. Applebaum presented a method to adaptively optimize the signal-to-noise ratio of an antenna array, which has later been termed the Applebaum or Howells-Applebaum array in the radar community. Widrow presented a method to automatically adjust array weights based on the least-mean-squares (LMS) algorithm.

Capon's method is known as the Capon beamformer, the minimum variance (MV) beamformer, or the minimum variance distortion-less response (MVDR) beamformer in the literature. We have used the name "minimum variance" in our work. The method finds an optimal set of weights based on the recorded data, hence the name *adaptive* beamformer. The goal is to minimize both off-axis interference and noise by a given optimization criterion. The idea is simple: Find the weights that minimize the variance of $z(t)$ in (9), with the constraint that the signal that originates from the steering direction or focal point of the array is passed with unit gain. Hence, the signal in focus is passed undistorted, while the channels are combined such that interference and noise is minimized. First, note that we can express the variance of the output of the narrowband delay-and-sum beamformer in (16) as:

$$\begin{aligned} E\{|z(t)|^2\} &= E\{|\mathbf{w}^H\mathbf{X}(t)|^2\} && (20)\\ &= \mathbf{w}^H\mathbf{R}\mathbf{w}, && (21)\end{aligned}$$

where $\mathbf{R} = E\{\mathbf{X}\mathbf{X}^H\}$ is the spatial covariance matrix. We can formulate the minimum variance beamformer optimization as [3]:

$$\begin{aligned} \min_{\mathbf{w}} \quad & \mathbf{w}^H\mathbf{R}\mathbf{w} \\ \text{subject to} \quad & \mathbf{w}^H\mathbf{a}(k_x) = 1, \end{aligned} \tag{22}$$

where $\mathbf{a}(k_x)$ is the steering vector. As mentioned in Section 1, $\mathbf{a}(k_x)$ represents a propagation direction, thus the constraint forces the beampattern of the solution to have gain one in the steering direction, meaning that $W(k_x) = 1$. The solution to (22) is:

$$\mathbf{w} = \frac{\mathbf{R}^{-1}\mathbf{a}(k_x)}{\mathbf{a}(k_x)^H\mathbf{R}^{-1}\mathbf{a}(k_x)}. \tag{23}$$

Fig. 1(a) shows an example of the output of a MV beamformer operating on a linear array with 32 elements and $\lambda/2$ spacing. Two signals were present in the wavefield (with angles-of-arrival of 0 and 5 degrees), and the SNR per channel was 20 dB. The dashed line shows the corresponding DAS beamformer output

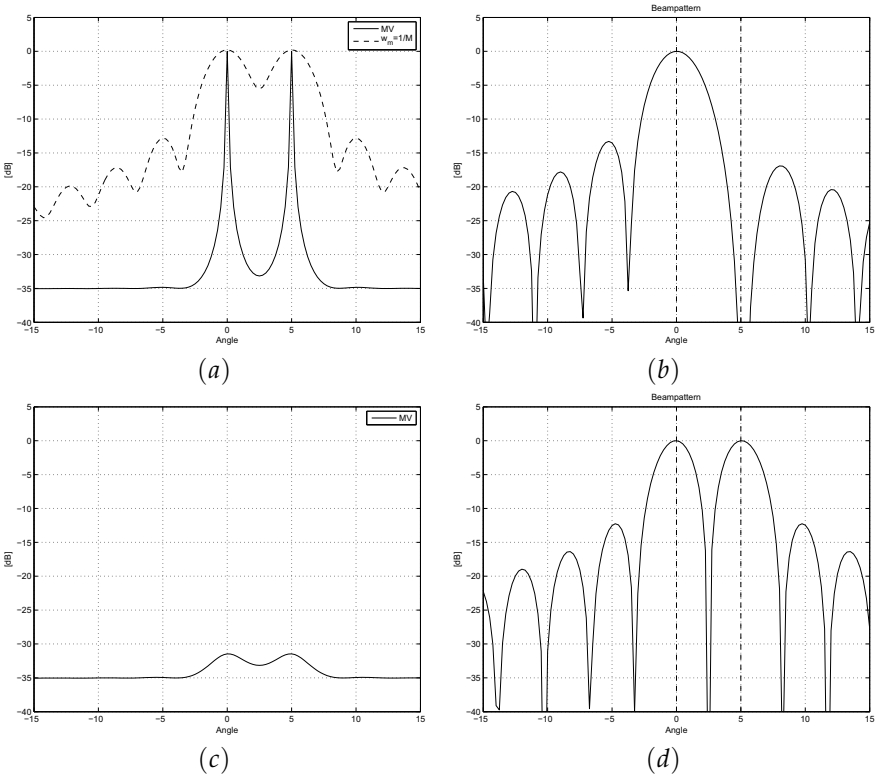


Fig. 1: Examples of MV beamformer performance (a) Responses for uncorrelated sources (dashed lines shows DAS for $w_m = 1/M$) (b) Beampattern when steering towards $\theta = 0$. (c), and (d) shows corresponding plots for correlated sources.

for $w_m = 1/M$. Fig. 1(b) shows the beampattern when we focus towards $\theta = 0$. We see that the gain is one (0 dB) in the steering direction, and that a null is formed at the angle-of-arrival of the interfering signal.

The performance of the MV beamformer depends on the number of elements, the SNR, and the number of signals present in the wavefield. The more sensors, the more degrees of freedom the MV beamformer has to form nulls in directions of interfering signals. Poor SNR will constrain the achievable resolution. Noise (if white) gives a contribution to the spatial covariance matrix which is proportional to the identity matrix, and can be interpreted as a *diagonal loading* term. The noise will constrain the norm of the weight vector and decrease the resolution, but at the same time give a more robust solution. We address diagonal loading and robust MV beamforming in Section 3.2.

3.1 Signal Cancellation

Signal correlation (or coherence) may lead to signal cancellation in the MV beamformer [6], which means that the amplitude of the signal in focus may be underestimated. To see why this can occur, we assume two monochromatic signals, $s_0(t)$ and $s_1(t)$ that propagate in two different directions given by k_0 and k_1. We assume that we focus towards s_0. For simplicity we assume noise-free recordings. The measurement vector can be described as:

$$\mathbf{X}(t) = s_0(t) \cdot \mathbf{a}(k_0) + s_1(t) \cdot \mathbf{a}(k_1) \tag{24}$$

The output of a beamformer operating on these measurements is:

$$\begin{aligned} z(t) &= \mathbf{w}^H \mathbf{X}(t) && (25)\\ &= s_0(t)\mathbf{w}^H\mathbf{a}(k_0) + s_1(t)\mathbf{w}^H\mathbf{a}(k_1) && (26)\\ &= W(k_0)s_0(t) + W(k_1)s_1(t) && (27)\\ &= s_0(t) + W(k_1)s_1(t) && (28) \end{aligned}$$

As we steer towards s_0, the constraint in (22) forces $W(k_0) = 1$. The variance of $z(t)$ becomes:

$$\begin{aligned} E\{|z(t)|^2\} &= E\left\{|s_0(t) + W(k_1)s_1(t)|^2\right\} && (29)\\ &= E\left\{|s_0(t)|^2\right\} + 2\mathrm{Re}\left\{E\left\{s_0(t)s_1^*(t)\right\}\right\} + W(k)^2 E\left\{|s_1(t)|^2\right\}, && (30) \end{aligned}$$

where $\mathrm{Re}\{\cdot\}$ denotes the real part. To minimize interference, the MV solution should form a null in the direction of s_1, *i.e.* $W(k_1) = 0$. The variance of the output of the beamformer in (30) would then be equal to the variance of $s_0(t)$. However, we see that if $s_0(t)$ and $s_1(t)$ are correlated, the second term in (30) may be non-zero and negative. Hence, there may exist a solution in which the variance is smaller than $E\{|s_0(t)|^2\}$. ***E.g.*** if $E\{|s_0(t)|^2\}$, $E\{|s_1(t)|^2\}$ and the cross-correlation between s_0(t) and $s_1(t)$ is equal to one, the variance of $z(t)$ becomes zero if $W(k_1) = -1$. The MV beamformer will favour $W(k_1) = -1$ over $W(k_1) = 0$, as the output of the beamformer has a smaller variance. In this situation signal cancellation has occurred. We can make a similar argument for broadband sources, except that the second term will depend on the cross-correlation at different lags, and the third term in (30) will include a filtered version of $s_1(t)$. Fig. 1(c) shows examples of the output of a MV beamformer when the source signals are correlated. The array parameters were the same as in Fig. 1(a). We see that signal cancellation occurs when trying to estimate the amplitude of the two signals. Fig. 1(d) shows the beampattern of the MV weights when steering towards 0 degrees. We see that the beampattern has two peaks corresponding to the angles of arrival of the sources. Signal cancellation occurs because the phase at the location of the second source is shifted.

In Paper V we show that if we apply the MV beamformer using only a single snapshot of the wavefield, signal cancellation may occur as well. In both cases,

this effect can be mitigated by spatial smoothing (subarray averaging) [7, 8], which we explain in Section 4.2.

3.2 Robust Minimum Variance Beamforming

A weakness of the MV beamformer is robustness towards uncertainties in the steering vector or gain differences across the array. If the steering vector in (23) is slightly wrong, e.g. because the assumed propagation velocity is wrong, the beamformer will try to suppress the signal of interest, as the only requirement is that a signal coming from the steering direction is passed with zero dB gain. There are several ways to make the beamformer more robust, and we will review a few of them here.

A natural extension is to force zero dB gain in directions close to the steering angle, giving extra constraints in the optimization. We may solve:

$$\begin{aligned} \min_{\mathbf{w}} \quad & \mathbf{w}^H \mathbf{R} \mathbf{w} \\ \text{subject to} \quad & \mathbf{w}^H \mathbf{C} = \mathbf{f}, \end{aligned} \tag{31}$$

where $\mathbf{C} = \begin{bmatrix} \mathbf{a}(k_x - \delta) & \mathbf{a}(k_x) & \mathbf{a}(k_x + \delta) \end{bmatrix}$ and $\mathbf{f} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$. Each constraint reduces the degrees of freedom by one. This approach may severely diminish the resolution. Another approach is to constrain the derivative of the beampattern, e.g. by setting:

$$\mathbf{C} = \begin{bmatrix} \mathbf{a}(k_x) & \frac{d\mathbf{a}(k_x)}{dk_x} \end{bmatrix}, \tag{32}$$

and $\mathbf{f} = \begin{bmatrix} 1 & 0 \end{bmatrix}$ in (31). Forcing the derivative to be zero, means that the beampattern will have its maximum in $\mathbf{a}(k_x)$, such that the mainlobe will be more or less symmetric around the steering angle. This may reduce the ability to define edges. We show in Papers III and IV that asymmetric mainlobes in the beampattern are beneficial for edge definition.

Several of the proposed robust methods belong to the class of *diagonal loading*. Diagonal loading means that a small term is added to the diagonal of the covariance matrix in (23) prior to inversion, replacing $\mathbf{R}$ by $\mathbf{R} + \epsilon \mathbf{I}$. Diagonal loading can seen as "adding" white noise to the data before solving the MV optimization (note that the covariance matrix of a wavefield consisting of white noise is a diagonal matrix, with the variance of the noise on the diagonal). By increasing the noise level artificially, the MV solution approaches that of a rectangular weighted DAS, which can be seen if we insert a diagonal matrix into (23). We could define ϵ up front, but it may be difficult to find a representative value as the signal statistics change across the image. We have chosen a simple approach, in which ϵ is proportional to the trace of $\mathbf{R}$:

$$\epsilon = \frac{\Delta}{L} \mathrm{tr}\{\mathbf{R}\}, \tag{33}$$

where L is the dimension of $\mathbf{R}$. In that way, ϵ is proportional to the average variance of the received signals.

To avoid unstable solutions, we may constrain the norm of the weight vector. This is achieved by adding the constraint $||\mathbf{w}||^2 \leq \beta$ in (22). It turns out that solution to the optimization problem is in the form of diagonal loading [9]. With this approach it may also be difficult to determine the value of β.

It has also been proposed to model the uncertainties in the steeringvector explicitly in the optimization. One example is to solve the optimization [10]:

$$\min_{\mathbf{w}} \quad \mathbf{w}^H \mathbf{R} \mathbf{w} \tag{34}$$

$$\text{subject to} \quad \text{Re}\{\mathbf{w}^H \mathbf{a}\} \geq 1 \quad \forall \mathbf{a} \in \varepsilon, \tag{35}$$

where $\text{Re}\{\cdot\}$ denotes the real part, and ε is an ellipsoid that covers the possible values of $\mathbf{a}$ due to uncertainties about the true steering vector. Several similar modifications to the optimization have been proposed [11–13]. These approaches belong to a class of "extended" diagonal loading methods.

In Paper I we show that by decreasing the length of the subarrays using subarray averaging, we also obtain a more robust MV beamformer. We explain subarray averaging in Section 4.2.

3.3 Other Adaptive Beamformers

A natural extension to the MV beamformer is to optimize a set of channel filters instead of just array weights. The minimum variance optimization problem using FIR filters was formulated by Frost and is often referred to as the *Frost beamformer* [14]. By using channel filters we can form nulls along a line in frequency-wavenumber space, corresponding to a broadband propagating signal. Hence, the Frost beamformer is often referred to as a broadband beamformer. As delays cannot be applied using complex exponentionals in the form of a steering vector, the first step of the Frost beamformer is to delay the measurements to steer in a particular direction. The optimization can then be formulated as:

$$\begin{aligned} \min_{h_0,\ldots,h_{M-1}} \quad & E\left\{ \left| \textstyle\sum_m h_m[n] * x[n - \Delta_m] \right|^2 \right\} \\ \text{subject to} \quad & \textstyle\sum_m h_m[n] = \delta[n], \end{aligned} \tag{36}$$

where $h_1, \ldots, h_{M-1}$ are FIR filters, $(*)$ denotes convolution, and $\delta[n]$ is the Kroenecker delta function (we have now assumed that we operate on sampled signals in the temporal dimension). The constraint forcing the sum of the FIR filters to be the delta function, ensures that the signal in focus is passed undistorted. It may be replaced by a different impulse response if temporal filtering of the beamformer output is desirable. Frost formulated (36) in vector form. He also suggested to solve it iteratively using the Lagrange multiplier as cost function, to avoid matrix inversion. Griffiths and Jim [15] later reformulated the problem in the form of a generalized sidelobe canceller.

The signal cancellation effects in the minimum variance beamformer has led to several proposals on how to mitigate it. In [6] Widrow *et al.* describes this phenomena and suggests a cure in the form of the *Duvall beamformer*, named after the second author. The idea behind this beamformer is that if the signal we steer towards is not present in the data, it cannot be cancelled. The first step in the Duvall beamformer is to apply delays to steer towards the signal of interest. A new set of measurements is formed by subtracting neighboring sensors, attempting to subtract the signal of interest. The MV optimization problem is then solved with the "signal-free" measurements. It turns out, however, that this beamformer can only handle two correlated sources.

In [16] Bresler *et al.* suggest approaches that completely eliminate signal cancellation in the presence of correlated signals. They present three distinct beamformers, corresponding to different optimality criteria, by exploiting an underlying narrowband signal model. To avoid signal cancellation they use a maximum likelihood approach to find the signal-free covariance matrix. However, as this step of the method relies on narrowband model, the method is not directly transferable to the broadband case, unless frequencies are treated separately.

The amplitude and phase estimator (APES) is an adaptive spectral estimator (or beamformer) suggested by Li and Stoica [17]. It can be formulated as [18]:

$$\begin{aligned} \min_{\mathbf{w},\alpha} \quad & \frac{1}{M-L+1} \sum_{l=0}^{M-L+1} \left| \mathbf{w}^H \mathbf{X}_l - \alpha e^{jk_x \bar{x}_l} \right|^2 \\ \text{subject to} \quad & \mathbf{w}^H \mathbf{a} = 1. \end{aligned} \tag{37}$$

The optimization is very similar to the MV beamformer using subarray averaging, except that we now force the filter output to be as close as possible to a sinusoid with wavenumber k_x. It can be shown that (37) can be solved by the optimization:

$$\begin{aligned} \min_{\mathbf{w}} \quad & \mathbf{w}^H \mathbf{Q} \mathbf{w} \\ \text{subject to} \quad & \mathbf{w}^H \mathbf{a}(k_x) = 1, \end{aligned} \tag{38}$$

where $\mathbf{Q} = \hat{\mathbf{R}} - \mathbf{g}(k_x)\mathbf{g}(k_x)^H$ and

$$\mathbf{g}(k_x) = \sum_{l=0}^{M-L+1} \mathbf{X}_l e^{jk_x \bar{x}_l}. \tag{39}$$

We can interpret $\mathbf{Q}$ as an estimate of the signal-free covariance matrix, where the subtraction is given by the Fourier estimate of the signal in focus. The idea is somewhat similar to the Duvall beamformer, but the method works in the general case. The APES method gives better amplitude estimates than MV, but has lower resolution. The *adaptive single snapshot beamformer* for noisy data, by Ali and Schreib [19], solves a very similar optimization problem as APES.

3.4 Other High-Resolution Methods

We will briefly mention a few high-resolution methods which do not fall under the category of beamforming, but some of them can be used for imaging. To further improve the resolution of the MV beamformer, methods like the *eigenvector method*, *MUSIC* and *ESPRIT* have been suggested [20–22]. In their original form, they are, however, only direction-of-arrival estimators and do not estimate the amplitude of the propagating signals. Hence, they cannot be used to estimate reflectivity. Also, they require an estimate of the number of signals present in the wavefield, and rely on a narrowband signal model. Some preprocessing schemes have, however, been suggested for broadband signals, including use of focusing filters and spatial resampling [23, 24]. We have used the latter preprocessing technique in Paper VI for the problem of blind source separation.

More promising methods for medical ultrasound is the minimum mean squared error (MMSE) and maximum a posteriori (MAP) estimation techniques applied to non-destructive testing with ultrasound by Lingvall and Olofsson [25]. The image is divided into a grid of possible scatterers, and the acoustic field in each of the grid points is modeled. The most likely scattering amplitudes, given the model and the sampled data, are estimated. Hence, the methods rely on an accurate model of the imaging system. A similar method (TONE), which is also based on a MAP estimator, has been applied to medical ultrasound by Viola and Walker [26]. This is a near-field, broadband extension of the narrowband method SPOC [27]. These methods require a lot of computations, and real-time implementations may be unrealistic for a long time.

4 Medical Ultrasound Imaging

The use of ultrasound as a clinical tool dates back to the 1950s. The first developments were *amplitude mode* (A-mode) ultrasound, which used a single transducer element acting as both a transmitter and a receiver. The measured reflections were displayed as time series, corresponding to a one-dimensional scan. Later, *brightness mode* (B-mode) imaging was developed, where the image points were displayed in gray scale with brightness proportional to the amplitude of the reflections, and a two-dimensional cross section of the body was shown. Through technological advances like real-time imaging and electronic focus, ultrasound gained widespread use as a diagnostic tool in the 1970s. The latest technology is three-dimensional (3D) imaging, using two-dimensional (2D) transducer arrays, showing objects in three dimensions rather than just a cross-section. See [28] for a historical review of the developments in ultrasound imaging.

Medical ultrasound 1D arrays typically consists of 64 to 256 elements, and operates at frequencies from 2 to 10 MHz. Images are formed by pulse-echo imaging, meaning that an active source transmits an ultrasound beam and the reflected echoes are processed to form an image. DAS is the standard beamforming technique used. Images may be formed by various scan types: In a *linear scan* only a group of transducer elements is active at one time, and the image is formed by activating a new set elements across the transducer array. A *sector scan* uses 100-200 beams which are steered in different angles. The beams may be steered mechanically or electronically by delaying the signals transmitted from different elements. With a sector scan we may obtain a large image using a small aperture. Fig. 2 shows examples of sector scan images. The *curved scan* combines the simple electronics of the linear scan with the wide image of the sector scan, and uses a curved array where only a portion the aperture is active at one time.

The standard display is B-mode. Other modes include *motion mode* (M-mode), which displays the echos from one scan line over time, *color flow*, in which estimates of the blood flow velocities are displayed, *harmonic imaging*, which images harmonics generated by non-linear propagation, and 3D imaging, which has already been mentioned. See [29, 30] for introductions to medical ultrasound imaging.

In this project we have only considered linear 1D arrays forming B-mode images using sector scans, where the beams have been steered electronically.

4.1 Broadband Near-Field Minimum Variance Beamforming

In Section 3 we presented a narrowband, far-field minimum variance beamformer. In medical ultrasound imaging broadband pulses are transmitted and

Jonas, week 9

Janine, week 18

Fig. 2: Examples of B-mode images, obtained with sector scans.

the beamformer operates in the extreme near-field of the transducer, so we cannot use these equations directly. As delays cannot simply be described with phase delays as for the steering vector in (22), we first delay the data, similar to the first step in the DAS beamformer, transforming the data from near- to far-field. The delayed measurement vector is of the form:

$$\mathbf{X}[n] = \begin{bmatrix} x_0\left[n - \Delta_0[n]\right] \\ x_1\left[n - \Delta_1[n]\right] \\ \vdots \\ x_{M-1}\left[n - \Delta_{M-1}[n]\right] \end{bmatrix}. \tag{40}$$

Note that the delays, $\Delta[n]$ are time-dependent, as dynamic focusing is used on reception. We can express the beamformer in vector form as:

$$z(t) = \mathbf{w}^H[n]\mathbf{X}[n], \tag{41}$$

where $\mathbf{w}^H[n] = \begin{bmatrix} w_0[n] & w_1[n] & \cdots & w_{M-1}[n] \end{bmatrix}$. Note that the weights are also time-dependent. After delays, signals propagating from the focal point of the receive beam can be described by the steeringvector, $\mathbf{a} = \begin{bmatrix} 1 & \ldots & 1 \end{bmatrix}^T$. We can then solve (22) using this steering vector as constraint, together with an estimate of the spatial covariance matrix. The short pulses and the rapidly varying signal statistics present a challenge for the estimation. We address estimation of this matrix in the next section.

The assumption about uncorrelated signals is violated in ultrasound imaging, due to presence of an active source. The returned echos from different scatterers will be correlated as they originate from the same source signal. Hence, signal cancellation may occur. However, the fact that the source is focused, is an advantage compared to passive systems. Most of the reflected energy will appear around the direction of the scan line, leaving large regions in the view of the array where high sidelobes may be allowed.

The computational overhead of the MV beamformer presents another challenge to a practical imaging system, due to the real-time requirements. We discuss the real-time implications in Paper III and address them using a simplified adaptive beamformer in Paper IV.

4.2 Estimation of the Spatial Covariance Matrix

In practice the spatial covariance matrix in (22) must be estimated. The true covariance matrix is replaced by an estimate of the form:

$$\hat{\mathbf{R}} = \frac{1}{N}\sum_l \mathbf{X}_l\mathbf{X}_l^H, \tag{42}$$

where N is the number of snapshots. For ergodic signals, where the signal statistics are independent of time, we may average over many temporal samples,

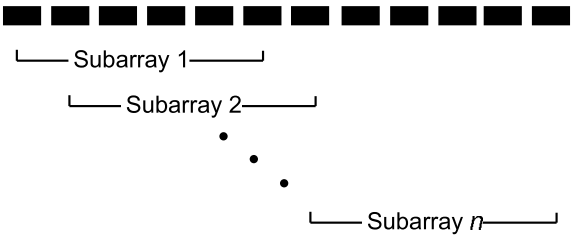


Fig. 3: Illustration of division into subarrays. The array is divided into (overlapping) subarrays of length L, and the spatial covariance matrices of the subarrays are averaged.

using $\mathbf{X}_l = \mathbf{X}[l]$ in (42) (where $\mathbf{X}[l]$ is the lth temporal snapshot of the wavefield). In medical ultrasound, the transmitted pulses are short and the signal statistics may change rapidly. We may instead average in the spatial dimension by dividing the array in (overlapping) subarrays [7, 8]. We use $\mathbf{X}_l = \mathbf{X}_l[n]$ in (42), where $\mathbf{X}_l[n]$ represents the measurements from subarray l at temporal sample n:

$$\mathbf{X}_l[n] = \begin{bmatrix} x_l[n - \Delta_l[n]] \\ x_{l+1}[n - \Delta_{l+1}[n]] \\ \vdots \\ x_{l+L-1}[n - \Delta_{l+L-1}[n]] \end{bmatrix}, \tag{43}$$

where L is the length of the subarrays. The division into subarrays is illustrated in Fig. 3. In Paper I we used this estimate, which works well for point scatterers, but may lead to very different speckle statistics compared to DAS beamforming. In Paper II we introduced averaging in the temporal domain (corresponding to depth) as well, giving speckle statistics that are similar to DAS. This corresponds to using:

$$\mathbf{X}_l = \begin{bmatrix} \mathbf{X}_l[n-K] & \cdots & \mathbf{X}_l[n] & \cdots & \mathbf{X}[n+K] \end{bmatrix}, \tag{44}$$

which means that we average over $2K+1$ temporal samples for each subarray.

In medical ultrasound we may average across other dimensions as well. We briefly explain the different options which have been exploited by other authors. Sasso and Cohen-Bacrie [31] suggested to average across consecutive transmit beams. This was also exploited in [32]. As the width of the transmit beam varies with depth, a point in the image may be insonified several times across a scan. The transmit beams will have different focus, but the receive beams are adjusted to focus at a particular point in the image. Hence, we can estimate the covariance matrix using $\mathbf{X}_l = \mathbf{X}_{\theta_l}(p)$, where $\mathbf{X}_{\theta_l}(p)$ represents a transmit beam transmitted in angle θ_l, and p represents the focal point of the receive beam.

We may also average the received data from different frames in the image, setting $\mathbf{X}_l = \mathbf{X}_{F_l}(p)$, where F_l denotes frame l. Target motion may cause

incoherence in data captured in different frames, which may affect the estimate.

Another way to estimate the covariance matrix, which is exploited in [33], is by a synthetic aperture approach. With a single transmit (Tx) element and all receiver (Rx) elements active, we obtain one sample of the wavefield. We then switch to the next Tx element to obtain the next sample, and so on. We estimate the covariance matrix by averaging the collected samples. Again, target motion may be an issue for fast moving objects.

These ways of estimating the covariance matrix may be used in combination with the robust methods presented in Section 3.2. In our work we have applied subarray averaging together with a small diagonal loading term.

5 Adaptive Beamforming in Medical Ultrasound Imaging: State-of-the-Art

In recent years, several authors have investigated use of the minimum variance beamformer or other adaptive beamformers in medical ultrasound imaging. In 2002, Mann and Walker [34] applied the Frost beamformer to a medical ultrasound scenario, where they demonstrated improved definition of a single point scatter and improved contrast in a cyst phantom using experimental RF data. It is important to note that the use of a single point does not reveal the signal cancellation phenomena that may occur in the Frost and MV beamformers. In 2005, Sasso and Cohen-Bacrie [31] applied the minimum variance beamformer to simulated data of a point scatterer close to a cyst. They argued that subarray averaging was required to make the method work, due to correlated echos. With this scheme they demonstrated improved point definition and increased contrast. By averaging over consecutive scan lines, they further improved the results.

In Paper VII we showed some preliminary results from simulated and experimental data of wire targets, and experimental data of a heart phantom. Viola and Walker [35] applied the Frost, Duvall and the adaptive single snapshot beamformer to point targets. They also applied SPOC, which is a high-resolution method mentioned in Section 3.4. The method was further refined in [26, 36], and renamed to *time-domain optimized near-field estimator* (TONE).

Wang *et al.* [33] used a synthetic aperture approach in combination with robust minimum variance beamformers, which they applied to simulated data of point targets and experimental data of point-, cyst- and heart phantoms. They demonstrated improved resolution and contrast, and showed that their robust adaptive beamformer could outperform delay-and-sum using half the transducer size.

In Paper I we demonstrated increased resolution by our implementation of the minimum variance beamformer, and investigated the sensitivity to perturbations in acoustic velocity. In Paper II we studied the differences in speckle statistics using DAS and MV beamformers. In Paper III and Paper VIII we showed that the MV beamformer may give other benefits than simply higher resolution. Without sacrificing image quality, we showed that the transducer size and channel count could be reduced, frame rates could be increased (in combination with parallel beamforming or by reducing the number of focal zones), or a virtual increase in penetration depth could be achieved by lowering the transmit frequency.

Holfort *et al.* presented their frequency domain implementation of the MV beamformer in [37]. They split the received signals in different frequency bands, performed independent MV beamforming per band, and combined them.

This gives $N_f - 1$ times the number of MV weights per sample (where N_f is the number of frequency bands), and thus increases the degrees of freedom, but gives a more complex method. They applied their method to a synthetic aperture beamformer, where they first formed "low-resolution" images using single transmit elements, and then combined the low-resolution images to a high-resolution image.

Vignon and Burcher presented their implementation of the MV beamformer in [32]. The method has similarities with the method in [33] in that the covariance matrix is formed by insonifying the medium with different beams. However, they use focused beams, which compared to the synthetic aperture approach is beneficial for the SNR, and for the generation of harmonics (for harmonic imaging). Also, this method will be less sensitive to target motion. They were the first to show *in vivo* images using adaptive methods in ultrasound imaging, both from a heart and from an abdominal scan.

In Paper IX we investigated the MV beamformer's sensitivity to phase aberrations. We showed that the MV beamformer gave a substantial decrease in mainlobe width without increase in sidelobe level for weak and intermediate aberrations. The performance was equal or better than conventional beamforming for all realistic aberrations.

In Paper IV we showed results from what we have termed the *low complexity adaptive beamformer*, which is based on the minimum variance beamformer, but chooses the apodization functions from a set of predetermined windows. The method is only an approximation to the MV beamformer, but has much lower complexity. It still displays great improvements over DAS beamforming.

6 Blind Source Separation

Blind source separation (BSS) is a statistical method to recover signals or sources from several observations of mixtures of them. The classical example of BSS is the cocktail party problem, in which several people (sources) speak simultaneously in a room, like in a cocktail party. The problem amounts to recovering the different voices (source signals) using several microphones. BSS has practical applications in fields like telecommunications, speech enhancement and bio-medicine. An introduction to the field can be found in [38]. The fact that the separation is "blind" means that no a priori information is available about the mixture, except maybe how many sources there are in the mix. In the original formulation we do not know anything about the position of the sensors either.

The separation is often performed by independent component analysis (ICA), and the terms ICA and BSS are often used interchangeably. There exists several algorithms performing ICA, like FastICA [39] and JADE [40]. We briefly describe the classes of methods within the ICA framework.

6.1 Independent Component Analysis

By the general definition, ICA of the random vector $\mathbf{X}$ consists of finding a linear transform $\mathbf{s} = \mathbf{W}\mathbf{X}$ such that the components in $\mathbf{s} = \begin{bmatrix} s_1 & \cdots & s_n \end{bmatrix}$ are as independent as possible in the sense of maximizing some function $F(s_1, \cdots, s_n)$ that measures independence [41]. ICA assumes a model of the form:

$$\mathbf{X} = \mathbf{A}\mathbf{s}, \tag{45}$$

where $\mathbf{X}$ contains the observations of the mixtures of sources, $\mathbf{A}$ describes the mix and $\mathbf{s}$ are the sources we aim to extract. Hence, we assume an instantaneous mix of the sources. The contrast function, $F(s_1, \ldots, s_n)$, must be some measure of independence and the choice is dependent on statistical properties, convergence, etc.

A successful separation of sources using ICA requires that:

- the sources, s_i, are non-Gaussian (with the possible exception of one)
- the number of of observations are at least as many as the number of sources
- the mixing matrix, $\mathbf{A}$, is of full column rank.

With all assumptions fulfilled, a successful separation by ICA still leaves potential problems: The sources are returned in a random order with an arbitrary scaling, referred to as permutation and scaling inconsistencies.

The ICA model assumes an instantaneous mix as in (45), but the mixes in the case of propagating waves (like the cocktail party problem) are convolutive. The classical approach to transform the model from a convolutive to an

instantaneous mix, is to treat individual frequency bands separately. For narrowband sources we can describe the propagation delays as complex scalars, hence we can solve for each frequency band separately:

$$\mathbf{X}(\omega) = \mathbf{A}(\omega)\mathbf{s}(\omega). \tag{46}$$

We perform the separation for each frequency band and reconstruct the sources by combining the extracted subbands. The permutation and scaling inconsistencies now become a problem: The different subbands for one particular source may be extracted in a random order. Also, the arbitrary scaling of the subbands makes it difficult to reconstruct the signals. In Paper VI we suggest a method to transform the convolutive model to an instantaneous model by a spatial preprocessing method known as spatial resampling [24]. All frequencies of a particular source are transformed onto the same wavenumber, and the original convolutive model becomes instantaneous in the time domain. Hence, we process all frequency bands simultaneously, and avoid the permutation and scaling problems. The separation process is not completely blind anymore, as we have to know the positions of the sensors to perform the spatial resampling. However, it is not uncommon to have this information available.

6.2 Relation to Adaptive Beamforming

The approach to solve the BSS problem is often purely statistical, and the potential benefits of array processing techniques are often ignored. However, some authors have incorporated beamforming techniques to solve the convolutive BSS problem [42, 43]. It has been argued that frequency domain BSS of convolutive mixtures of speech is fundamentally limited by the performance of an adaptive beamformer [44]. To link the two fields, we observe that the separation matrix $\mathbf{W}$ will contain the row vectors $\mathbf{w}_1, \ldots, \mathbf{w}_n$. From a beamforming perspective we can view the rows of $\mathbf{W}$ as window functions in a beamformer. Upon successful completion of ICA, the projections of the observations onto $\mathbf{w}_i$, $s_i = \mathbf{w}_i\mathbf{X}$, will give statistically independent s_i. It means that the beampattern of the vector $\mathbf{w}_i$ will contain zeros in the propagation directions of the other sources, s_j $(i \neq j)$. The gain in the direction of source i will be non-zero (but may be different from one). We see the similarities to adaptive beamforming. Adaptive beamformers use the recorded data to estimate signals coming from specific directions. The separation matrix, $\mathbf{W}$ may be found by stacking the weights of the adaptive beamformer for the propagation directions of $s_1, \ldots, s_n$:

$$\mathbf{W}_{MV} = \begin{bmatrix} \mathbf{w}_1^H \\ \vdots \\ \mathbf{w}_n^H \end{bmatrix}. \tag{47}$$

For uncorrelated sources at high SNR the beampattern of $\mathbf{w}_i^H$ will contain zeros in directions of s_j for all $j \neq i$, and have gain one in direction of s_i. If so, we have extracted the different sources successfully. In contrast to ICA, we only require the sources to be uncorrelated. If solved in the frequency domain, we avoid the scaling and permutation inconsistencies of ICA. However, the separation is no longer blind: We need knowledge about the positions of the sensors for the steering vector in the constraint in (22). Also, we need to know, or estimate, the locations of the different sources. This may be done by scanning across the full view of the microphone array.

7 Summary of Papers

Paper I

Adaptive Beamforming Applied to Medical Ultrasound Imaging
Johan-Fredrik Synnevåg, Andreas Austeng, and Sverre Holm
Published in *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, August 2007

In the first contribution we explain how we have applied the minimum variance beamformer to medical ultrasound. We demonstrate how the subarray size and diagonal loading parameters affect the resolution of the beamformer in a typical medical ultrasound scenario. We also show that the lack of robustness the minimum variance beamformer displays, can cause problems if we assume incorrect wavefield parameters. We demonstrate that a large subarray size and small diagonal loading parameter severely affect the amplitude estimate if the assumed acoustic velocity is incorrect. We also show that if we reduce the subarray size or increase the diagonal loading, we can make the beamformer robust. We demonstrate improved image resolution on both simulated and experimental RF data of wire targets, and on experimental RF data from a heart phantom.

Paper II

Speckle Statistics in Adaptive Beamforming
Johan-Fredrik Synnevåg, Carl Inge Colombo Nilsen, and Sverre Holm
Published in *Proceedings of IEEE Ultrasonics Symposium*, October 2007

In the second contribution we compare the speckle statistics from images formed by delay-and-sum and minimum variance beamformers. We examine how the estimates of the spatial covariance matrix affect the speckle patterns in the minimum variance images. We show that if we use the spatial dimension only to estimate the covariance matrix, areas that appear relatively homogeneous in delay-and-sum images, appears as collection of point scatterers in the minimum variance images – reducing the overall brightness of the speckle. By using both spatial and temporal averaging, the speckle distribution appears similar to delay-and-sum, without compromising the improved resolution offered by the minimum variance beamformer. Results from both simulations and experimental RF data are shown.

Paper III

Benefits of Minimum Variance Beamforming in Medical Ultrasound Imaging
Johan-Fredrik Synnevåg, Andreas Austeng, and Sverre Holm
Submitted to *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*

In the third contribution we demonstrate that the high-resolution properties of the minimum variance beamformer can be exploited in other ways than simply improving image resolution. In this paper we investigate three scenarios which would normally lead to reduced image quality, unless compensated for. We show that by using a minimum variance beamformer on reception, we can reduce the transducer size and number of transducer elements without sacrificing image quality. Also, we show that we can compensate for the wider transmit beams used in combination with parallel receive beams to increase frame-rate. Hence, we can increase frame-rate without reducing the quality of the image. Finally, we show that we can increase penetration without sacrificing lateral resolution by lowering the transmit frequency and compensate for the reduced resolution using a minimum variance beamformer on reception.

Paper IV

A Low Complexity Data-Dependent Beamformer
Johan-Fredrik Synnevåg, Andreas Austeng, and Sverre Holm
Submitted to *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*

In the first three contributions we show that the minimum variance beamformer offers improvement in resolution of medical ultrasound images. But the improvement comes at a high cost. The computational overhead required by the method may prevent a real-time implementation in medical ultrasound systems at present. In Paper IV we suggest a simpler method, which adapts to the recorded data, but selects the transducer apodization from a number of predefined windows. The computational overhead will depend on the number of windows. In this paper we show results using four and twelve windows, and demonstrate significant improvement compared to delay-and-sum. The performance is not as good as for the minimum variance beamformer, due to the limited solution space.

Paper V

On Single Snapshot Minimum Variance Beamforming
Johan-Fredrik Synnevåg and Are F. C. Jensen
Submitted to *Signal Processing*

In Paper V we analyze the minimum variance beamformer when only a single snapshot of the wavefield is used to optimize the weights. We show that single snapshot beamforming reveals the same problems with signal cancellation as if sources are coherent. We express subarray averaging as a decorrelation filter, which allows us to predict scenarios that may lead to signal cancellation. Also, we present a slightly altered optimization problem, which avoids signal cancellation altogether.

Paper VI

Blind Source Separation for Convolutive Mixtures Using Spatially Resampled Observations
Johan-Fredrik Synnevåg and Tobias Dahl
Published in *Proceedings of the 14th European Signal Processing Conference*, September 2006

Paper VI is a contribution to the field of blind source separation. This field is related to adaptive beamforming, although the approach to find the solution is quite different. We may perform blind source separation using independent component analysis as long as the mixes are instantaneous. However, for propagating waves, the mixes are convolutive. One way to get around this, is to solve the problem for individual frequency bands, which introduces scaling and permutation problems. The separated sources are returned in a random order and thus requires a remixing of the different frequency bands to reconstruct the signals. Also, different scaling of the frequency bands makes reconstruction of the signals difficult. In this paper we propose to use an array of sensors and use spatial resampling of the observations to transform the model from a convolutive to an instantaneous mix, thus allowing standard methods to perform the actual separation. In that way all frequency bands are treated simultaneously, and we avoid the permutation and scaling problems. We apply the method to the cocktail-party problem with two sources, and demonstrate 15 dB separation.

8 Discussion

Papers I-IV concerns different aspects of adaptive beamforming in medical ultrasound. In the first paper we show how the minimum variance beamformer can be applied to medical ultrasound imaging. In the second paper we address its effect on speckle, and in the third we show how the method can be exploited. The fourth paper concerns the real-time requirements, where we suggest a simplified approach. Several approaches on how to apply adaptive beamformers to medical ultrasound have been suggested in recent years. All form the spatial covariance matrix in different ways, by averaging in different domains and by different ways of applying diagonal loading. We believe that ours is a fairly simple approach in that we perform all the processing on the same scan line, and use a simple way to apply diagonal loading.

The results in Paper II leads to an interesting question: What should speckle look like? The delay-and-sum beamformer gives a certain distribution of the brightness levels, but it is not necessarily "correct". The desirable speckle patterns in medical ultrasound are those with low variance – quite the opposite of what minimum variance beamformer gives, unless proper measures are taken. In the spirit of adaptive beamforming it may be possible to formulate similar optimization problems targeted at speckle reduction, but as the method stands right now, it is more likely to increase speckle variance than to reduce it.

Paper III takes a different approach to adaptive beamforming, by showing other benefits than reduction of mainlobe width and sidelobe level. We have focused on three enhancements in this paper, but we can think of more. In Paper VIII we have also shown that we can reduce the number of focal zones by applying the minimum variance beamformer on reception, which also increases the frame rate. We may also use the adaptive former for speckle reduction, by averaging the output power of different subarrays. The reduced resolution caused by the reduced apertures of the different subarrays may then be compensated for by the adaptive beamformer.

Paper IV takes a practical approach to adaptive beamforming, and combines minimum variance beamforming theory with more classical window design. Some of the windows we use are, however, quite different from the classical windows, in that they have maximum sensitivity outside of the steering angle. The proposed low-complexity adaptive beamformer works well in a medical ultrasound scenario, and may also be a candidate for other applications where processing speed is critical. We are, however, somewhat fortunate in medical ultrasound, as the focused source causes most of the energy to be reflected from a small range of angles. Hence, we do not need very many windows to adapt to typical scenarios.

Paper V gives us an interpretation of subarray averaging in a single snapshot context. In [1, 45] similar derivations for narrowband, coherent sources can be found.

Introduction

Paper VI is a small contribution to the field of blind source separation of convolutive mixtures. Although the method has strict limitations, it gives an example of how array processing techniques can be exploited in this field.

9 Conclusion and further research

We can summarize the contributions of the thesis as follows:

- We have successfully applied the minimum variance beamformer to medical ultrasound imaging.
- We have shown that the method can be made robust through controlling one of two parameters: subarray size or diagonal loading
- We have shown that by using a single snapshot for the minimum variance optimization, the speckle patterns, which are important for the images, are very different from conventional beamforming. The speckle becomes more similar by averaging over the temporal dimension as well.
- We have exploited the minimum variance beamformer for other enhancements than simply improving image resolution and contrast. We have demonstrated that smaller transducers, higher frame rates or higher penetration is possible without sacrificing image quality compared to delay-and-sum.
- We have suggested a low-complexity adaptive beamformer which is easy to implement and may suit the real-time requirements.
- We have shown how subarray averaging affects the performance of a single snapshot minimum variance beamformer. We have also suggested a modified optimization which avoids signal cancellation
- We have suggested use of spatial resampling as preprocessing for the problem of blind source separation of convolutive mixtures.

Adaptive beamforming has received considerable attention in the recent years, and several implementations of the method has occurred in the literature. The results are encouraging, but we lack proof that the methods will improve diagnostics. Future research should assess the methods with clinical objects.

In this project, we have only considered 1D arrays, but 2D arrays are becoming increasingly popular. Extending the methods to 2D arrays gives new challenges regarding the implementation. An implementation on a 2D array requires a significant increase in the number of computations, and the low-complexity adaptive beamformer may be a candidate in such a scenario.

We have used a rather ad-hoc window design for the low-complexity adaptive beamformer in Paper IV, based on both classical windows and the knowledge we have gained through use of the minimum variance beamformer. Future research should further refine the window design – also directed at different applications. Also, by analyzing which windows are actually used in different scenarios, we can learn more about optimal apodization functions for ultrasound imaging.

References

[1] Don H. Johnson and Dan E. Dudgeon. *Array Signal Processing: Concepts and Techniques*. Simon & Schuster, 1992.

[2] F. Bryn. Optimum signal processing of three-dimensional arrays operating on Gaussian signals and noise. *The Journal of the Acoustical Society of America*, 34(3):289–297, March 1962.

[3] J. Capon. High-resolution frequency-wavenumber spectrum analysis. *Proc. IEEE*, 57:1408–1418, August 1969.

[4] B. Widrow, P.E. Mantey, L.J. Griffiths, and B.B. Goode. Adaptive antenna systems. *Proceedings of the IEEE*, 55(12):2143–2159, Dec. 1967.

[5] S. Applebaum. Adaptive arrays. *Antennas and Propagation, IEEE Transactions on*, 24(5):585–598, Sep 1976.

[6] B. Widrow, K. Duvall, R. Gooch, and W. Newman. Signal cancellation phenomena in adaptive antennas: Causes and cures. *Antennas and Propagation, IEEE Transactions on [legacy, pre - 1988]*, 30(3):469–478, May 1982.

[7] J. E. Evans, J. R. Johnson, and D. F. Sun. High resolution angular spectrum estimation techniques for terrain scattering analysis and angle of arrival estimation. *Proc. 1st ASSP Workshop Spectral Estimation, Hamilton, Ont., Canada*, pages 134–139, 1981.

[8] Tie-Jun Shan, Mati Wax, and Thomas Kailath. On spatial smoothing for direction-of-arrival estimation of coherent signals. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 33(4):806–811, August 1985.

[9] Jian Li, P. Stoica, and Zhisong Wang. On robust Capon beamforming and diagonal loading. *IEEE Transactions on Signal Processing*, 51(7):1702–1715, July 2003.

[10] R.G. Lorenz and S.P. Boyd. Robust minimum variance beamforming. *Signal Processing, IEEE Transactions on*, 53(5):1684–1696, May 2005.

[11] S.A. Vorobyov, A.B. Gershman, and Zhi-Quan Luo. Robust adaptive beamforming using worst-case performance optimization: a solution to the signal mismatch problem. *Signal Processing, IEEE Transactions on*, 51(2):313–324, Feb 2003.

[12] P. Stoica, Zhisong Wang, and Jian Li. Robust capon beamforming. *Signal Processing Letters, IEEE*, 10(6):172–175, June 2003.

[13] Jian Li, P. Stoica, and Zhisong Wang. Doubly constrained robust capon beamformer. *Signal Processing, IEEE Transactions on*, 52(9):2407–2423, Sept. 2004.

[14] O. L. Frost. An algorithm for linearly constrained adaptive array processing. *Proc. IEEE*, 60:926–935, August 1972.

[15] L. J. Griffiths and C. Jim. An alternative approach to linearly constrained adaptive beamforming. *Proc. IEEE*, AP-30:27–34, January 1982.

[16] Y. Bresler, V.U. Reddy, and T. Kailath. Optimum beamforming for coherent signal and interferences. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 36(6):833–843, Jun 1988.

[17] Jian Li and P. Stoica. An adaptive filtering approach to spectral estimation and sar imaging. *Signal Processing, IEEE Transactions on*, 44(6):1469–1484, Jun 1996.

[18] P. Stoica, Hongbin Li, and Jian Li. A new derivation of the APES filter. *Signal Processing Letters, IEEE*, 6(8):205–206, Aug 1999.

[19] M.E. Ali and F. Schreib. Adaptive single snapshot beamforming: a new concept for the rejection of nonstationary and coherent interferers. *Signal Processing, IEEE Transactions on*, 40(12):3055–3058, Dec 1992.

[20] D. Johnson and S. DeGraaf. Improving the resolution of bearing in passive sonar arrays by eigenvalue analysis. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 30(4):638–647, Aug 1982.

[21] R. Schmidt. Multiple emitter location and signal parameter estimation. *Antennas and Propagation, IEEE Transactions on*, 34(3):276–280, Mar 1986.

[22] R. Roy, A. Paulraj, and T. Kailath. Esprit–a subspace rotation approach to estimation of parameters of cisoids in noise. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 34(5):1340–1342, Oct 1986.

[23] S. Sivanand, Jar-Ferr Yang, and M. Kaveh. Focusing filters for wide-band direction finding. *IEEE Transactions on Signal Processing*, 39(2):437–445, February 1991.

[24] Jeffrey Krolik and David Swingler. The performance of minimax spatial resampling filters for focusing wide-band arrays. *IEEE Transactions on Signal Processing*, 39(8):1899–1903, August 1991.

[25] F. Lingvall and T. Olofsson. On time-domain model-based ultrasonic array imaging. *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, 54(8):1623–1633, August 2007.

[26] F. Viola, M.A. Ellis, and W.F. Walker. Time-domain optimized near-field estimator for ultrasound imaging: Initial development and results. *Medical Imaging, IEEE Transactions on*, 27(1):99–110, Jan. 2008.

[27] R. Bethel, B. Shapo, and H. Van Trees. Single snapshot spatial processing: optimized and constrained. *Sensor Array and Multichannel Signal Processing Workshop Proceedings, 2002*, pages 508–512, Aug. 2002.

[28] D. A. Merton. Diagnostic medical ultrasound technology: A brief historical review. *Journal of Diagnostic Medical Sonography*, 13(5):10–23, 1997.

[29] Sverre Holm. Medisinsk ultralydavbildning. *Fra Fysikkens Verden*, pages 4–11, 1999.

[30] J. A. Jensen. Medical ultrasound imaging. *Progress in Biophysics and Molecular Biology*, 93(1):153–165, jan 2007.

[31] Magali Sasso and Claude Cohen-Bacrie. Medical ultrasound imaging using the fully adaptive beamformer. *Acoustics, Speech and Signal Processing, 2005. Proceedings (ICASSP '05). IEEE International Conference on*, 2:489–492, March 2005.

[32] F. Vignon and M.R. Burcher. Capon beamforming in medical ultrasound imaging with focused beams. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 55(3):619–628, March 2008.

[33] Zhisong Wang, Jian Li, and Renbiao Wu. Time-delay- and time-reversal-based robust Capon beamformers for ultrasound imaging. *IEEE Transactions on Medical Imaging*, 24:1308–1322, Oct. 2005.

[34] J. A. Mann and W. F. Walker. A constrained adaptive beamformer for medical ultrasound: Initial results. *Ultrasonics Symposium, 2002. Proceedings. 2002 IEEE*, 2:1807–1810, October 2002.

[35] F. Viola and W. F. Walker. Adaptive signal processing in medical ultrasound beamforming. *Proc. IEEE Ultrasonics Symposium*, 4:1980–1983, Sept. 2005.

[36] M.A. Ellis, F. Viola, and W.F. Walker. 4b-1 diffuse targets for improved contrast in beamforming adapted to target. *Ultrasonics Symposium, 2007. IEEE*, pages 216–219, Oct. 2007.

[37] I. K. Holfort, F. Gran, and J. A. Jensen. Minimum variance beamforming for high frame-rate ultrasound imaging. *Proc. IEEE Ultrasonics Symposium*, pages 1541–1544, Oct. 2007.

[38] J.-F. Cardoso. Blind signal separation: statistical principles. *Proceedings of the IEEE*, 86(10):2009–2025, Oct 1998.

[39] A. Hyvarinen. Fast and robust fixed-point algorithms for independent component analysis. *Neural Networks, IEEE Transactions on*, 10(3):626–634, May 1999.

[40] J.F. Cardoso and A. Souloumiac. Blind beamforming for non-gaussian signals. *Radar and Signal Processing, IEE Proceedings F*, 140(6):362–370, Dec 1993.

[41] Aapo Hyvärinen. Survey on independent component analysis. *Neural Computing Surveys*, 2:94–128, 1999.

[42] L.C.Parra and C.V. Alvino. Geometric source separation: Merging convolutive source separation with geometric beamforming. *IEEE Transactions on Speech and Audio Processing*, 10(6):352–362, September 2002.

[43] Lucas Parra and Craig Fancourt. An adaptive beamforming perspective on convolutive blind source separation. In *Noise Reduction in Speech Applications*. CRC Press LLC, 2002.

[44] Shoko Araki, Ryo Mukai, Shoji Makino, Tsuyoki Nishikawa, and Hiroshi Saruwatari. The fundamental limitation of frequency domain blind source separation for convolutive mixtures of speech. *IEEE Transactions on Speech and Audio Processing*, 11(2):109–116, March 2003.

[45] H. L. Van Trees. *Optimum Array Processing*. Wiley, New York, 2002.

I

Adaptive Beamforming Applied to Medical Ultrasound Imaging

Johan-Fredrik Synnevåg, Andreas Austeng, and Sverre Holm

Abstract

We have applied the minimum variance (MV) adaptive beamformer to medical ultrasound imaging and shown significant improvement in image quality compared to delay-and-sum (DAS). We demonstrate reduced mainlobe width and suppression of sidelobes on both simulated and experimental RF data of closely spaced wire targets, which gives potential contrast and resolution enhancement in medical images. The method is applied to experimental RF data from a heart phantom, in which we show increased resolution and improved definition of the ventricular walls.

A potential weakness of adaptive beamformers is sensitivity to errors in the assumed wavefield parameters. We look at two ways to increase robustness of the proposed method; spatial smoothing and diagonal loading. We show that both are controlled by a single parameter which can move the performance from that of a MV beamformer to that of a DAS beamformer. We evaluate the sensitivity to velocity errors and show that reliable amplitude estimates are achieved while the mainlobe width and sidelobe levels are still significantly lower than for the conventional beamformer.

1 Introduction

Delay-and-sum (DAS) beamforming is the standard technique in medical ultrasound imaging. An image is formed by transmitting a narrow beam in a number of angles and dynamically delaying and summing the received signals from all channels. The sidelobe level of the DAS beamformer can be controlled using aperture shading, resulting in increased contrast at the expense of resolution. In contrast to the predetermined shading in DAS, adaptive beamformers use the recorded wavefield to compute the aperture weights. By suppressing interfering signals from off-axis directions and allowing large sidelobes in directions in which there is no received energy, the adaptive beamformers can increase resolution.

The minimum variance (MV) adaptive beamformer [1,2] and subspace-based methods have mostly been studied in narrowband applications. Extensions to broadband imaging include preprocessing with focusing- and spatial resampling filters, allowing narrowband methods to be used on broadband data [3, 4]. The MV beamformer has been applied to medical ultrasound imaging by several authors [5–9]. Mann and Walker [5] have used a constrained adaptive beamformer on experimental data of a single point target and a cyst phantom demonstrating improved contrast and resolution. Sasso and Cohen-Bacrie [6] have applied a MV beamfomer on a simulated data-set, showing improved contrast in the final image. Viola and Walker [9] have investigated use of several adaptive beamformers on simulated data and demonstrated improved resolution. Wang *et al.* [7] have applied a robust MV beamformer to medical ultrasound imaging using a synthetic aperture focusing approach. In their method the spatial covariance matrix, which is required to find the optimal aperture shading, is calculated by averaging appropriately delayed data from each individual transmit element. Although this method allows dynamic focus on both transmission and reception, the loss in signal-to-noise ratio (SNR) caused by synthetic aperture focusing may be unacceptable in medical imaging applications. Also, data captured with different transmit elements may not be coherent due to target motion. In [8], we show preliminary results from the present method, where full dynamic focus was applied through synthetic focus of individual transmitters. In this paper, we apply fixed focus on transmission. That allows us to evaluate the performance of the beamformer outside the transmit focal point, giving more realistic results for medical ultrasound imaging. We also consider ways of increasing the robustness of the MV beamformer, which ensures reliable amplitude estimates and at the same time improves performance.

In Section 2, we present the method and show how robustness is achieved through either spatial smoothing or diagonal loading. In Section 3 we compare the MV beamformer to DAS on simulated and experimental RF data from closely spaced wire targets, and on experimental RF data of a heart phantom. We also evaluate the sensitivity of the method to errors in acoustic velocity and

show that reliable amplitude estimates can be achieved by either of the robust methods. We discuss our findings in Section 4 and draw conclusions in Section 5.

2 Methods

2.1 Signal Model and Minimum Variance Beamformer

We assume an array of M elements, each recording a signal $x_m(t)$. We consider $P+1$ scatterers, each reflecting a signal, $s_p(t)$. We assume that $s_0(t)$ originates from the focal point of the receiver, and that the other reflectors are sources of interference. Under these assumptions, the mth time-delayed channel can be described as:

$$x_m(t) = \frac{1}{r_{m,0}} s_0(t) + \sum_{p=1}^{P} \frac{1}{r_{m,p}} s_p(t) * \delta(t - \tau_{m,p}) + n_m(t), \tag{1}$$

where $r_{m,p}$ is the distance from reflector p to sensor m, $\delta(t)$ is the Dirac delta-function, $\tau_{m,p}$ is the delay from reflector p to sensor m, $n_m(t)$ is noise on channel m, and $(*)$ is the convolution operator. The M observations are ordered in a vector:

$$\mathbf{X}(t) = \begin{bmatrix} x_0(t) \\ x_1(t) \\ \vdots \\ x_{M-1}(t) \end{bmatrix}, \tag{2}$$

which represents the observed wavefield. After delaying each channel to focus at a point in the image, the goal of the adaptive beamformer is to compute the optimal aperture shading before combining the channels. The output of the beamformer is a weighted sum of the spatial measurements:

$$z(t) = \sum_{m=0}^{M-1} w_m(t) x_m(t) = \mathbf{w}(t)^H \mathbf{X}(t), \tag{3}$$

where $w_m(t)$ is the aperture weight for sensor m, and $\mathbf{w}(t) = \begin{bmatrix} w_0(t) & w_1(t) & \cdots & w_{M-1}(t) \end{bmatrix}^H$. The minimum variance beamformer seeks to minimize the variance (power) of $z(t)$:

$$P(t) = E[|z(t)|^2], \tag{4}$$

while maintaining unit gain in the focal point. This optimization problem can be formulated as:

$$\begin{aligned} &\min_{\mathbf{w}(t)} \mathbf{w}(t)^H \mathbf{R}(t) \mathbf{w}(t) \\ &\text{subject to } \mathbf{w}(t)^H \mathbf{a} = 1, \end{aligned} \tag{5}$$

where:

$$\mathbf{R}(t) = E\left[\mathbf{X}(t)\mathbf{X}(t)^H\right] \tag{6}$$

is the spatial covariance matrix and $\mathbf{a}$ is the equivalent of the steering vector in narrowband applications. Since the data already have been delayed to focus at the point of interest, $\mathbf{a}$ is simply a vector of ones. The solution to (5) is:

$$\mathbf{w}(t) = \frac{\mathbf{R}(t)^{-1}\mathbf{a}}{\mathbf{a}^H\mathbf{R}(t)^{-1}\mathbf{a}}. \tag{7}$$

Applying these weights to (3) gives the MV amplitude estimate.

2.2 Estimation of the Spatial Covariance Matrix

In practice the covariance matrix, $\mathbf{R}(t)$, in (7) is replaced by the sample covariance matrix, $\hat{\mathbf{R}}(t)$. Because the transmitted pulses in medical ultrasound imaging are short and non-stationary, $\mathbf{R}(t)$ is rapidly changing with time. Hence, the estimate must be calculated from a single, or only a few, temporal samples. If synthetic aperture focusing (SAFT) is used as in [7], the estimate can be obtained by averaging the covariance matrices of appropriately delayed observations from each individual transmit element. Target motion and SNR requirements, however, may prohibit use of SAFT in practical medical ultrasound systems and require all elements to transmit simultaneously. To obtain a good estimate, we instead use spatial smoothing [10], in which the array is divided into overlapping subarrays, and the covariance matrices for all subarrays are averaged. This technique also is used to avoid signal cancellation in narrowband applications when sources are correlated or coherent. The estimated covariance matrix at time t then becomes:

$$\hat{\mathbf{R}}(t) = \frac{1}{M-L+1}\sum_{l=0}^{M-L}\bar{\mathbf{X}}_l(t)\bar{\mathbf{X}}_l^H(t), \tag{8}$$

where:

$$\bar{\mathbf{X}}_l(t) = \begin{bmatrix} x_l(t) \\ x_{l+1}(t) \\ \vdots \\ x_{l+L-1}(t) \end{bmatrix}, \tag{9}$$

and L is the subarray length. By using $\hat{\mathbf{R}}(t)$ in (7) we see that the number of coefficients in $\mathbf{w}(t)$ is reduced, which limits the degrees of freedom to suppress interfering signals and noise. There is then a trade-off between accurate estimation of the covariance matrix and the length of the spatial filter. As we discuss in Section 3, decreasing the subarray length leads to a more robust solution, but resolution is decreased.

After computation of the optimal aperture weights using the sample covariance matrix, we obtain the amplitude estimate as:

$$\hat{z}(t) = \frac{1}{M-L+1} \sum_{l=0}^{M-L} \mathbf{w}(t)^H \bar{\mathbf{X}}_l(t). \tag{10}$$

The amplitude estimate can be viewed as a sum of $M-L+1$ directive elements. Hence, the minimum variance solution approaches DAS as the subarray length is decreased. The overlapping subarrays give some more weight to the central elements, corresponding to a nonuniform shading.

2.3 Robust Minimum Variance Beamforming

Due to the high resolution of the minimum variance beamformer at high SNR, a large number of scan lines may be required to avoid angular undersampling and to ensure robustness against errors in the wavefield parameters. We expect that such errors will affect the peaks in the amplitude estimates and lead to underestimation of the reflectivity of the targets. The constraint in (5) only assures that reflections originating from the focal point of the receiver are passed with unit gain – others are suppressed. Wrong assumptions of acoustic velocity or phase aberrations will, for instance, lead to targets appearing slightly out of focus. The MV beamformer will try to minimize these reflections. By increasing robustness, we constrain the level of suppression outside the focal point, and allow reflections appearing slightly out of focus to pass through the beamformer.

We consider two ways to increase robustness of the proposed method in this paper; decreasing the subarray length when computing (8), and diagonal loading of the estimated covariance matrix. As discussed in Section 2, the proposed method approaches the DAS beamformer as the subarray length, L, is decreased. In the extreme case where $L = 1$, (10) becomes the DAS solution with uniform aperture shading. Hence, by increasing the number of subarrays used to form the sample covariance matrix, we obtain an increasingly robust solution. However, increased robustness comes at the expense of resolution. The choice of subarray length should ensure that the covariance matrix estimate is invertible, which sets the upper limit on L to [11]:

$$L \leq M/2. \tag{11}$$

We show results for several values of L in Section 3.

A second common way to increase robustness of the minimum variance beamformer is to add a constant, ϵ, to the diagonal of the covariance matrix before evaluating (7). This means that $\hat{\mathbf{R}}(t)$ is replaced by $\hat{\mathbf{R}}(t) + \epsilon\mathbf{I}$. There exists several methods for calculating the value of ϵ based on the uncertainty in the model parameters [12]. We use a simple approach, in which the amount of diagonal loading is proportional to the power in the received signals [13].

We can view diagonal loading as adding spatially white noise to the recorded wavefield before computing the aperture weights. Increased noise level will constrain the sidelobe levels in directions in which there are no interfering signals, and thereby limit the level of suppression in directions in which interfering reflections appear. As white noise becomes dominant, the MV solution approaches the DAS beamformer with uniform shading. We can see this if we consider a wavefield of spatially white noise. The covariance matrix is then proportional to the identity matrix, $\mathbf{R}(t) = \sigma_n^2\mathbf{I}$, where σ_n^2 is the variance of the noise. From (7) we get:

$$\mathbf{w}(t) = \frac{(\sigma_n^2\mathbf{I})^{-1}\mathbf{a}}{\mathbf{a}^H(\sigma_n^2\mathbf{I})^{-1}\mathbf{a}} = \frac{\mathbf{a}}{\mathbf{a}^H\mathbf{a}} = \frac{1}{M}\mathbf{a}, \tag{12}$$

which is simply an unweighted sum of the elements. In the following results the amount of diagonal loading is found as:

$$\epsilon = \Delta \cdot \mathrm{tr}\{\hat{\mathbf{R}}(t)\}\mathbf{I}, \tag{13}$$

where Δ is a constant and $\mathrm{tr}\{\cdot\}$ is the trace operator. The amount of diagonal loading is then given by the power in the received signals and the predetermined value Δ. Δ determines the relative weight given to the optimal solution in (7) and the DAS solution in (12). Setting $\Delta = 1/L$ corresponds to equal weighting of the two, in the sense that the traces of $\hat{\mathbf{R}}(t)$ and $\epsilon\mathbf{I}$ are the same [13]. We also can interpret Δ as a parameter controlling the relative artificial noise level in the data. Setting $\Delta = 1/L$ corresponds to adding spatially white noise with the same average variance as the received signals. As we will see in Section 3, we obtain a robust solution using $\Delta = 1/L$.

We now have two ways of increasing robustness of the proposed method, which both move the performance from that of a MV beamformer to that of an unweighted DAS beamformer.

3 Results

We have applied the MV beamformer to both simulated and experimental RF data and compared the results to the DAS beamformer. All transmitter and receiver combinations were simulated or recorded. The first processing steps were common to both beamformers: We synthesized fixed focus on transmission and dynamic focus on reception by delay and sum of the recorded data from each individual transmitter. Delays were implemented by upsampling the received signals and selecting the sample closest to the theoretically predicted delay. We then computed the discrete time analytic signal of each channel. This was done to avoid symmetric beampatterns in the MV solution, which could decrease resolution. The receiver channels then were summed for the DAS beamformer. For the MV beamformer, the optimal aperture weights were calculated and

applied before summation. All MV beamformer results used subarray length $L \leq M/2$. Unless a diagonal loading parameter is given, we used $\Delta = 1/100L$ to ensure a well conditioned covariance matrix.

3.1 Simulated Data

We simulated an 18.5 mm, 96 element, 4 MHz transducer using Field II [14], imaging a number of pairwise reflectors located at depths 30-80 mm. The reflectors were separated by 2 mm. Transmit focus was 60 mm. We added white, Gaussian noise to each receiver channel before beamforming. SNR was approximately 40 dB per channel for the reflectors in focus.

Fig. 1 shows images obtained with the DAS and MV beamformers displayed over 55 dB dynamic range. Figs. 1(b)-(d) show the MV beamformer for different subarray lengths. We see that as L increases the reflectors are better resolved, whereas for smaller L the image is closer to DAS. Fig. 2 shows the steered responses at two different depths. Fig. 2(a) shows the reflectors at the focal point of the transducer and Fig. 2(b) shows the deepest reflectors. We see that there is a significant reduction in sidelobe level for the MV beamformer at both depths. Sidelobes rapidly decay towards the background noise level for all values of L, and they are between 15 and 20 dB lower than for DAS. We see that the mainlobe width decreases as subarrays get longer. For the reflectors at the transmit focal point, the width is less than 1/4 of DAS for all examples. The wires are resolved by about 8 dB for DAS and up to 40 dB for MV. At 80 mm, which is away from the transmit focal point, DAS is unable to resolve the reflectors, while MV can resolve them by up to 30 dB. For $L = 18$, the MV beamformer is unable to separate the reflections, but the sidelobes are still about 20 dB lower, giving better definition of edges.

3.2 Sensitivity to Velocity Errors

As mentioned, a potential weakness of adaptive beamformers is lack of robustness against errors in the assumed wavefield parameters. The following results show the MV beamformers' sensitivity to errors in acoustic velocity for the different robust methods. We used the same simulated data as in Section 3.1, and beamformed the data with overestimation of the acoustic velocity by 5%. An error of this size covers the uncertainty in medical ultrasound imaging, in which the acoustic velocities range from about 1440 m/s in fat to about 1570 m/s in blood [15]. Fig. 3 shows the steered response at depths 60 and 80 mm for DAS and MV for different subarray lengths. For $L = 48$, we see that the peak amplitudes are grossly underestimated at both depths, due to the high level of suppression of reflections appearing out of focus. We also see that, as the subarray length is decreased, the MV amplitude estimates approach those of DAS. For $L = 18$ at 60 mm, the mainlobe width is approximately the same as

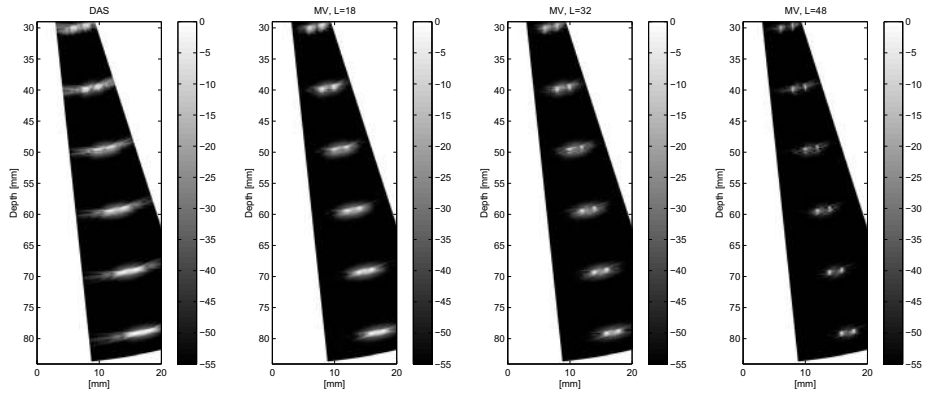


Fig. 1: Simulated wire targets using an 18.5 mm, 96 element, 4 MHz transducer. Transmit focus was 60 mm and dynamic receiver focus was applied. (a) DAS, (b) MV (L=18), (c) MV (L=32), and (d) MV (L=48).

DAS, but the sidelobes are about 12 dB lower. At 80 mm, the mainlobe width is narrower as well, and the amplitude estimates are approximately the same.

Fig. 4 shows corresponding results for different amounts of diagonal loading. We see that, as the regularization parameter, Δ, increases, the MV amplitude estimates approach those of DAS. As for spatial smoothing, diagonal loading trades off mainlobe width for robustness in the amplitude estimates. Again, the sidelobe level for all examples of Δ is significantly lower than for DAS.

3.3 Experimental Data: Wire Targets

We recorded experimental RF data with a specially programmed System FiVe scanner (GE Vingmed Ultrasound, Horten, Norway) using an 18.5 mm, 96 element, 3.5 MHz transducer driven at 4 MHz. The data were sampled at 20 MHz and upsampled by a factor 8 before delays were applied. The wires were separated by 2 mm. We used subarray length, $L = 48$ and $\Delta = 1/L$ for the MV beamformer. Fig. 5 shows the steered responses of the DAS and MV beamformers at two different depths, for which the latter is at the transmit focal point. We see that the sidelobes are significantly lower for the MV beamformer. In Fig. 5(a), which is away from the focal point, the mainlobe width is only slightly narrower due to the high amount of diagonal loading. At the focal point in Fig. 5(b), the mainlobe width is also reduced, demonstrating improved performance on experimental data.

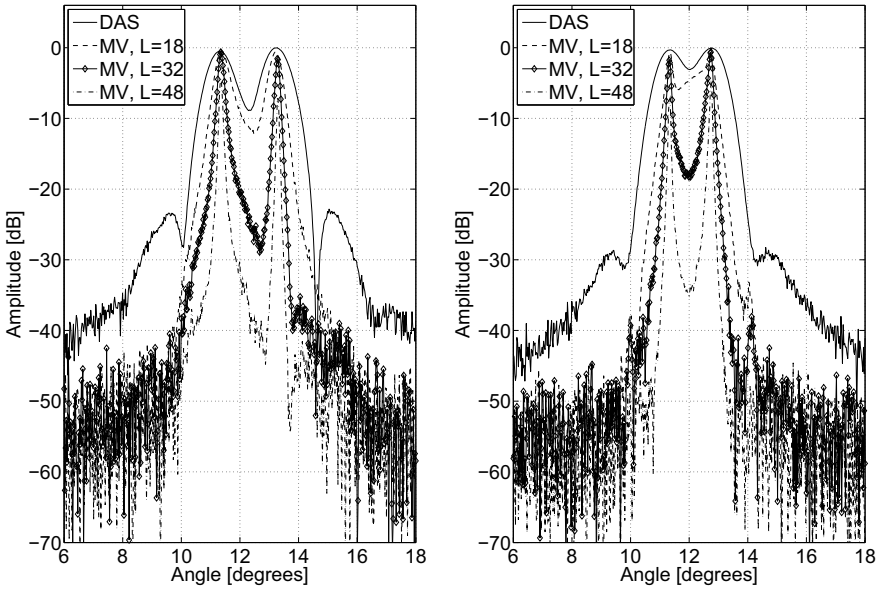


Fig. 2: Steered response of DAS and MV beamformers, using an 18.5 mm, 96 element, 4 MHz transducer, at depths (a) 60 mm and (b) 80 mm. Transmit focus was 60 mm and dynamic receiver focus was applied.

3.4 Experimental Data: Heart Phantom

The demonstrated properties of the MV beamformer on wire targets should lead to improved contrast and resolution in medical ultrasound images. We applied the method to experimental RF data from a heart phantom. The data-set was obtained from the Biomedical Ultrasound Laboratory, University of Michigan*. Data were recorded with a 64 element, 3.33 MHz transducer, and sampled at 17.76 MHz. Data were upsampled by a factor 8 before delays were applied. We used subarray length $L = 32$, and diagonal loading with $\Delta = 1/10L$ for the MV beamformer. Transmit focus was 60 mm. Fig. 6 shows the images obtained with the DAS and MV beamformers displayed over 55 dB dynamic range. We see that the resolution in generally much better in the MV image. We also see that the definition of the ventricular walls is improved, owing to the narrow beam and low sidelobe level.

*Ultrasound RF data-set heart from the Biomedical Ultrasound Laboratory, University of Michigan, Apr. 2006. Available at http://bul.eecs.umich.edu/, April 2006

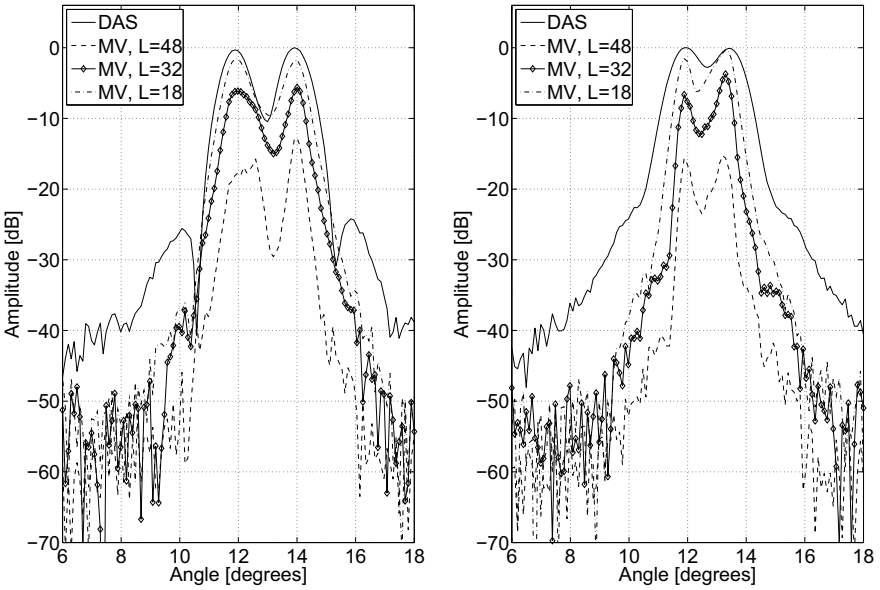


Fig. 3: Steered response of DAS and MV beamformers, using an 18.5 mm, 96 element, 4 MHz transducer in which images were formed with 5% error in acoustic velocity: (a) 60 mm and (b) 80 mm depth. Transmit focus was 60 mm and dynamic receiver focus was applied.

4 Discussion

The performance of the MV beamformer is dependent on the SNR. As SNR decreases, resolution will decrease in a manner similar to increasing the amount of diagonal loading. However, from (12) we see that in the absence of model-errors, the MV beamformer will perform equal or better than the unshaded DAS, provided that images are formed with sufficient angular sampling. Figs. 1 and 2 show that the MV beamformer achieved much higher resolution and lower sidelobes than DAS for $L = M/2$. Also, as subarray length was decreased, the resolution approached that of DAS. Still, the sidelobes were significantly lower than for DAS for all examples of L.

The results in Figs. 3 and 4 confirm that the MV beamformer underestimates the amplitudes of the reflected signals in the presence of wavefield perturbations. By decreasing the subarray length or applying diagonal loading reliable amplitude estimates were achieved, but the ability to resolve the targets was affected. We see from the figures that the main trade-off when choosing values for L and Δ is between mainlobe width and robustness. Sidelobes were significantly

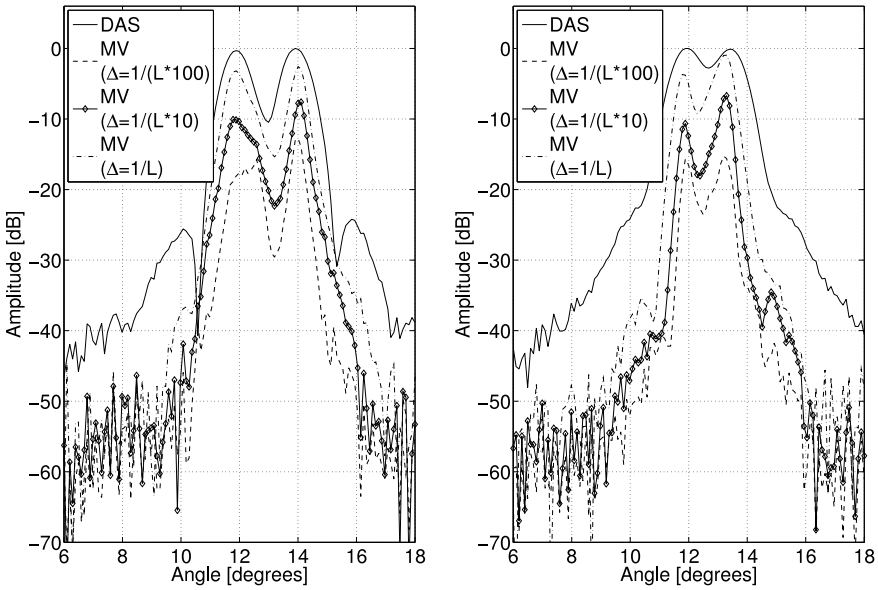


Fig. 4: Steered response of DAS and MV beamformers with different amount of diagonal loading, using an 18.5 mm, 96 element, 4 MHz transducer: (a) 60 mm and (b) 80 mm depth. Images were formed with 5% error in acoustic velocity. Transmit focus was 60 mm and dynamic receiver focus was applied.

lower than DAS for all examples of these parameters.

Setting $L = 18$ for the 96 element transducer in Fig. 3 gave amplitudes close to DAS, but resolution was only slightly improved. Using $L = M/2$ and $\Delta = 1/L$ in Fig. 4 gave similar results. A velocity shift of 5%, however, may be more than the typical estimation error. Beamforming the experimental heart-phantom data in Fig. 6 gave good results using $L = M/2$ and $\Delta = 1/10L$, in the sense that resolution was improved and the peak amplitudes in the images for DAS and MV were approximately the same.

The improved performance of the MV beamformer comes at a computational cost. As the delay-step is common to both beamformers, the computational overhead comes from calculation and application of the aperture weights. The computational complexity of direct computation of (7) is approximately $O(L^3)$. Increasing robustness using shorter subarrays, therefore, has a computational advantage over diagonal loading.

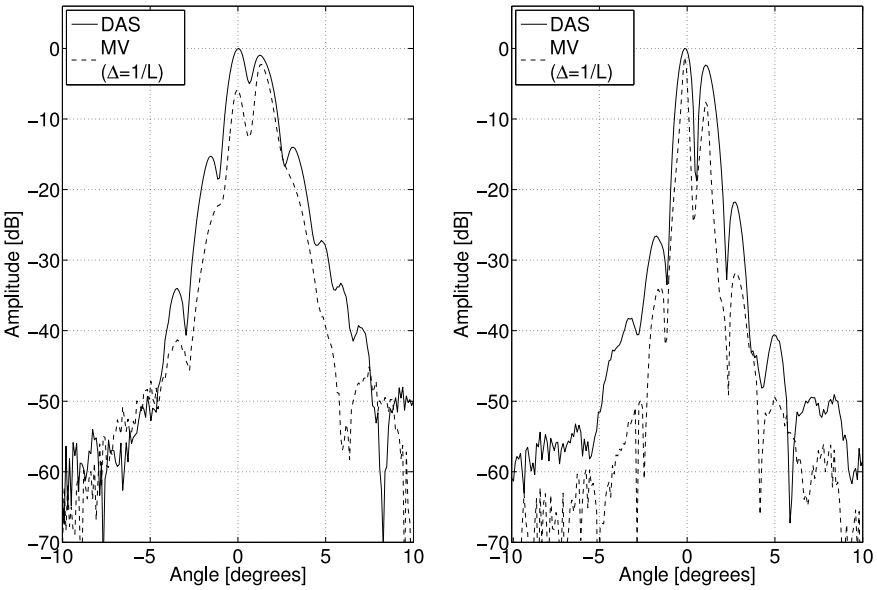


Fig. 5: Steered response of DAS and MV beamformers applied to experimental RF data, using an 18.5 mm, 96 element, 3.5 MHz transducer driven at 4 MHz: (a) 46 mm and (b) 56 mm depth. Transmit focus was 56 mm and dynamic receiver focus was applied.

5 Conclusions

We have successfully applied the minimum variance beamformer to medical ultrasound imaging and shown significant performance improvement compared to DAS. Results have been demonstrated on both simulated and experimental RF data. We have demonstrated two ways of increasing robustness of the method, which can give reliable amplitude estimates while still improving resolution and contrast. The method was applied to RF data from a realistic image, and showed significant improvement in image quality.

References

[1] F. Bryn. Optimum signal processing of three-dimensional arrays operating on Gaussian signals and noise. *The Journal of the Acoustical Society of America*, 34(3):289–297, March 1962.

[2] J. Capon. High-resolution frequency-wavenumber spectrum analysis. *Proc. IEEE*, 57:1408–1418, August 1969.

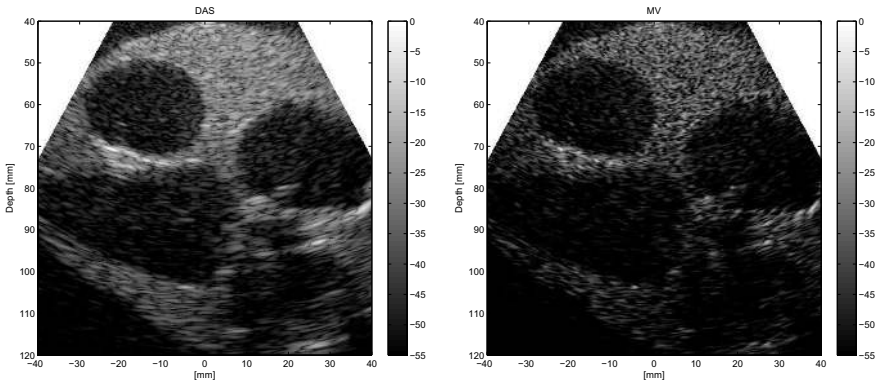


Fig. 6: Images obtained from experimental RF data from a heart-phantom using a 15.4 mm, 64 element, 3.33 MHz transducer: (a) DAS and (b) MV with $L = 32$ and $\Delta = 1/10L$. Transmit focus was 60 mm and dynamic receiver focus was applied.

[3] Jeffrey Krolik and David Swingler. The performance of minimax spatial resampling filters for focusing wide-band arrays. *IEEE Transactions on Signal Processing*, 39(8):1899–1903, August 1991.

[4] S. Sivanand, Jar-Ferr Yang, and M. Kaveh. Focusing filters for wide-band direction finding. *IEEE Transactions on Signal Processing*, 39(2), Februrary 1991.

[5] J. A. Mann and W. F. Walker. A constrained adaptive beamformer for medical ultrasound: Initial results. *Ultrasonics Symposium, 2002. Proceedings. 2002 IEEE*, 2:1807–1810, October 2002.

[6] Magali Sasso and Claude Cohen-Bacrie. Medical ultrasound imaging using the fully adaptive beamformer. *Acoustics, Speech and Signal Processing, 2005. Proceedings (ICASSP '05). IEEE International Conference on*, 2, March 2005.

[7] Zhisong Wang, Jian Li, and Renbiao Wu. Time-delay- and time-reversal-based robust Capon beamformers for ultrasound imaging. *Transactions on Medical Imaging*, 14:1308–1322, Oct. 2005.

[8] J.-F. Synnevåg, A. Austeng, and S. Holm. Minimum variance adaptive beamforming applied to medical ultrasound imaging. *Proc. IEEE Ultrasonics Symposium*, 2:1199–1202, Sept. 2005.

[9] F. Viola and W. F. Walker. Adaptive signal processing in medical ultrasound beamforming. *Proc. IEEE Ultrasonics Symposium*, 4:1980–1983, Sept. 2005.

[10] Tie-Jun Shan, Mati Wax, and Thomas Kailath. On spatial smoothing for direction-of-arrival estimation of coherent signals. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 33(4), August 1985.

[11] P. Stoica and Randolph Moses. *Introduction to Spectral Analysis*, chapter 5. Prentice Hall, 1997.

[12] Jian Li, P. Stoica, and Zhisong Wang. On robust Capon beamforming and diagonal loading. *IEEE Transactions on Signal Processing*, 51(7):1702–1715, July 2003.

[13] Ali Özbek. Adaptive seismic noise and interference attenuation method. United States Patent No.: US 6,446,008 B1, September 2002.

[14] J. A. Jensen. Field: A program for simulating ultrasound systems. *Medical & Biological Engineering & Computing*, 34:351–353, 1996.

[15] Bjørn A. J. Angelsen. *Ultrasound Imaging. Waves, Signals and Signal Processing*, chapter 1. Emantec, Norway, 2000.

II

Speckle Statistics in Adaptive Beamforming

Johan-Fredrik Synnevåg, Carl Inge Colombo Nilsen, and Sverre Holm

Abstract

We have examined the statistics of the speckle patterns in images formed using delay-and-sum and minimum variance beamforming. We show how the estimate of the spatial covariance matrix used in the latter beamformer affects the speckle patterns. By using purely spatial averaging in the estimate the speckle statistics are quite different from delay-and-sum: Areas that are apparently homogeneous in delay-and-sum images appear as a collection of point scatterers in the minimum variance images, reducing the overall brightness of the speckle. We show that temporal averaging reduces this effect without compromising the improved spatial resolution offered by the minimum variance beamformer. Results from both simulations and experimental RF data are shown.

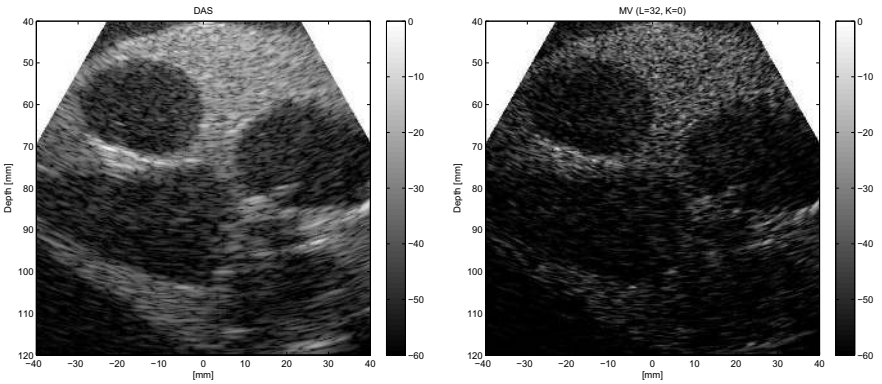


Fig. 1: Heart phantom imaged by a 15.4 mm, 64 element 3.33 MHz transducer. (a) DAS and (b) MV ($L = 32$, $K = 0$). Transmit focus was 60 mm.

1 Introduction

A number of authors have applied adaptive beamforming to medical ultrasound imaging [1–6]. Superior resolution has been demonstrated on point scatterers, but little attention has been paid to the effect on so called "speckle", which is the grainy appearance of homogeneous areas in ultrasound images. Speckle in ultrasound B-mode images has been researched since Burckhardt's initial paper in 1978 [7]. This research has resulted in a number of theories seeking to describe and explain the statistical properties of speckle, as well as many suggested methods for speckle reduction.

In this paper we investigate the difference in speckle statistics on images formed by delay-and-sum (DAS) and minimum variance (MV) beamformers. In particular, we examine the effects of temporal and spatial averaging on system resolution and speckle statistics of the MV beamformer. Fig. 1, which shows images of a heart phantom formed by DAS and MV, demonstrates the starting point of our investigation: while giving better edge definition, the MV beamformer in Fig. 1(b) gives the impression of point scatterers in homogeneous tissue. These structures appear more homogeneous and brighter in Fig. 1(a) which shows the image formed using DAS. Hence, the MV image displays a lower contrast.

2 Methods

2.1 The MV Beamformer

We assume an M element transducer in which $x_m[n]$ is the sampled output from element m. The output of a general beamformer operating on these measurements can be described as:

$$z[n] = \sum_{m=0}^{M-1} w_m[n] x_m[n - \Delta_m] = \mathbf{w}^H[n]\mathbf{X}[n], \tag{1}$$

where $w_m[n]$ is a (complex) weight, Δ_m is a delay applied to channel m to focus at a specific point in the image, $\mathbf{w}^H[n] = [\ w_0[n] \ \cdots \ w_{M-1}[n]\ ]$ and $\mathbf{X}[n] = [\ x_0[n-\Delta_0] \ \cdots \ x_{M-1}[n-\Delta_{M-1}]\ ]^T$. For a DAS beamformer the weights, $\mathbf{w}[n]$, are predetermined with the objective of minimizing sidelobes while keeping a narrow mainlobe. For the MV beamformer, $\mathbf{w}[n]$ is calculated from the recorded data by minimizing the variance of $z[n]$ while maintaining unit gain at the focal point. The analytical solution to this problem is given by [8]:

$$\mathbf{w}[n] = \frac{\mathbf{R}^{-1}[n]\mathbf{a}}{\mathbf{a}^H\mathbf{R}^{-1}[n]\mathbf{a}}, \tag{2}$$

where $\mathbf{R}[n]$ is the spatial covariance matrix, and $\mathbf{a}$ is the steering vector. Since the signals in $\mathbf{X}[n]$ have been delayed, $\mathbf{a}$ is a vector of ones.

In practice $\mathbf{R}[n]$ must be estimated, either by averaging in the temporal domain, the spatial domain, or both. Averaging in the spatial domain is done by dividing the transducer into (overlapping) subarrays and averaging the spatial covariance matrices of each subarray. The general covariance matrix estimate averaged over $2K+1$ temporal samples and $M-L+1$ subarrays of length L is given by:

$$\hat{\mathbf{R}}[n] = \frac{1}{(2K+1)(M-L+1)} \sum_{k=-K}^{K} \sum_{l=0}^{M-L} \mathbf{X}_l[n-k]\mathbf{X}_l^H[n-k], \tag{3}$$

where $\mathbf{X}_l[n] = [\ x_l[n-\Delta_l] \ \cdots \ x_{l+L-1}[n-\Delta_{l+L-1}]\ ]^T$. The subarray length, L, determines the degrees of freedom in the MV estimate. A smaller L gives a more robust estimate at the expense of resolution. The MV amplitude estimate is found by averaging the output of each subarray beamformer:

$$z_{MV}[n] = \frac{1}{M-L+1} \sum_{l=0}^{M-L} \mathbf{w}^H[n]\mathbf{X}_l[n]. \tag{4}$$

2.2 Speckle Statistics

As in optics, speckle is described as the result of scattering a transmitted signal by a large number of sub-wavelength sized particles that are densely distributed in space with random distances following a uniform probability density. "Densely" is meant with respect to the resolution cell size of the imaging system in question. The distribution of particles is considered spatially uncorrelated. When adding the phase shifted contributions from all these scatterers, the result is a signal with a Rayleigh distributed magnitude.

The spatial correlation of the resulting image does not reflect that of the original, uncorrelated particle distribution, but rather gives the impression of the existence of structures in the homogeneous tissue. Research suggests that these patterns reflect the resolution properties of the imaging system itself and no specific properties of the tissue (e.g. [9]). This means that the parameters of the intensity distribution are affected when changing the aperture by for instance increasing the array size or applying a weighting function. As adaptive beamformers like the MV beamformer apply time-varying weights that are calculated using the (estimated) spatial covariance matrix of the received signals, this will affect speckle statistics differently than DAS beamformers with constant weights. A clear difference is that the “resolution” of a MV beamformer is not fixed; it cannot be connected to the properties of a static beampattern like in the case of DAS. Instead it depends, for instance, on the accuracy of the covariance matrix estimate.

3 Results

3.1 Simulations

We simulated two phantoms using Field II [10]: To estimate the speckle statistics obtained by the DAS and MV beamformers we simulated 100,000 point scatterers randomly distributed within an area of 20x20x20 mm. All reflection coefficients of the scatterers were set equal. In the second phantom we used the same distribution of point scatterers, but added four strong reflectors, leading to speckle images with four distinct points.

We simulated an 18.5 mm, 96 element, 4 MHz transducer. All transmitter and receiver combinations were simulated, allowing dynamic focus on both transmission and reception. We synthesized dynamic transmit focus by combining appropriately delayed receiver channels from each transmit element. The new set of receiver channels were then dynamically delayed. Delays were implemented by upsampling and selecting the sample closest to the theoretically predicted delay. For the DAS beamformer, the receiver channels were then summed. For MV, the aperture weights were calculated and applied.

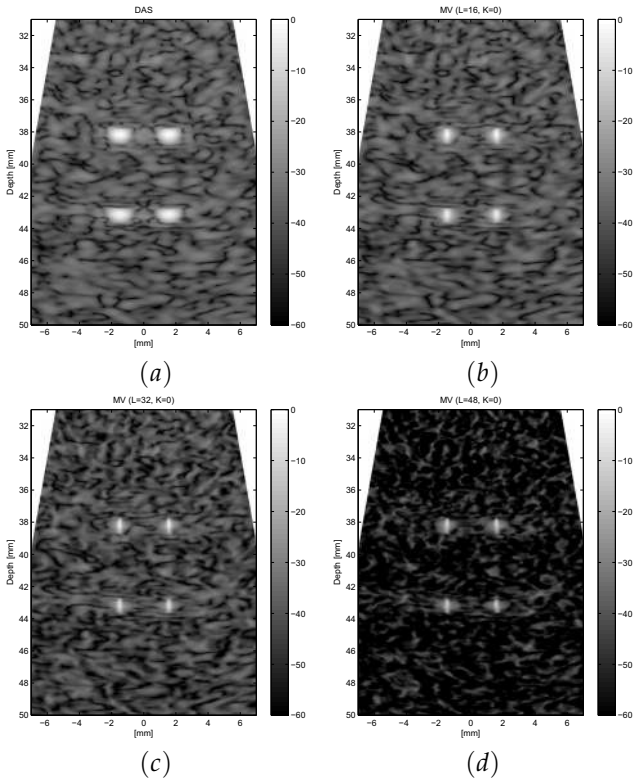


Fig. 2: Images of speckle and point scatterers using an 18.5 mm, 96 element, 4 MHz transducer: (a) DAS (b) MV ($L = 16$, $K = 0$), (c) MV ($L = 32$, $K = 0$) (d) MV ($L = 48$, $K = 0$).

3.2 Speckle Patterns

Fig. 2 shows images of the four point scatterers surrounded by speckle formed by DAS and MV for different subarray lengths. All MV results use $K = 0$, corresponding to no temporal averaging. We see that as L increases, the point scatterers are better defined. We also see that the intensity of the speckle is reduced when the subarray length increases. For $L = 48$ the speckle grains seem to be resolved into individual point scatterers. This may be an undesirable effect in medical ultrasound imaging where speckle regions are important for the image formation.

To investigate the statistics of the speckle in Fig. 2 we used the simulation of the phantom without the four strong reflectors. Fig. 3 shows the magnitude distributions for the images consisting of pure speckle for DAS and MV with different subarray lengths. We see that the distributions resemble Rayleigh

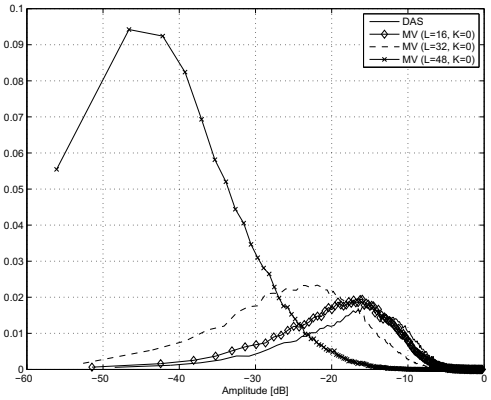


Fig. 3: Magnitude distributions for the speckle in Fig. 2.

distributions for all beamformers, but with different average magnitude and variance. For the "one-sample" ($K = 0$) MV beamformers the average magnitude is reduced as L increases.

Fig. 4 shows corresponding images as in Fig. 2, but with different amount of temporal averaging for the MV beamformer. Fig. 4(a) shows DAS and Fig. 4(b)-(d) shows MV using $L = 48$ and different values of K. We see that as the amount of temporal averaging increases, the resolution of the system remains approximately the same, but the intensity of the speckle approaches that of DAS. This can also be seen in Fig. 5 which shows the speckle statistics for DAS and MV using the same parameters. As K increases, the average magnitude increases, and the distribution approaches that of DAS.

3.3 Heart Phantom

We applied the DAS and MV beamformers to experimental RF data from a heart phantom. The data were obtained from the Biomedical Ultrasound Laboratory, University of Michigan.* The RF data were recorded with a 64 element, 3.33 MHz transducer. All transmitter and receiver combinations were available, and we synthesized fixed transmit focus at 60 mm. Fig. 1 and Fig. 6 shows the images obtained with DAS and MV for different amount of temporal averaging. Both MV results use $L = 32$. The MV image in Fig. 1(b), in which $K = 0$, displays a better definition of the ventricular walls than for the DAS in Fig. 1(a), but the tissue appears less homogeneous and the average intensity of the speckle is lower. This results in an image with lower contrast. The MV image in Fig. 6, in which $K = 11$, displays similar resolution as for $K = 0$, but the intensity of

*Ultrasound RF data-set 'heart' from the Biomedical Ultrasound Laboratory, University of Michigan. Available at http://bul.eecs.umich.edu/, April 2006.

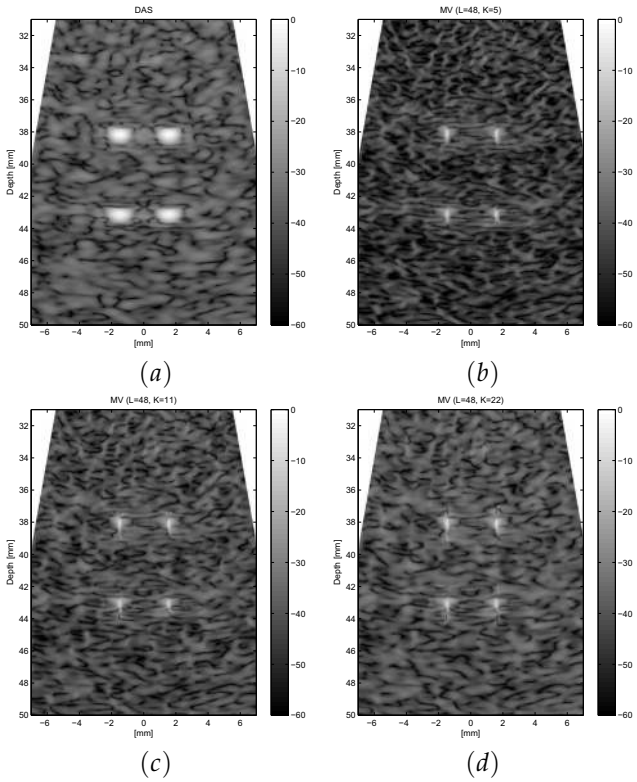


Fig. 4: Images of speckle and point scatterers using an 18.5 mm, 96 element, 4 MHz transducer: (a) DAS (b) MV ($L = 48$, $K = 5$), (c) MV ($L = 48$, $K = 11$) (d) MV ($L = 48$, $K = 22$).

the speckle is closer to DAS. Thus, resolution is improved while the contrast is retained.

4 Discussion

The spatially averaged MV results in Fig. 2 show that we achieve increased resolution when subarray length increases. However, increasing L has a significant effect on speckle: From Fig. 3 we see that the average magnitude of the speckle is reduced by approximately 8 dB for $L = 32$ and approximately 30 dB for $L = 48$ compared to DAS.

Fig. 4 shows that by using temporal as well as spatial averaging, the average magnitude of the speckle approaches that of DAS as K increases. At the same time we retain the the good resolution of Fig. 2(d). From Fig. 5 we see that for

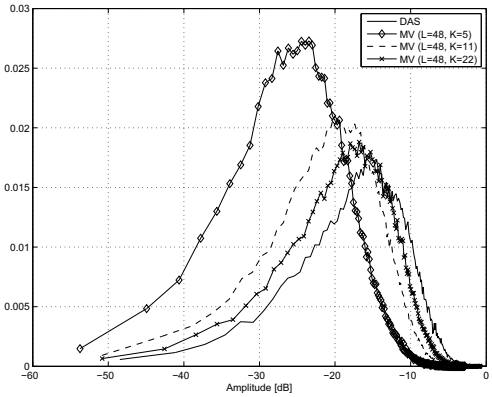


Fig. 5: Magnitude distributions for the speckle in Fig. 4.

$K = 5$, 11 and 22, which corresponds to temporal averaging over 1/2, 1 and 2 pulse lengths, the average speckle magnitude is approximately 10, 4 and 2 dB lower than for DAS, respectively. This shows that similar speckle brightness as DAS is achievable while improving the spatial resolution of the system.

Although the period over which the transmit pulses in medical ultrasound can be considered stationary is short, these results suggest that the one-sample MV beamformers ($K = 0$) do not capture the statistics of a speckle process. The transmit pulse merely acts as a filter on a random, stationary process, forcing use of temporal averaging to capture the spatial covariance of the backscattered signals. While giving an acceptable estimate of the spatial covariance of echoes from strong, isolated scatterers, it yields a poor estimate of the more stationary backscatter from homogeneous tissue. However, by use of temporal averaging the speckle intensity in MV images is similar to DAS, while the good resolution of the beamformer is retained.

5 Conclusion

We have examined the difference in speckle statistics on images formed by delay-and-sum and minimum variance beamformers. We have shown that the differences depend on how the spatial covariance matrix required by the minimum variance beamformer is estimated. By averaging both in the spatial and temporal domain, similar speckle statistics as delay-and-sum is achieved without compromising the improved spatial resolution offered by the minimum variance beamformer.

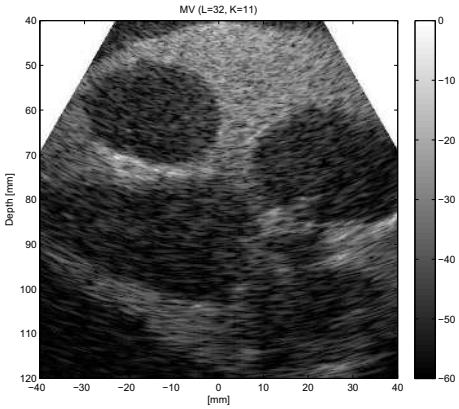


Fig. 6: Heart phantom imaged by a 15.4 mm, 64 element 3.33 MHz transducer using MV beamforming ($L = 32$, $K = 11$). Transmit focus was 60 mm.

6 Acknowledgment

We want to thank Dr Kjell Kristoffersen, GE Vingmed Ultrasound, for pointing out the differences in speckle statistics, and for several helpful discussions.

References

[1] J. A. Mann and W. F. Walker, "A constrained adaptive beamformer for medical ultrasound: Initial results," *Ultrasonics Symposium, 2002. Proceedings. 2002 IEEE*, vol. 2, pp. 1807–1810, October 2002.

[2] M. Sasso and C. Cohen-Bacrie, "Medical ultrasound imaging using the fully adaptive beamformer," *Acoustics, Speech and Signal Processing, 2005. Proceedings (ICASSP '05). IEEE International Conference on*, vol. 2, pp. 489–492, March 2005.

[3] J.-F. Synnevåg, A. Austeng, and S. Holm, "Minimum variance adaptive beamforming applied to medical ultrasound imaging," *Proc. IEEE Ultrasonics Symposium*, vol. 2, pp. 1199–1202, Sept. 2005.

[4] F. Viola and W. F. Walker, "Adaptive signal processing in medical ultrasound beamforming," *Proc. IEEE Ultrasonics Symposium*, vol. 4, pp. 1980–1983, Sept. 2005.

[5] Z. Wang, J. Li, and R. Wu, "Time-delay- and time-reversal-based robust Capon beamformers for ultrasound imaging," *IEEE Transactions on Medical Imaging*, vol. 24, pp. 1308–1322, Oct. 2005.

[6] J.-F. Synnevåg, A. Austeng, and S. Holm, "Adaptive beamforming applied to medical ultrasound imaging," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 54, no. 8, pp. 1606–1613, Aug. 2007.

[7] C. B. Burckhardt, "Speckle in ultrasound b-mode scans," *IEEE Transactions on Sonics and Ultrasonics*, vol. SU-25, pp. 1–6, January 1978.

[8] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, pp. 1408–1418, August 1969.

[9] R. F. Wagner, S. W. Smith, J. M. Sandrik, and H. Lopez, "Statistics of speckle in ultrasound b-scans," *IEEE Transactions on Sonics and Ultrasonics*, vol. 30, pp. 156–163, May 1983.

[10] J. A. Jensen, "Field: A program for simulating ultrasound systems," *Medical & Biological Engineering & Computing*, vol. 34, pp. 351–353, 1996.

III

Benefits of Minimum Variance Beamforming in Medical Ultrasound Imaging

Johan-Fredrik Synnevåg, Andreas Austeng, and Sverre Holm

Abstract

Recently, significant improvement in image resolution has been demonstrated by applying adaptive beamforming to medical ultrasound imaging. In this paper we have used the minimum variance beamformer to show how the low sidelobe levels and narrow beamwidth of adaptive methods can be utilized, not only to increase resolution, but to enhance imaging in several ways: By using a minimum variance beamformer instead of delay-and-sum on reception, reduced aperture, higher frame rates or increased depth of penetration can be achieved without sacrificing image quality. We demonstrate comparable resolution on images of wire targets and a cyst phantom obtained with a 96 element, 18.5 mm transducer using delay-and-sum, and a 48 element, 9.25 mm transducer using minimum variance. To increase frame rate, fewer and wider transmit beams in combination with several parallel receive beams may be used. We show comparable resolution to delay-and-sum using minimum variance, 1/4th of the number of transmit beams and four parallel receive beams, potentially increasing the frame rate by four. Finally, we show that by lowering the frequency of the transmitted beam and beamforming the received data with the minimum variance beamformer, increased depth of penetration is achieved without sacrificing lateral resolution.

1 Introduction

Recently, several authors have applied adaptive beamforming to medical ultrasound imaging and shown significant improvement in image resolution compared to delay-and-sum (DAS) [1–8]. The main difference from conventional delay-and-sum beamforming is that the weights are not predetermined, but signal dependent. In addition they may be complex, thus allowing for beampatterns that are asymmetrical around the center of the beam.

Adaptive beamforming methods such as the Capon and Frost beamformers were originally developed for passive direction finding [9, 10]. In a system with active transmission, such as medical ultrasound imaging, the presence of a coherent source generates received echoes which also are coherent, but the standard adaptive methods assume incoherent echoes. Without any modification, the methods may be useful for beamwidth reduction and sidelobe suppression in simple scenarios like a cyst or isolated point targets, as demonstrated by one of the first to apply minimum variance beamforming to medical ultrasound, Mann and Walker [1]. But for more complex scenarios, several measures need to be taken in order to handle signal coherence in the returning echos. These measures are in the form of various forms of averaging and smoothing of the estimates.

Subaperture averaging is the main measure that has been developed for handling coherent echos. It was first introduced by Evans et al. [11]. In this context subaperture averaging means that the spatial covariance matrix is estimated by averaging the covariance matrices of different subapertures. This technique was applied to medical ultrasound imaging by Sasso and Cohen-Bacrie [2], but only demonstrated on cyst images. The imaging scenario that best demonstrates the need for subaperture averaging is two very close targets, closer than the limit which can be resolved by DAS. Using this scenario, we have demonstrated that better resolution than DAS was possible even with coherent echos [3].

A second measure, originally used to make high-resolution methods more robust, is *diagonal loading*. This is a regularization method that adds a constant to the diagonal of the estimate of the covariance matrix. Wang et al. used an extension of this technique and demonstrated robustness to element amplitude errors on cyst and single point targets in medical ultrasound imaging [5]. In [3] we showed how it makes the method more robust to differences between actual and assumed velocity of sound. Both subaperture averaging and diagonal loading have the desirable property that variation of a single parameter allows one to adjust the method so that it falls back to conventional delay-and-sum beamforming [6].

A third way to do smoothing is to use the time dimension. In [2], Sasso and Cohen-Bacrie proposed to combine averaging from frame to frame with averaging of the covariance estimate over neighboring beams. We are concerned that this may affect resolution both in time and space. Instead we have

shown that averaging over a small range gate is one method for getting speckle statistics that resembles DAS beamforming speckle statistics [12]. The range gate was comparable in extent to the transmit pulse, and therefore has negligible effect on spatial resolution. In [12] we showed that without temporal smoothing the current method displays lower average intensity and lower speckle SNR than corresponding DAS images. Similar observations were made in [8]. Holfort et al. also did implicit time averaging since they split the transducers bandwidth into independent bands by FFT, performed independent minimum variance beamforming per band, and combined them [7].

Our method for minimum variance beamforming at present is based on subaperture averaging, diagonal loading, and time averaging in range. It can handle coherent echos, it is fairly robust to parameter mismatch, and the images have speckle which resembles that of conventional imaging.

In all publications to date the focus has been to increase image resolution and contrast. In this paper we show that the properties of the minimum variance (MV) beamformer can be utilized to enhance medical ultrasound imaging in other ways. We show that it is possible to reduce transducer size, increase frame rate or increase penetration depth without sacrificing image quality compared to DAS. Reduction in transducer size is possible since the resolution of the MV beamformer is much better than that of DAS. A smaller aperture is therefore sufficient to achieve comparable results. We also demonstrate that increased frame rate can be achieved without loss of resolution using fewer and wider transmit beams combined with several parallel receive beams processed with the MV beamformer. The resolution lost by the wide transmit beams is compensated for. Finally, we show that increased depth of penetration is achieved in attenuating media if we lower the center frequency of the transmit beam and compensate for the lost resolution by a MV beamformer.

In the next section we present how we apply the MV beamformer to medical ultrasound imaging and present three ways to utilize its properties. In Section 3 we show results from each of the proposed enhancements. We discuss our findings in Section 4 and draw conclusions in Section 5.

2 Methods

2.1 Minimum Variance Beamformer

We assume a transducer consisting of M elements, and beamformers operating on the sampled output of each channel, $x_m[n]$. The output of a beamformer focusing at a point in the image can be written as a weighted sum of the delayed measurements,

$$z[n] = \sum_{m=0}^{M-1} w_m[n] x_m[n - \Delta_m[n]], \tag{1}$$

where $w_m[n]$ is a (complex) weight and $\Delta_m[n]$ is the delay for element m at sample n. We have included time-dependence for the weights, as they may vary with image depth. In vector form, (1) can be written as

$$z[n] = \mathbf{w}[n]^H \mathbf{X}[n], \tag{2}$$

where

$$\mathbf{X}[n] = \begin{bmatrix} x_0[n - \Delta_0[n]] \\ x_1[n - \Delta_1[n]] \\ \vdots \\ x_{M-1}[n - \Delta_{M-1}[n]] \end{bmatrix}, \tag{3}$$

$$\mathbf{w}[n] = \begin{bmatrix} w_0^*[n] \\ w_1^*[n] \\ \vdots \\ w_{M-1}^*[n] \end{bmatrix}, \tag{4}$$

* denotes the complex conjugate, and ^{H} denotes conjugate transpose. Note that the variance of $z[n]$ can be written as

$$E\left[|z[n]|^2\right] = \mathbf{w}[n]^H \mathbf{R}[n] \mathbf{w}[n], \tag{5}$$

where $E[\cdot]$ denotes the expectation, and $\mathbf{R}[n] = E\left[\mathbf{X}[n]\mathbf{X}[n]^H\right]$ is the spatial covariance matrix. Application of predetermined weights is usually a tradeoff between mainlobe width and sidelobe level of the reception beam. Near the transducer a window gradually falling towards the edges of the aperture (e.g. Hamming) is often used for sidelobe suppression, while a rectangular window is often used at greater depths for maximum penetration. Instead of predetermined weights, the MV beamformer calculates the aperture shading from the recorded data. The weights are found by minimizing the variance of $z[n]$, while forcing unit gain at the focal point, leading to the optimization problem:

$$\begin{aligned} &\min_{\mathbf{w}[n]} \mathbf{w}[n]^H \mathbf{R}[n] \mathbf{w}[n] \\ &\text{subject to } \mathbf{w}[n]^H \mathbf{a} = 1, \end{aligned} \tag{6}$$

where $\mathbf{a}$ is the *steering vector*. Since the data has been delayed, $\mathbf{a}$ is simply a vector of ones. Different element sensitivities could also be included in $\mathbf{a}$, giving entries different from one. Alternatively, uncertainties in the sensitivities could be handled by adding constraints [13]. Eq. (6) has an analytical solution given by [9]:

$$\mathbf{w}[n] = \frac{\mathbf{R}[n]^{-1}\mathbf{a}}{\mathbf{a}^H \mathbf{R}[n]^{-1}\mathbf{a}}. \tag{7}$$

In contrast to the fixed beampatterns provided by predetermined aperture shading, the MV beamformer will adapt to the recorded data, and allow

large sidelobes in directions where there is little energy, and place zeros in directions where there is interference. With complex data, beampatterns that are asymmetric around the steering angle may also arise as the optimal solutions, as the only constraint is that the response has unit gain at the focal point of the receive beam.

In order to demonstrate the effect of adaptive weights, we show two examples of the resulting angular responses (*beampatterns)* of the apodization functions. The beampatterns were calculated for the center frequency of the transducer by applying the spatial Fourier transform to the apodization functions. Fig. 1 shows the beampatterns of the weights calculated by the MV beamformer. The beampattern of a rectangular window has been included for reference. In Fig. 1(a) we image two closely spaced wire targets at $\pm$ 1 degree angle, and steer towards the one at +1 degree (see Fig. 3 for examples of images). The solid vertical line indicates the angle of steering. We see that signals propagating from this direction are passed with unit gain. The dotted vertical line indicates the incoming angle of the reflection from the interfering scatterer. We see that the response of the MV weights is close to zero at this angle. On the other hand, in regions where there is no signal energy, the beamformer allows large sidelobes (see for instance from 2-7 degrees). We also see that the mainlobe of the beampattern is asymmetric around the steering angle.

Fig. 1(b) shows an example where we image a cyst surrounded by tissue and steer close to the edge of the cyst (see Fig. 5 for examples of images). The solid horizontal line indicates the cyst region, and the solid vertical line indicates the angle of steering. There is no signal scattering inside the cyst, so a large sidelobe is present here. We also see that the mainlobe is asymmetric to achieve maximum signal suppression from the surrounding tissue which is outside of the focal point.

Estimation of the spatial covariance matrix, $\mathbf{R}[n]$ in (6), requires some sort of averaging of the observed signals. In this application, averaging in the spatial domain is required for two reasons:

- Short and non-stationary transmit pulses reflecting off point scatterers lead to a non-stationary $\mathbf{X}[n]$, which limits the use of temporal averaging.
- Signals from different scatterers may be correlated, as all reflections originate from the same transmit pulse. This can lead to signal cancellation, and may be avoided by subarray averaging [11, 14].

In [6] we estimated the covariance matrix by averaging in the spatial domain only. However, for objects consisting of unresolvable micro-structures (leading to speckle), averaging solely in the spatial dimension may give very different image statistics compared to DAS [12]. The speckle variance tends to increase, such that tissue appears less homogeneous in the MV images. Although "minimum variance" beamforming may suggest that the speckle variance should be reduced, "minimum variance" in this context refers to the output of the

beamformer at the focal point of the receive beam. The beamformer is designed to suppress signals that are incoherent across the aperture, resulting in larger areas of lower scattering amplitudes compared to DAS. We believe that by using purely spatial averaging, which works very well on isolated point scatterers, we do not capture the statistics of the more stationary process of speckle generation. In [12] we show that by applying temporal averaging - as well as spatial averaging - similar speckle statistics are retained. The general estimate of the covariance matrix with subarrays of length L and temporal averaging over $2K+1$ samples then becomes:

$$\hat{\mathbf{R}}[n] = \frac{1}{(2K+1)(M-L+1)} \sum_{k=-K}^{K} \sum_{l=0}^{M-L} \bar{\mathbf{X}}_l[n-k]\bar{\mathbf{X}}_l[n-k]^H, \tag{8}$$

where

$$\bar{\mathbf{X}}_l[n] = \begin{bmatrix} x_l[n] \\ x_{l+1}[n] \\ \vdots \\ x_{l+L-1}[n] \end{bmatrix}. \tag{9}$$

By using $\mathbf{R}[n] = \hat{\mathbf{R}}[n]$ in (7) we find the aperture shading to be applied to each subarray, and get the amplitude estimate for the total array as:

$$\hat{z}[n] = \frac{1}{M-L+1} \sum_{l=0}^{M-L} \mathbf{w}[n]^H \bar{\mathbf{X}}_l[n]. \tag{10}$$

The choice of L determines the degrees of freedom in the MV estimate, and trades off resolution for robustness. A smaller L gives a more robust estimate at the expense of resolution. It is important to note that although the covariance matrix, and hence $\mathbf{w}[n]$, is estimated from samples $n-K,\ldots,n+K$, the amplitude estimate only depends on observations at sample n. Hence, there is no smoothing in the time-domain and we retain the temporal resolution.

The high resolution of the MV beamformer at high signal-to-noise ratios (SNR) makes it sensitive to wavefield perturbations and may also require dense angular sampling to avoid amplitude loss. In [6] we discussed two ways to increase robustness of the MV estimate: Decrease the subarray length, L, and apply diagonal loading. Diagonal loading means that $\mathbf{R}[n]$ is replaced by $\mathbf{R}[n] + \epsilon\mathbf{I}$ in (7). We choose ϵ to be proportional to the power of the observations:

$$\epsilon = \Delta \cdot \mathrm{tr}\{\mathbf{R}[n]\}, \tag{11}$$

where $\mathrm{tr}\{\cdot\}$ is the trace operator and Δ is typically (much) less than $1/L$. See [6] for a discussion of the choice of parameters.

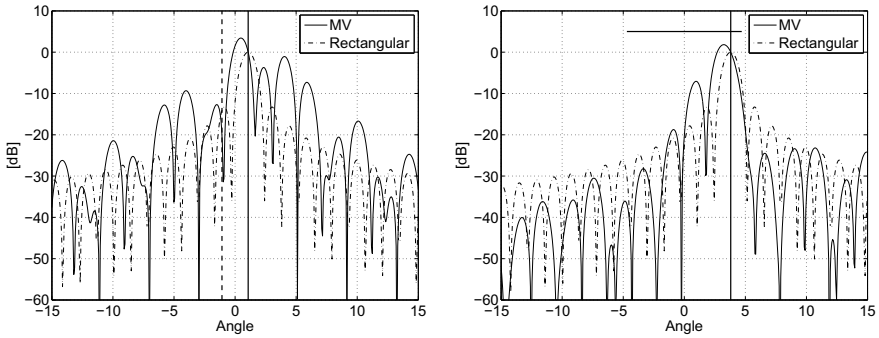


Fig. 1: Beampatterns of the MV beamformer: (a) Wire targets (b) Cyst. The dashed line is the corresponding beampattern of a rectangular window. The vertical lines in (a) indicate the location of the wire targets, where the solid line is the steering angle. The solid vertical line in (b) indicates the steering angle, and the horizontal line indicates the cyst.

2.2 Applications

Besides increasing image resolution using conventional transducers, which we showed in [6], we propose to utilize the properties of the MV beamformer for three other enhancements; reduce transducer size, increase frame rate and increase depth of penetration without sacrificing image quality. The high resolution of the MV beamformer allows for smaller transducers, and may allow use of a single transducer for several applications. In Section 3 we show results from the MV beamformer using a transducer of half the size of DAS. A similar experiment was performed in [5]. However, they use a synthetic aperture approach and form the covariance matrix by averaging data from each individual transmit element. The loss in signal-to-noise ratio due to the synthetic aperture focusing may be unacceptable in medical ultrasound imaging. Also, target motion may cause incoherence between data captured with different transmit elements. In this paper we use a focused transducer and perform all MV beamforming on reception, giving realistic results for practical medical ultrasound systems.

By using parallel receive beamforming (PRB) we can increase frame rate in medical ultrasound imaging [15]. With this technique fewer transmit (Tx) beams combined with multiple parallel receive (Rx) beams are used to increase frame rate. However, to avoid gain loss we must transmit wider beams, leading to lower resolution and resulting in a tradeoff between image quality and frame rate. Fig. 2 demonstrates the principle behind PRB. Fig. 2(a) shows the one-way beampatterns (Tx or Rx) of four consecutive lines using conventional DAS beamforming. The angular sampling in this example is twice that of

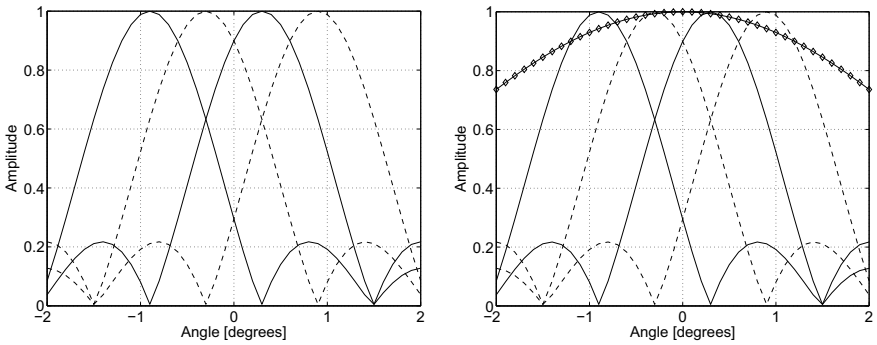


Fig. 2: (a) One-way beampatterns of four consecutive Tx (or Rx) beams (at the center frequency). The angular sampling in this example is twice of what is required by the Rayleigh criterion. (b) Beampatterns of one Tx beam (solid with diamonds) and four parallel Rx beams (solid and dashed).

the Rayleigh criterion. Fig. 2(b) shows the corresponding one-way Tx and Rx beampatterns using PRB. Here only one wide beam is transmitted, and four parallel beams are used on reception, increasing the frame rate by four. The wide Tx beam will however affect resolution. In Section 3 we show that by using a MV beamformer on reception, the loss in resolution caused by wider transmit beams can be compensated for. Note that near the focal point of the transducer, the two-way beampattern will be shifted with respect to the peak of the receive beam. This effect can be compensated for by steering off axis [16]. We have neglected this effect in our simulations, as the uncorrected examples are sufficient to demonstrate the principle.

We achieve increased depth of penetration in attenuating media if the frequency of the transmitted beam is reduced. Lowering the transmit frequency will however affect the resolution of the image. In the following section we show that the lateral resolution lost by lowering the transmit frequency can be compensated for using an MV beamformer on reception. The axial resolution is however also affected, but is usually much better than the lateral resolution, so a reduction can be afforded.

3 Results

We have simulated three different phantoms using Field II [17]: A phantom of wire targets located at 30-80 mm depth, and two cyst phantoms with four strong point scatterers embedded (which we refer to as *cyst phantom 1* and *2*). The cyst phantoms were generated by placing 135,000 equal amplitude point scatterers randomly within a region of 100x10x50 mm for cyst phantom 1, and 300,000

scatterers in a region of 220x10x110 mm for cyst phantom 2. The number of scatterers per resolution cell was more than 10, which is recommended to simulate speckle [18]. The distribution of the scatterer positions was uniform, and within the region of the cysts we set the amplitudes to zero. Four strong point scatterers were placed behind the cysts. In phantom 1 the cyst was centered at 35 mm and had a diameter of 6 mm. The point scatterers were separated by 3 mm laterally. In phantom 2 the cyst was centered at 95 mm and had a diameter of 10 mm. The scatterers were separated by 6 mm. In the simulations of cyst phantom 2 we included attenuation of 0.5 dB/cm/MHz. For all simulations we added white, Gaussian noise corresponding to approximately 60 dB SNR with respect to the strongest reflector in the image.

We have simulated 3.5 MHz transducers with approximately 85 percent -6 dB fractional bandwidth. The width of the arrays was either 9.25 mm or 18.5 mm and the height was 10 mm. Unless otherwise stated, the transducer was excited by 2 periods of a sine wave at the center frequency. We applied fixed focus on transmission and dynamic focus on reception. The received signals were sampled at 20 MHz and filtered by a 1.5-5.5 MHz bandpass filter. We added noise after filtering. For both beamformers delays were implemented by upsampling to 160 MHz and selecting the sample closest to the theoretically predicted delay. We computed the analytic signals by applying the Hilbert transform to the channel outputs, thus allowing complex solutions in the MV beamformer. The difference in the two beamformers was in the combination of the channels: For DAS, we applied either a rectangular or a Hamming window to the delayed channels before summation. For MV, weights were calculated using (7) and applied using (10). For all MV results we used diagonal loading parameter $\Delta = 1/(100L)$.

3.1 Decreasing Transducer Size

We have simulated a 48 element, 9.25 mm transducer and a 96 element, 18.5 mm transducer running at 3.5 MHz. We used MV beamforming on the smallest transducer. Fig. 3 shows results for the wire targets, were Tx focus was 50 mm. Fig. 3(a) and (b) shows the 18.5 mm probe using DAS beamforming and rectangular and Hamming apodization, respectively. Fig. 3(c) shows the corresponding image using the 9.25 mm probe running MV with subarray length $L = 24$. In the simple point target scenario we obtain a good estimate of the covariance matrix by averaging in the spatial domain only, hence we used $K = 0$. We see that the resolution obtained by the 9.25 mm probe is better than the 18.5 mm. Fig. 4 shows the steered responses at depths 40 mm and 60 mm. For both depths we see that the beamwidth of the 9.25 mm probe is less than 1/4th of the 18.5 mm using a rectangular window, and that the sidelobes are comparable to the Hamming weighted DAS. We see a slight shift in the location of the peaks for the MV result at 60 mm, which due to energy leaking through the sidelobes of the smaller array.

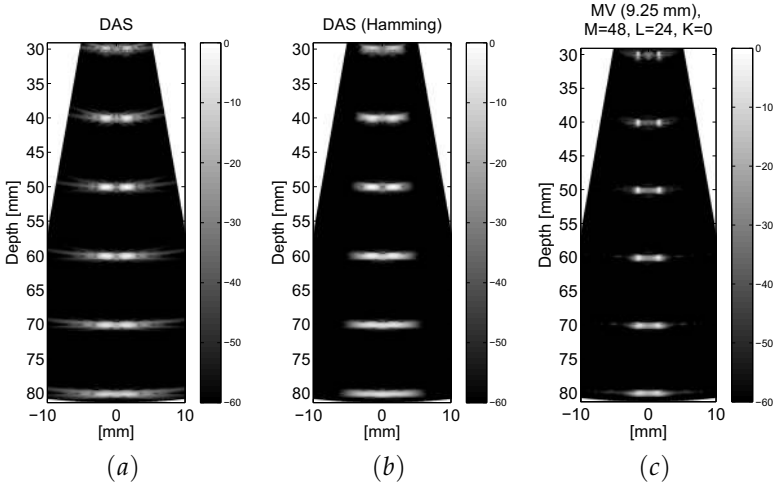


Fig. 3: Simulated wire targets using 3.5 MHz transducers. (a) DAS using 96 element, 18.5 mm transducer and rectangular window (b) DAS using 96 element, 18.5 mm transducer and Hamming window (c) MV using 48 element, 9.25 mm transducer ($M = 48$, $L = 24$, $K = 0$). Tx focus was 50 mm.

Fig. 5 shows results for cyst phantom 1, where Tx focus was 35 mm. Figs. 5(a)-(c) shows the 9.25 mm and 18.5 mm probes using DAS with different apodizations. Fig. 5(d) shows the 9.25 mm probe using MV beamforming with $L = 24$ and $K = 11$. Using $K = 11$ corresponds to temporal averaging over approximately the length of the transmit pulse. We see that the lateral resolution lost due to the reduced transducer size using DAS in Fig. 5(a), is compensated for by the MV beamformer in Fig. 5(d).

3.2 Higher Frame Rates by Combining Parallel Receive Beams with MV Beamforming

It has been common in ultrasound imaging to use transmit beams of comparable beamwidth to the receive beams. As frame rate is inverse proportional to the number of transmit beams required to scan an image, many scanners now use fewer transmit beams than receive beams. This is especially interesting in imaging of fast-moving structures, e.g. in cardiology. The price to pay for this is a requirement for several receive beams per transmit beam in order to sample the image properly. This requires parallel receive beamformers, e.g. 2-4 parallel beams. This can be found in many commercial ultrasound scanners today, and the number of parallel beams may increase even more with newer 3D scanners. With the MV beamformer, there is a potential for decreasing the number of

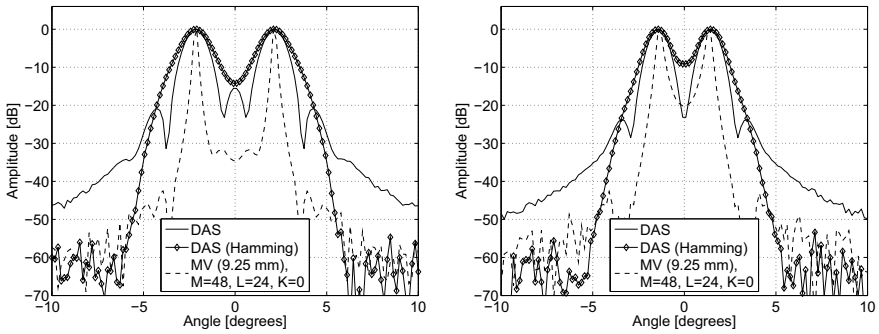


Fig. 4: Steered response of DAS and MV beamformers, using 3.5 MHz transducers at depths (a) 40 mm and (b) 60 mm. DAS was used with a 96 element, 18.5 mm transducer (using rectangular and Hamming window) and MV with a 48 element, 9.25 mm transducer ($L = 24$, $K = 0$). Tx focus was 50 mm.

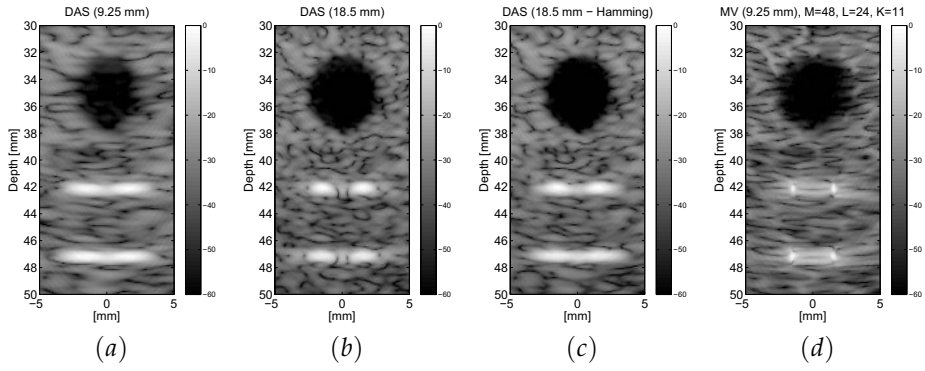


Fig. 5: Simulated cyst phantoms using 3.5 MHz transducers. (a) DAS with 48 element, 9.25 mm, transducer (rectangular window) (b) DAS with 96 element, 18.5 mm transducer (rectangular window) (c) DAS with 96 element, 18.5 mm transducer (Hamming window) (d) MV with 48 element, 9.25 mm transducer ($L = 24$ and $K = 11$). Tx focus was 35 mm.

transmit beams further and therefore we have simulated the 96 element, 3.5 MHz, 18.5 mm transducer using wider Tx beams and several parallel MV receive beams. As reference we formed non-parallel DAS images using 32 Tx and 32 Rx beams over a sector of 20 degrees. For MV we used only 1/4th of the aperture on transmission, giving a transmit beam approximately four times as wide as for DAS. The loss in amplitude due to fewer transmit elements was compensated for. We used 8 Tx beams and 4 parallel Rx beams per transmission. A sector image with DAS thus requires four times as many transmissions as the MV sector image. Hence, the parallel MV images may be formed four times faster than the non-parallel DAS; given that there is sufficient processing power in the receive beamformer for doing several MV beamformers in parallel. Fig. 6 shows the wire target images obtained from the different beamformers with 50 mm Tx focus, where we used $L = 24$ and $K = 0$ for the MV beamformer. We see that the wires are better resolved for MV despite the wide transmit beam. Fig. 7 shows the steered responses at 40 mm and 60 mm. We see that the mainlobe of the MV beamformer is narrower and the sidelobes are comparable to the Hamming weighted DAS.

Fig. 8 shows results for cyst phantom 1, where Tx focus was 35 mm. Here we have included DAS using parallel receive beams to demonstrate how the wide transmit beams degrade the image. Fig. 8(a) shows DAS using PRB. Compared to Fig. 8(b), which uses four times the number of Tx beams, we see that the cyst appears smaller and that the resolution of the point scatterers is lower. By using MV and PRB, which is shown in Fig. 8(d) with $L = 24$ and $K = 11$, we see that the resolution lost by the wide transmit beam is compensated for: The cyst is about the same size as in Fig. 8(b), and the point scatterers are better defined.

Note that since we only use 8 Tx beams in Fig. 6(c) and Fig. 8(d) versus 32 Tx beams in Figs. 6(a) and (b), and Figs. 8(b) and (c), respectively, the images can be formed four times as fast, given a real-time implementation of the MV method. Hence, we can increase the frame rate by a factor four, without sacrificing image quality. We discuss the real-time aspects of the method in Section 4.

To gain further insight into MV beamforming it may be interesting to compare parallel and non-parallel MV beamforming for the simulations in the previous example. In Fig. 9 we have included examples of receive beampatterns for the two cases. The beampatterns were calculated in the same way as the ones in Fig. 1 The focal point of the receive beam was at the edge of the cyst. Fig. 9(a) shows the MV beampattern using PRB, and Fig. 9(b) shows the MV beampattern where the full array was used on transmission. The dashed lines show the Tx beampatterns at the center frequency. Note that the steering of the Tx beam is different in the two examples, as the Tx beam in Fig. 9(a) is at the center of four parallel receive beams. We see that the parallel MV beamformer must compensate for the wide mainlobe and large first side lobe of the Tx beam, and thus has lower sensitivity than the non-parallel beamformer in these regions (see angles from 6 to 8 degrees and from 10 to 14 degrees). We also see that the non-parallel MV beamformer compensates for the first sidelobe

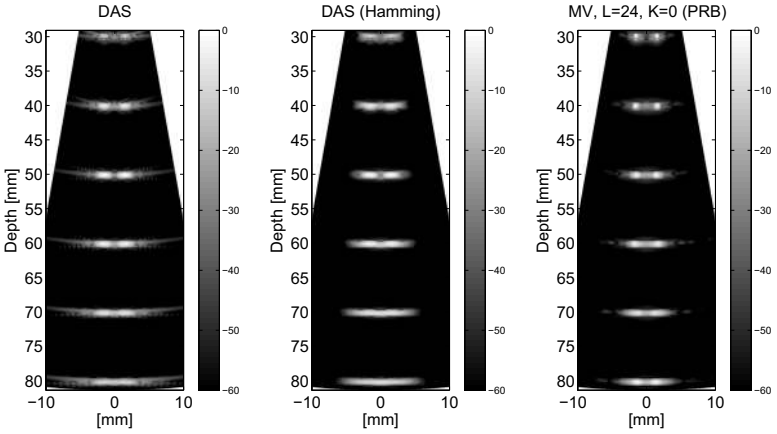


Fig. 6: Simulated wire targets using a 18.5 mm, 96 elements, 3.5 MHz transducer. (a) DAS using 32 Tx/Rx lines (rectangular window) (b) DAS using 32 Tx/Rx lines (Hamming window) (c) MV using 8 Tx and 32 Rx lines ($L = 24$, $K = 0$). Tx focus was 50 mm.

of the narrower Tx beam used in the example in Fig. 9(b) (see at approximately 6.5 degree angle in Fig. 9(b)). Note that although the MV beamformer can to some extent compensate for the Tx beam, a more focused beam is an advantage as the MV beamformer then gets more freedom to suppress off-axis scattering.

3.3 Increased Depth of Penetration

Figs. 10 and 11 shows results from the 96 element, 3.5 MHz, 18.5 mm transducer for cyst phantom 2, which has attenuation included. The transducer was either excited by 1 period of a 1.75 MHz sine wave, or 2 periods at 3.5 MHz. Due to the impulse response of the transducer the effective center frequency was approximately 2.5 MHz when the transducer was excited at 1.75 MHz. We compensated for the reduced gain due to the lower sensitivity of the transducer when transmitting at 1.75 MHz. Tx focus was 100 mm. In Fig. 10(a) and (b) we have used DAS and compared excitation at 1.75 MHz and 3.5 MHz, respectively. We see that at 1.75 MHz excitation, the reflected signals are stronger due to the frequency dependent attenuation. In Fig. 10(d) we have compensated for the propagation loss of DAS in Fig. 10(b) to get a fair comparison of the resolution. We see that the resolution of the 1.75 MHz DAS is poorer than for 3.5 MHz. In Fig. 10(c) we have simulated 1.75 MHz excitation and MV receive beamforming with $L = 40$ and $K = 11$. We see that the resolution lost by the lower transmission frequency in Fig. 10(a) is compensated for. Compared to Fig. 10(d) the lateral resolution is about the same and the point scatterers are

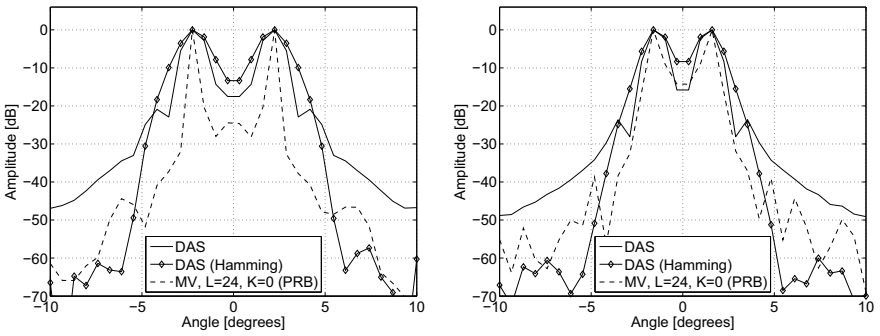


Fig. 7: Steered response of DAS and MV beamformers using a 18.5 mm, 96 elements, 3.5 MHz transducer at depths (a) 40 mm and (b) 60 mm. DAS used 32 Tx/Rx lines and MV used 8 Tx and 32 Rx lines, $L = 24$ and $K = 0$. Tx focus was 50 mm.

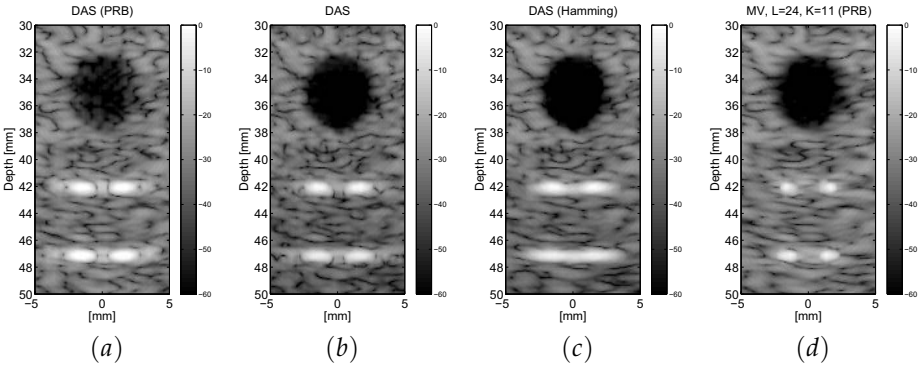
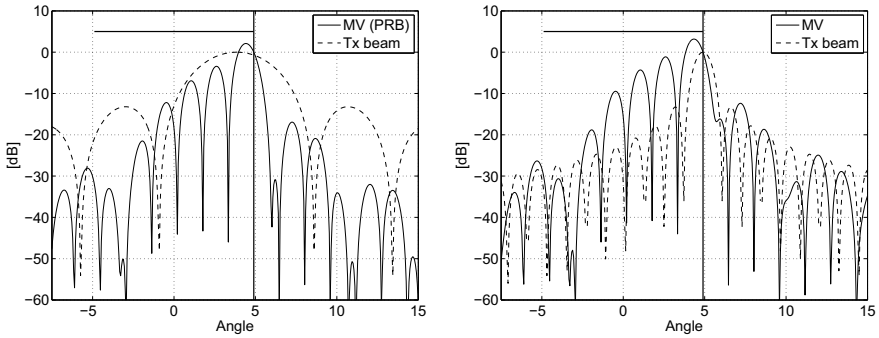


Fig. 8: Simulated cyst phantoms using a 18.5 mm, 96 element, 3.5 MHz transducers. (a) DAS (PRB, rectangular window), (b) DAS (rectangular window) (c) DAS (Hamming window) (d) MV (PRB, $L = 24$ and $K = 11$). Tx focus was 35 mm.



(a) (b)

Fig. 9: Beampatterns of MV beamformers. (a) Parallel MV beamformer (b) Non-parallel MV beamformer. The dashed line is the beampattern of the Tx beam at the center frequency. The solid vertical lines indicates the steering angle, and the horizontal line indicates the cyst.

much better defined. However, we see some distortion of the edges of the cyst. Fig. 11 shows the steered responses at 107 mm depth. We see that the peak amplitude is about 11 dB lower for the 3.5 MHz transmission resulting in 11 dB lower SNR. We also see that the MV beamformer produces narrower peaks, even though the transmission frequency is lower.

4 Discussion

The receive beampatterns in Fig. 1 are the key to understanding the improved performance of the MV beamformer. The echos from the steering direction are passed with unit gain (0 dB), and the mainlobe width and sidelobe levels are adapted to minimize off-axis energy. The mainlobes of the MV receive beams in these examples are asymmetric and their maximums are shifted away from the steering angle. Although this will result in two-way beampatterns that are more sensitive outside of the steering direction, the maximums are only placed at these locations because there is no reflected energy there. In Fig. 1(a) the peak is between the two scatterers and in Fig. 1(b) the peak is inside the cyst. Hence, geometric distortion will not occur despite the shifted peaks. We also see that the MV receive beams have low sensitivity where off-axis energy is present. For every point in the image new weights, and hence new receive beams, are calculated. In contrast to DAS, where beampatterns describe the imaging capabilities of the system, the beampatterns of the MV beamformer only shows how off-axis energy is suppressed for a specific sampled wavefield at a specific point. Hence, the beampatterns must always be considered in conjunction with the recorded data.

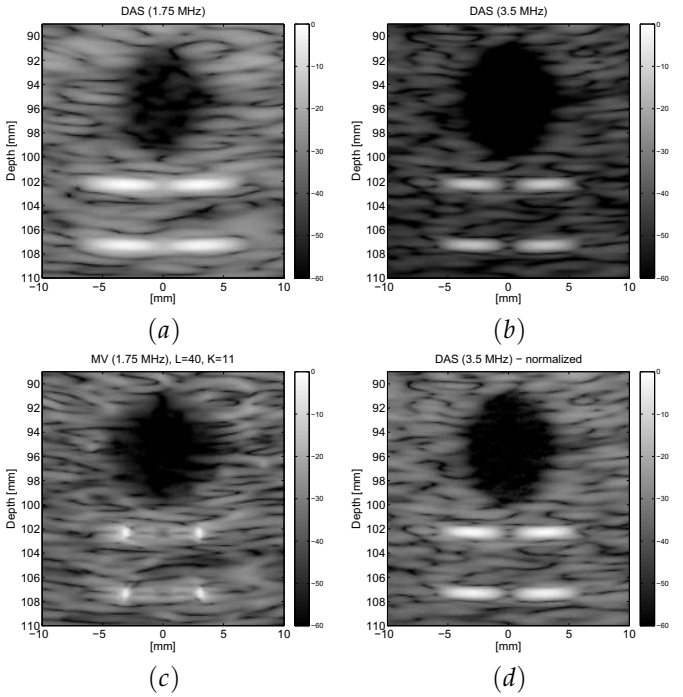


Fig. 10: Simulated cyst phantoms using a 18.5 mm, 96 element transducer. (a) DAS (1.75 MHz), (b) DAS (3.5 MHz) and (c) MV (1.75 MHz, $L = 40$ and $K = 11$). (d) DAS (same as (b), but the propagation loss was compensated for). Tx focus was 100 mm.

From Figs. 3 and 6 we see that for the simple phantoms consisting of wire targets, MV performs better than DAS despite that a smaller transducer or a wider transmit beam is used. The strong point scatterers in the cyst phantoms in Figs. 5, 8 and 10 are also much better defined using MV. For the cysts in these figures, we get similar definition as for DAS. When imaging cyst boundaries, off-axis energy is scattered from a range of angles rather than from isolated points, and we have less freedom to suppress interference. Hence, we cannot expect as much improvement as for the point scatterers.

For the 9.25 mm transducer using subarray length $L = M/2$ in Figs. 3-5 we achieve very good resolution of the wire targets and point scatterers, and a similar definition of the cyst as for the 18.5 mm transducer. However, we see some edge distortion on the cyst in Fig. 5(d). To compensate for the reduced aperture size, we set the subarray size to half of the aperture to get as good lateral resolution as possible. This resulted in larger speckle variance and inconsistent edges along the cyst. Note that the cyst edge distortion is reduced in the example in Fig. 8(d) where we could choose a more conservative subarray

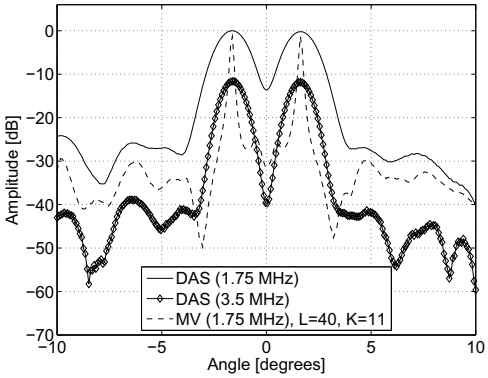


Fig. 11: Steered responses of DAS and MV ($L = 40$, $K = 11$) beamformers using a 18.5 mm, 96 element transducer at 107 mm depth with 0.5 dB/cm/MHz attenuation. Tx focus was 100 mm.

size ($L = M/4$).

Using parallel receive beams in Fig. 6(c) and Fig. 8(d) we needed a robust amplitude estimate due to the relatively few scan lines, and used $L = M/4$ for both the wire targets and the cyst phantom. We see from the figures that the resolution of the wire targets and point scatterers is better for the parallel MV than for the non-parallel DAS. The cyst definition is about the same. In these examples we have used only 8 Tx beams for the MV results, versus 32 Tx beams for DAS. Hence, the parallel MV images may be formed four times as fast. We have thus demonstrated that by combining PRB and MV beamforming, we can increase frame rate without sacrificing image quality. Both parallel receive beamforming and MV beamforming will, however, increase the computational load of the system. To achieve an increase in frame rate which scales with the number of parallel beams, we must be able to process the parallel beams equally fast as a single beam in a non-parallel system. With the added complexity of the MV method, the combination certainly presents a challenge for a real-time system. But since the parallel beams operate on nearly the same data, there could be a potential saving in computations by exploiting these similarities. For instance, it may be possible to use the same covariance matrix estimate for all parallel receive beams. We may account for the differences in angle by using a complex steering vector in (7). We further discuss the real-time aspects of the MV method below.

To finish the discussion on the choice of parameters, we found that reducing the transmit frequency from 3.5 MHz to 1.75 MHz in Fig. 10 required longer subarrays to match the lateral resolution of DAS at 3.5 MHz, and we used $L = 40$. This resulted in some distortion of the edges of the cyst. We see that the point scatterers are much better defined.

The MV beamformer requires significantly more computations than DAS. The delay-step is common to both beamformers, so the difference lies in finding the aperture weights adaptively. The complexity of applying the weights and summing, which are the only extra operations required by DAS, is linear with the number of elements ($O(M)$). The matrix inversion needed to find the aperture weights for the MV beamformer has complexity proportional to the subarray size, $O(L^3)$. Hence, using smaller subarrays is more efficient, but reduces the resolution. With the improvements shown in this paper we have assumed that the MV algorithm can be run in real-time. However, an algorithm with computational complexity of $O(L^3)$ may not be suitable for real-time implementation at present. Further research should be directed at making approximations that can run in real-time. One useful observation is that for spatially stationary signals the covariance matrix becomes Toeplitz in theory. Such a matrix can be inverted in $O(L^2)$ operations. In our application it may be a good approximation to assume spatial stationarity as most of the energy is scattered from angles close to the steering angle. Thus we can force a Toeplitz structure by averaging the entries in the covariance matrix corresponding to the same lags. Whether this approximation will give acceptable performance remains to be seen. Another approach, suggested by Vignon and Burcher [8], is to use a number of predefined windows which fulfill the constraint in (6). The window that minimizes $\mathbf{w}[n]^H\mathbf{R}[n]\mathbf{w}[n]$ is chosen for sample n. If the number of predetermined windows is much smaller than M, the overhead required to find the aperture weights is small. The method will have approximately linear complexity and will be suitable for a real-time environment. Our initial studies of this approach show that we can achieve significant improvement over DAS [19].

In [6] we demonstrated increased resolution by applying the MV beamformer to ultrasound imaging. In this paper we have utilized the high resolution properties for other enhancements with the objective of matching the resolution of DAS. However, we can imagine combinations of the enhancements, e.g. a slight increase in both resolution and penetration depth by selecting a proper transmit frequency. It is also worth noting that the reduced channel count for the examples in Figs. 3-5 may allow for a cheaper system.

5 Conclusion

We have demonstrated that we can utilize the minimum variance beamformer to reduce transducer size, increase frame rates or increase penetration depth in medical ultrasound, without sacrificing image quality compared to delay-and-sum. We have shown that the minimum variance beamformer can compensate for the resolution lost by using smaller transducers, wider transmit beams or a lower transmit frequency, thus the increase in processing power may be well worth the effort. Future work will be directed at assessing the performance in a real-time setting and with clinical objects.

References

[1] J. A. Mann and W. F. Walker, "A constrained adaptive beamformer for medical ultrasound: Initial results," *Ultrasonics Symposium, 2002. Proceedings. 2002 IEEE*, vol. 2, pp. 1807–1810, October 2002.

[2] M. Sasso and C. Cohen-Bacrie, "Medical ultrasound imaging using the fully adaptive beamformer," *Acoustics, Speech and Signal Processing, 2005. Proceedings (ICASSP '05). IEEE International Conference on*, vol. 2, pp. 489–492, March 2005.

[3] J.-F. Synnevåg, A. Austeng, and S. Holm, "Minimum variance adaptive beamforming applied to medical ultrasound imaging," *Proc. IEEE Ultrasonics Symposium*, vol. 2, pp. 1199–1202, Sept. 2005.

[4] F. Viola and W. F. Walker, "Adaptive signal processing in medical ultrasound beamforming," *Proc. IEEE Ultrasonics Symposium*, vol. 4, pp. 1980–1983, Sept. 2005.

[5] Z. Wang, J. Li, and R. Wu, "Time-delay- and time-reversal-based robust Capon beamformers for ultrasound imaging," *IEEE Transactions on Medical Imaging*, vol. 24, pp. 1308–1322, Oct. 2005.

[6] J.-F. Synnevåg, A. Austeng, and S. Holm, "Adaptive beamforming applied to medical ultrasound imaging," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 54, no. 8, pp. 1606–1613, Aug. 2007.

[7] I. K. Holfort, F. Gran, and J. A. Jensen, "Minimum variance beamforming for high frame-rate ultrasound imaging," *Proc. IEEE Ultrasonics Symposium*, pp. 1541–1544, Oct. 2007.

[8] F. Vignon and M. Burcher, "Capon beamforming in medical ultrasound imaging with focused beams," *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 55, no. 3, pp. 619–628, March 2008.

[9] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, pp. 1408–1418, August 1969.

[10] O. L. Frost, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE*, vol. 60, pp. 926–935, August 1972.

[11] J. E. Evans, J. R. Johnson, and D. F. Sun, "High resolution angular spectrum estimation techniques for terrain scattering analysis and angle of arrival estimation," *Proc. 1st ASSP Workshop Spectral Estimation, Hamilton, Ont., Canada*, pp. 134–139, 1981.

[12] J.-F. Synnevåg, C. I. C. Nilsen, and S. Holm, "Speckle statistics in adaptive beamforming," *Proc. IEEE Ultrasonics Symposium*, pp. 1545–1548, Oct 2007.

[13] J. Li, P. Stoica, and Z. Wang, "On robust Capon beamforming and diagonal loading," *IEEE Transactions on Signal Processing*, vol. 51, no. 7, pp. 1702–1715, July 2003.

[14] T.-J. Shan, M. Wax, and T. Kailath, "On spatial smoothing for direction-of-arrival estimation of coherent signals," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 4, pp. 806–811, August 1985.

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

[16] T. Hergum, T. Bjastad, K. Kristoffersen, and H. Torp, "Parallel beamforming using synthetic transmit beams," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 54, no. 2, pp. 271–280, February 2007.

[17] J. A. Jensen, "Field: A program for simulating ultrasound systems," *Medical & Biological Engineering & Computing*, vol. 34, pp. 351–353, 1996.

[18] R. Wagner, M. Insana, and S. Smith, "Fundamental correlation lengths of coherent speckle in medical ultrasonic images," *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on*, vol. 35, no. 1, pp. 34–44, Jan 1988.

[19] J.-F. Synnevåg, S. Holm, and A. Austeng, "A low-complexity data-dependent beamformer," *to appear in Proc. IEEE Ultrasonics Symposium*, 2008.

IV

A Low Complexity Data-Dependent Beamformer

Johan-Fredrik Synnevåg, Andreas Austeng, and Sverre Holm

Abstract

The classical problem of choosing apodization functions for a beamformer involves a trade-off between mainlobe width and sidelobe level, *i.e.* a trade-off between resolution and contrast. To avoid this trade-off it has been suggested to apply adaptive beamforming, such as minimum variance beamforming, to medical ultrasound imaging. This has been an active topic of research in medical ultrasound imaging in the recent years, and several authors have demonstrated significant improvements in image resolution. However, the improvement comes at a considerable cost. Where the complexity of a conventional beamformer is linear with the number of elements ($O(M)$), the complexity of the minimum variance beamformer is up to $O(M^3)$. In this paper we have applied a method – based on an idea by Vignon and Burcher [1] – which is data-adaptive, but selects the apodization function between a number of predefined windows, giving linear complexity. In the proposed method we select an apodization function for each depth along a scan line based on the optimality criterion of the minimum variance beamformer. But unlike the minimum variance beamformer, which has an infinite solution space, we limit the number of possible outcomes to a set of predefined windows. The complexity of the method is then only P times that of the conventional, where P is the number of predefined windows. The suggested method gives significant improvement in image resolution at a low cost. The method is robust, can handle coherent targets and is easy to implement. It may also be used as a classifier as the selected window gives information about the object being imaged. We have applied the method to simulated data of wire targets and a cyst phantom, and to experimental RF data from a heart phantom using $P = 4$ and $P = 12$. The results show significant improvement in image resolution compared to delay-and-sum.

1 Introduction

The standard technique used to form images in medical ultrasound imaging is by delay-and-sum (DAS) beamforming. A focused beam is transmitted over a range of angles, and a receive (Rx) beam is dynamically focused by delaying and summing the recorded signals from the different channels. The mainlobe width and sidelobe level of the beams can be controlled by applying weights to the different channels. The weights form an apodization function or a window. In this paper we will focus on the receiver apodization, only. The fundamental trade-off when designing windows, is between mainlobe width and sidelobe level. A reduction in sidelobe level results in increased mainlobe width, trading off contrast for resolution. In practical systems, a window gradually falling towards the edges is often used close to the transducer to achieve low sidelobes, and a rectangular window is often used at larger depths to achieve good resolution and maximum sensitivity and penetration.

The transducer geometry, number of channels, transmit frequency and apodization function determine the imaging capabilities of a system, and is characterized by the *beampattern*. The ultimate goal is a beampattern with a narrow mainlobe and low sidelobes, but as mentioned, these are contradicting requirements. Instead, we may use adaptive beamforming, like the minimum variance (MV) beamformer, which takes the actual recorded wavefield into account to find the optimal apodization function. The key to the improved performance of the MV beamformer is that it will vary the apodization functions across the image, based on the recorded data. At a specific depth, the resulting beampatterns may contain large sidelobes in regions where off-axis scattering is low and be less sensitive where off-axis scattering is present.

Adaptive beamforming has been an active research topic in medical ultrasound imaging in recent years, and several authors have shown significant improvement in image quality [1–9]. See [9] for a review of the different contributions. All authors have demonstrated improvements in image quality. However, the improvements come at a cost. The complexity of applying weights and summing in conventional delay-and-sum beamforming, is linear with the number of elements ($O(M)$). The complexity of the MV beamformer is up to $O(M^3)$, following from the inversion of the spatial covariance matrix. So although the performance is good, the computational overhead is significant. In this paper, we propose an adaptive beamformer – based on and idea by Vignon and Burcher [1] – which uses several predefined windows and use the minimum variance optimization criterion to decide which window to use at a specific point in the image. The computational overhead scales with the number of predetermined windows compared to DAS. Hence, there is little overhead if we only use a few windows.

To fully understand the proposed method, it is important to understand the basics of the MV beamformer. We have organized the paper as follows: In the next section we first review the MV beamformer and show how we have

implemented it in medical ultrasound, before we present the proposed method. In section 3 we show results from simulated data of point targets and a cyst phantom, and from experimental RF data of a heart phantom. We discuss our findings in section 4 and draw conclusions in section 5.

2 Methods

2.1 The Minimum Variance Beamformer

We assume a transducer consisting of M elements, where $x_m[n]$ is the sampled output from element m. The output of a general beamformer can be written as:

$$z[n] = \sum_{m=0}^{M-1} w_m[n] x_m[n - \Delta_m[n]], \tag{1}$$

where $w_m[n]$ is a (complex) weight and $\Delta_m[n]$ is the delay applied to element m. In matrix form, $z[n]$ can be written as:

$$z[n] = \mathbf{w}^H[n]\mathbf{X}[n], \tag{2}$$

where

$$\mathbf{w}[n] = \begin{bmatrix} w_0^*[n] \\ w_1^*[n] \\ \vdots \\ w_{M-1}^*[n] \end{bmatrix}, \tag{3}$$

$$\mathbf{X}[n] = \begin{bmatrix} x_0[n - \Delta[n]] \\ x_1[n - \Delta[n]] \\ \vdots \\ x_{M-1}[n - \Delta[n]] \end{bmatrix}, \tag{4}$$

and * denotes the complex conjugate. The MV beamformer finds the optimal apodization $\mathbf{w}[n]$ by solving the optimization problem [10]:

$$\begin{aligned} \min_{\mathbf{w}[n]} E\left\{|z[n]|^2\right\} &= \mathbf{w}^H[n]\mathbf{R}[n]\mathbf{w}[n] \\ \text{subject to } \mathbf{w}^H[n]\mathbf{a} &= 1, \end{aligned} \tag{5}$$

where $E\{\cdot\}$ denotes the expectation, $\mathbf{R}[n] = E\{\mathbf{X}\mathbf{X}^H\}$ is the *spatial covariance matrix* and $\mathbf{a}$ is the *steering vector*. Since the data has been delayed, $\mathbf{a}$ is simply a vector of ones. The constraint corresponds to requiring $\sum_{m=0}^{M-1} w_m[n] = 1$. The solution to (5) is given by:

$$\mathbf{w}[n] = \frac{\mathbf{R}^{-1}[n]\mathbf{a}}{\mathbf{a}^H\mathbf{R}^{-1}[n]\mathbf{a}}. \tag{6}$$

Hence, the optimal apodization function $\mathbf{w}[n]$ is dependent on the inverse of $\mathbf{R}[n]$. The inversion requires in the order of $O(M^3)$ operations.

In practice the spatial covariance matrix must be estimated. As the returning echos from different targets will be correlated, measures must be taken to avoid signal cancellation [11]. We have successfully applied subarray averaging, meaning that the estimate is found by averaging the covariance matrices of different subarrays [7,12]. Averaging purely in the spatial dimension gives good results for simple images of point targets. However, for collections of scatterers that are much smaller than the resolution of the imaging system (leading to speckle), we have shown that averaging purely in the spatial dimension gives speckle statistics that are very different from DAS beamforming. The speckle appears less homogeneous and has a lower average intensity. By averaging in the temporal dimension (corresponding to depth), as well as along the transducer, similar speckle statistics are retained [13]. The general estimate for overlapping subarrays of length L, and averaging over $2K+1$ temporal samples becomes:

$$\hat{\mathbf{R}}[n] = \frac{1}{(2K+1)(M-L+1)} \sum_{k=-K}^{K} \sum_{l=0}^{M-L} \mathbf{X}_l[n+k]\mathbf{X}_l^H[n+k], \tag{7}$$

where

$$\mathbf{X}_l[n] = \begin{bmatrix} x_l[n-\Delta_l[n]] \\ x_{l+1}[n-\Delta_{l+1}[n]] \\ \vdots \\ x_{l+L-1}[n-\Delta_{l+L-1}[n]] \end{bmatrix}. \tag{8}$$

The choice of L is a trade-off between resolution and robustness of the estimate. Smaller L gives fewer degrees of freedom, and hence poorer resolution, but increases robustness towards signal cancellation. See [7] for a discussion of the choice of parameters and for results using different subarray lengths. Regarding the choice of K, we have achieved similar speckle statistics as DAS by averaging over a time-gate that is comparable to the length of the transmit pulse. See [13] for results using different values of K.

Another measure to increase robustness of the method is to apply *diagonal loading* to the covariance estimate before inversion. A small regularization term is added to the diagonal of the covariance matrix, such that $\mathbf{R}[n]$ is replaced by $\mathbf{R}[n] + \epsilon\mathbf{I}$ in (6). We have chosen ϵ to be proportional to the power of the observations:

$$\epsilon = \Delta \cdot \mathrm{tr}\left\{\mathbf{R}[n]\right\}, \tag{9}$$

where Δ typically is much less than $1/L$. See [7] for results using different Δ.

Even though the covariance matrix is estimated by averaging over many temporal samples (if $K > 0$), we apply the resulting weights to the current sample only. The MV amplitude estimate is given by:

$$\hat{z}[n] = \frac{1}{M-L+1} \sum_{l=0}^{M-L} \mathbf{w}[n]^H \bar{\mathbf{X}}_l[n]. \tag{10}$$

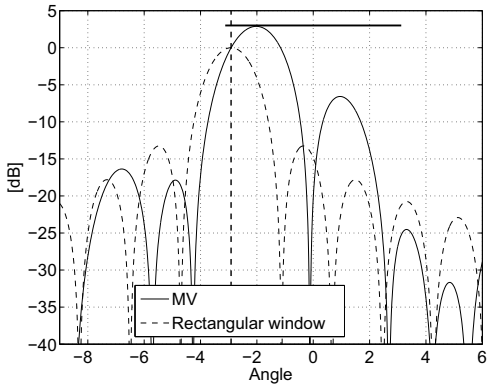


Fig. 1: Example of beampattern from a MV weighted 64 element, 2.5 MHz, 19.7 mm transducer (solid line). The dashed line shows the corresponding beampattern for a rectangular window. The horizontal line indicates a cyst.

Hence, no temporal smoothing occurs as a result of the temporal averaging.

To aid the window design of the proposed method it is instructive to look at the response of the MV weights in a typical medical ultrasound scenario. Fig. 1 shows the beampattern of a 64 element, 2.5 MHz, MV weighted array when focusing close to the edge of an anechoic cyst (see Fig. 8 for examples of images). The beampattern of a uniformly weighted array has been included for reference. The vertical dotted line indicates the steering angle. We see that the gain of the MV beamformer is 0 dB at this angle, following from the constraint in (5). The thick horizontal line indicates the cyst. We see that as there is no scattered energy inside the cyst, the MV solution allows large gains in this region, which in this example results in an asymmetric mainlobe. We also see that a large sidelobe is present inside the cyst (see at approximately 1 degree). The asymmetric mainlobe gives better suppression of the scattered energy outside the cyst, which in turn will give better definition of the edge.

2.2 Low Complexity Adaptive Beamforming

We now present the proposed method, which we have termed *low complexity adaptive* (LCA) beamforming. To reduce the complexity compared to MV, we limit the solution space of the MV beamformer by defining P predetermined windows, $\{\bar{\mathbf{w}}_0, \cdots, \bar{\mathbf{w}}_{P-1}\}$. We make sure that the windows fulfill the constraint in (5) by applying a proper normalization (*i.e.* we divide by the sum of the coefficients). For every sample along a scan line we apply one of the predefined windows,

$$\mathbf{w}[n] = \bar{\mathbf{w}}_p, \tag{11}$$

where $\bar{\mathbf{w}}_p$ is the pth predefined window. Which window to use is found by minimizing the cost function:

$$\min_p E\left\{\left|\bar{\mathbf{w}}_p^H \mathbf{X}[n]\right|^2\right\}, p = \{0, \cdots, P-1\}. \tag{12}$$

The solution to (12) is found by applying each of the P predefined windows to $\mathbf{X}[n]$ and choose the one that minimizes $E\{|\bar{\mathbf{w}}_p^H \mathbf{X}[n]|^2\}$. Hence, we avoid the inversion in (6). In practice, $E\{|\bar{\mathbf{w}}_p^H \mathbf{X}[n]|^2\}$ must be estimated. By defining:

$$z_p[n] = \bar{\mathbf{w}}_p^H \mathbf{X}[n], \tag{13}$$

we can estimate the variance as:

$$\hat{\sigma}^2_{z_p}[n] = \frac{1}{2K+1} \sum_{k=-K}^{K} \left|z_p[n+k]\right|^2. \tag{14}$$

For a single sample estimate ($K = 0$) we simply choose the window that minimizes $|\bar{\mathbf{w}}_p^H \mathbf{X}[n]|$. Compared to DAS we need to apply P windows instead of one, which requires P times the number of computations. For $K > 0$, the same complexity is achieved if we store the $2K+1$ last values of $z_p[n]$ for all p. We update the estimate in (14) as:

$$\begin{aligned}\hat{\sigma}^2_{z_p}[n+1] = \hat{\sigma}^2_{z_p}[n] - \\ \tfrac{1}{2K+1}\left(\left|z_p[n-K]\right|^2 - \left|z_p[n+K+1]\right|^2\right).\end{aligned} \tag{15}$$

In the next section we will see that, as for the MV beamformer, the single sample estimate works well for point targets, but to retain similar speckle as DAS, temporal averaging is required.

An example which illustrates the method is shown in Fig. 2, where we have imaged two closely spaced wire targets (see Fig. 6 for examples of full images). In Fig. 2(a) we show the steered responses using DAS with rectangular and Hamming apodization. In Fig. 2(b) we have chosen between the two windows, based on the optimization in (12) for every lateral point on the curve. DAS with rectangular weighting has been included for reference. We see that the LCA beamformer gives the resolution of the rectangular window and the low sidelobe level of the Hamming window. By using more windows we can further improve these results.

The key to the improved performance of the LCA beamformer is the design of the predetermined windows. It is natural to include the classical windows which trade off mainlobe width for sidelobe level, as the rectangular and Kaiser windows. We may also push the resolution of the rectangular window by increasing the weighting towards the edges of the aperture, which in turn gives higher sidelobes. Fig. 3 shows examples of three different windows: A

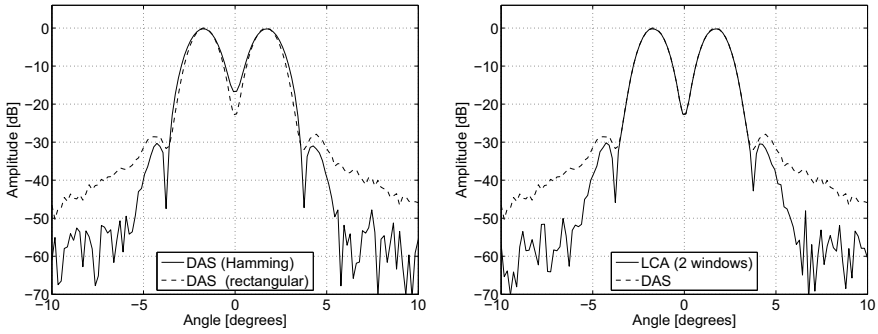


Fig. 2: Steered responses using a 64 element, 2.5 MHz, 19.7 mm transducer. (a) DAS with rectangular (dashed) and Hamming (solid) apodization (b) DAS with rectangular apodization (dashed) and LCA (solid) using two windows (rectangular and Hamming).

rectangular, Kaiser ($\beta = 2$, where β is the shape parameter of the Kaiser window) and "inverted" Kaiser ($\beta = 2$) window. The coefficients in the latter window is given by:

$$w_p[m] = \frac{1}{w_K[m]}, \tag{16}$$

where $w_K[m]$ is the mth coefficient of the Kaiser window. Fig. 4(a) shows the response of the inverted Kaiser window. We see that the mainlobe is narrower than for rectangular shading, but that the sidelobes have increased. We have also used our experience with the MV beamformer in the window design. We have designed windows with asymmetric responses to achieve good edge definition, based on the complex-valued MV solution in Fig. 1. The windows were designed by shifting the peak of the response of a Kaiser window:

$$w_p[m] = w_K[m]e^{-j2\pi\phi m}, \tag{17}$$

where ϕ determines the shift. The resulting window is normalized such that the gain is 0 dB for the steering angle. Fig. 4(b) shows an example of the asymmetric response of a Kaiser window ($\beta = 3$) which has been shifted using $\phi = 1/(2M)$ and normalized.

2.3 Robustness and Signal Cancellation

The two major problems with the MV beamformer is lack of robustness and signal cancellation. The method is less robust than DAS as the constraint in (5) only assures that reflections originating from the focal point of the receive beam are passed with gain one, others are suppressed. At high SNR, a target that

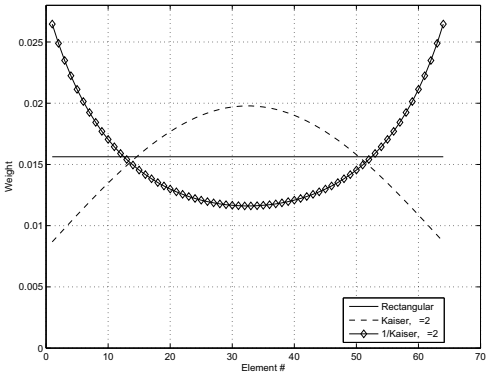
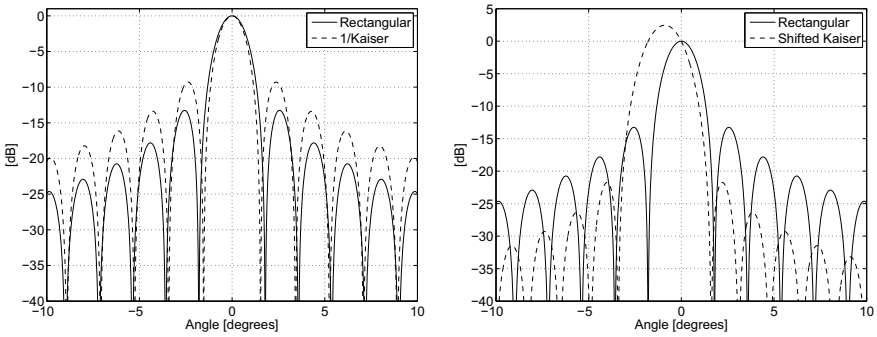


Fig. 3: Examples of window functions ($M = 64$): Rectangular (solid), Kaiser, $\beta = 2$ (dashed) and inverted Kaiser, $\beta = 2$ (solid with diamonds). The windows were normalized such that the sum of the weights is equal to one.



(a) (b)

Fig. 4: Examples of window responses. (a) Rectangular window (solid) and inverted Kaiser window, $\beta = 2$ (dashed). (b) Rectangular (solid) and shifted and normalized Kaiser window, $\beta = 3$ (dashed).

Table I: Windows used by the LCA beamformer.

#	Window	Parameters
1	Rectangular	
2	Kaiser	$\beta = 2$
3	Kaiser	$\beta = 3$
4	Kaiser	$\beta = 4$
5	"Inverted" Kaiser	$\beta = 2$
6	"Inverted" Kaiser	$\beta = 4$
7	Kaiser	$\beta = 3$, $\phi = 1/(2M)$
8	Kaiser	$\beta = 3$, $\phi = 3/(4M)$
9	Kaiser	$\beta = 3$, $\phi = 1/M$
10	Kaiser	$\beta = 3$, $\phi = -1/(2M)$
11	Kaiser	$\beta = 3$, $\phi = -3/(4M)$
12	Kaiser	$\beta = 3$, $\phi = -1/M$

is apparently slightly off axis may be severely attenuated, *e.g.* if the estimate of acoustic velocity is wrong. Also, since all reflections originate from the same source, the returning signals will be correlated, and signal cancellation may occur [11]. Both problems are mitigated by spatial smoothing or diagonal loading. In [7] we showed that the method could be made robust to errors in the assumed propagation velocity by decreasing L or increasing Δ – both at the expense of resolution.

The robustness of the LCA beamformer is given by the predetermined windows, and is controlled through the window design. The slope of the beampattern near the steering angle determines the attenuation of targets that are slightly off axis. By controlling the slope, we can control the maximum possible attenuation for a given error in acoustic velocity, or the attenuation of targets occurring between scan lines.

Signal cancellation may occur in the MV beamformer if the incoming signals are correlated. *E.g.* if the wavefield contains two correlated signals, the MV beampattern may contain two peaks with opposite phase, such that the signal in focus is cancelled by the off-axis interferer. The MV beamformer will favor such a solution, as the resulting variance is smaller than if a zero is placed at the location of the interference. As we have complete control over the predesigned windows, the LCA beamformer is not affected by signal correlation.

Fig. 5 shows an example where the lack of robustness affects the amplitude estimate of the MV beamformer. The figure shows the steered responses for two closely spaced targets using the MV beamformer with relatively long subarrays ($L = M/2$), together with LCA (using 12 windows) and DAS (see Fig. 6 for examples of images). We see that the amplitude estimate of the MV beamformer is approximately 9 dB lower than both LCA and DAS.

3 Results

3.1 Simulations

We have simulated two phantoms using Field II [14] – one consisting of pairs of wire targets and one cyst with point scatterers embedded. The wires were placed at 30-80 mm depth and separated by 3 mm laterally. The cyst phantom consisted of 300.000 point scatterers with equal reflectivity, randomly distributed over an area of 130x10x65 mm (with uniform distribution). The number of scatterers were more than 10 per resolution cell, which is recommended to simulate speckle [15]. The cyst was centered at 55 mm depth and had a diameter of 6 mm. Inside the cyst we set the amplitudes to zero. Pairs of bright point scatterers were placed in front of and behind the cyst, separated by 3 mm laterally. For all simulations we added white, Gaussian noise corresponding to approximately 40 dB SNR with respect to the strongest reflector in the image.

We simulated a 64 element, 2.5 MHz, 19.7 mm transducer with approximately 65 percent -6 dB fractional bandwidth. The height of the transducer was 10 mm. The transducer was driven by 2 cycles of a sine wave at the center frequency. We applied fixed focus on transmission, and dynamic focus on reception. The received signals were sampled at 20 MHz and filtered by a 1-4 MHz bandpass filter. We applied delays by upsampling the signals to 160 MHz and then selecting the sample closest to the theoretically predicted delay. We computed the analytic signals by applying the Hilbert transform to the delayed channels, thus allowing complex weights in the adaptive beamformers. For the conventional DAS beamformer, we applied either a rectangular or Hamming window and summed the channels. For the LCA beamformer, we used the 12 windows in Table I (unless otherwise stated) and found which window to apply by solving (12). For the MV beamformer we found the aperture shading using (6) and the amplitude estimate from (10). For all MV results we used diagonal loading parameter $\Delta = 1/100L$.

Fig. 6 shows images of the wire targets using the different beamformers. We used subarray length $L = 24$ for the MV beamformer. Tx focus was 50 mm. We see that both the LCA and MV beamformer gives significantly better resolution than DAS. Fig. 7 shows the steered responses at depths 50 and 80 mm depth. We see that both the sidelobe level and the mainlobe width is reduced by the adaptive beamformers. The adaptive beamformers do also resolve the targets at 80 mm, which the rectangular and Hamming weighted DAS beamformers are unable to.

Fig. 8 show images of the cyst phantom. Fig. 8(a) and (b) shows images using DAS with rectangular and Hamming apodization, respectively. We see that the cyst edges are better defined using the rectangular window and that the point targets are better resolved. Fig. 8(c) shows results from LCA using $K = 0$ and 4 predefined windows (windows 1, 4, 7 and 10 in Table I). We see that the definition of the cyst edges is slightly better than for the rectangular shading,

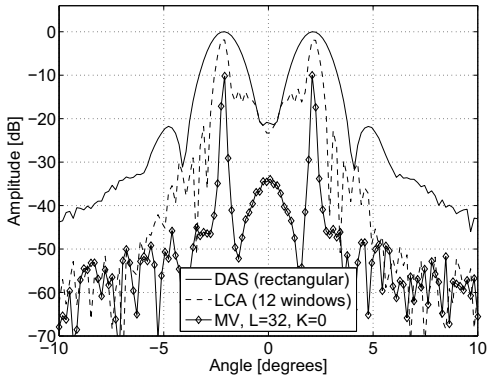


Fig. 5: Steered responses of DAS (solid), LCA (dashed), and MV (solid with diamonds) beamformers using a 64 element, 2.5 MHz, 19.7 mm transducer.

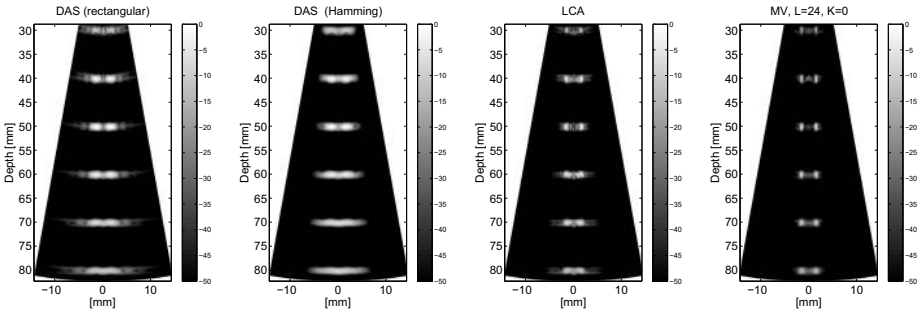


Fig. 6: Images of simulated data of wire targets using a 64 element, 2.5 MHz, 19.7 mm transducer (a) DAS (rectangular window) (b) DAS (Hamming window) (c) LCA (12 predefined windows) (d) MV ($L = 24$ and $K = 0$). Tx focus was 50 mm.

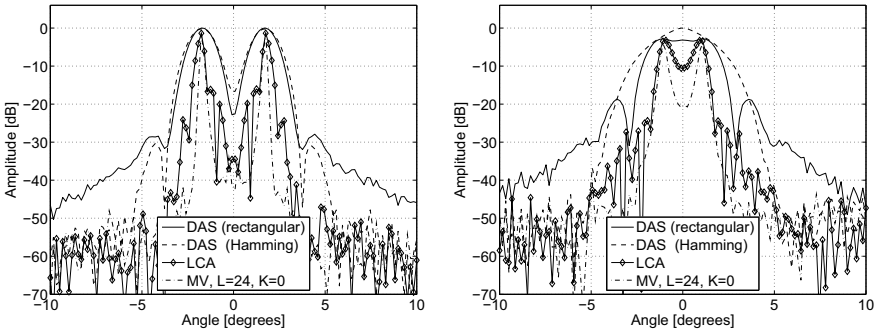


Fig. 7: Steered responses at different depths using a 64 element, 2.5 MHz, 19.7 mm transducer (a) 50 mm (b) 80 mm.

and that both the near and far point targets are better defined. The speckle appears similar, although with a slightly higher variance. Figs. 8(d) and (e) show LCA using 12 windows and $K = 0$ and $K = 15$, respectively. The cyst appears larger and closer to its actual size in these images. However, for $K = 0$, we see that the speckle appears less homogeneous and the overall brightness is reduced. Using $K = 15$, which means that we average over 31 samples when estimating the variance in (12), gives more homogeneous speckle. We see that the closer point targets are much better defined compared to the conventional beamformers, and that the far targets are slightly better resolved. Fig. 8(f) shows MV using $L = 24$. We see that the cyst definition is slightly poorer than LCA, but the point targets are better resolved, especially the deeper ones.

It is instructive to see which window the LCA beamformer chooses for the different points in the image. The choice of window may assist in classifying the object being imaged. Fig. 9 shows the chosen window type for each point in the image in Fig. 8(e). We have coded the image in three different shades and grouped the windows as follows: All windows with symmetric responses are coded in black (windows 1-6) and all windows with asymmetric responses which are shifted in the same direction are coded in gray (windows 7-9) and white (windows 10-12), respectively. We see that at close to and at the edges, asymmetric responses are chosen to get as good definition of the edges as possible. These windows are also chosen at the edges of the point scatterers. Between targets, windows with symmetric beampatterns are chosen, to suppress the off-axis targets on both sides.

3.2 Experimental Data

We have applied the different beamformers to an experimental RF data set from a heart phantom, obtained from the Biomedical Ultrasound Laboratory,

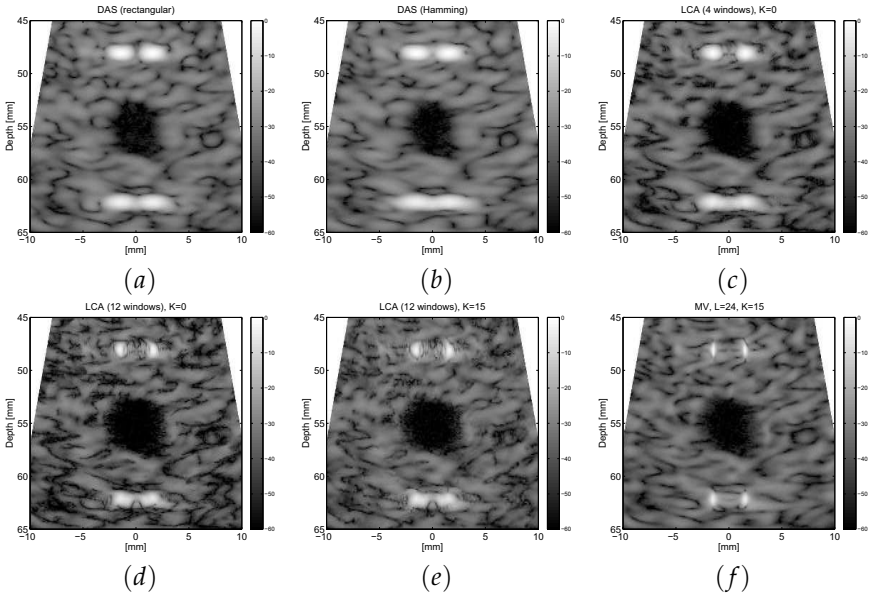


Fig. 8: Images from a simulated cyst phantom using a 64 element, 2.5 MHz, 19.7 mm transducer. (a) DAS (rectangular window) (b) DAS (Hamming window) (c) LCA (four predefined windows, $K = 0$) (d) LCA (12 predefined windows, $K = 0$) (e) LCA (12 predefined windows, $K = 15$) (f) MV ($L = 24$, $K = 15$). Tx focus was 50 mm.

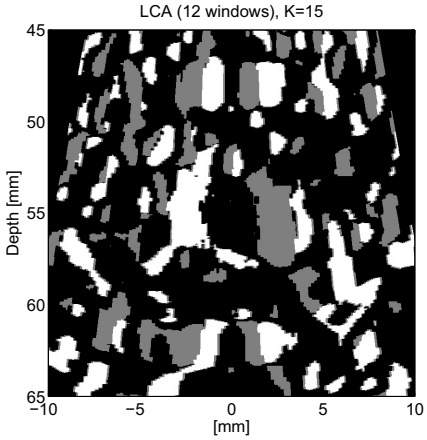
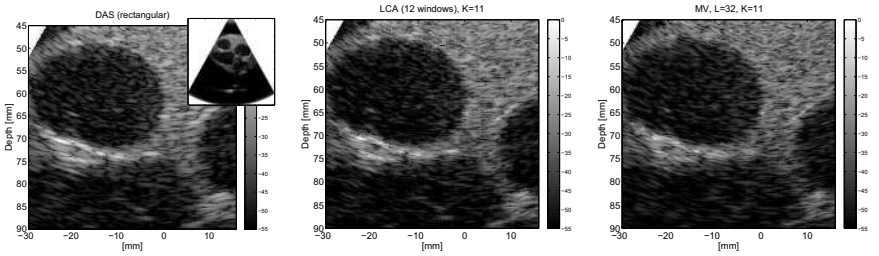


Fig. 9: Type of window that was chosen when forming the image in Fig. 8(e). Black indicates windows with symmetric beampatterns, gray and white indicates windows with asymmetric beampatterns.



(*a*) (*b*) (*c*)

Fig. 10: Images formed using experimental RF data and a 64 element, 3.33 MHz, 15.4 mm transducer. (a) DAS (rectangular window) (b) LCA (12 predefined windows, $K = 11$) (c) MV ($L = 32$, $K = 11$). Tx focus was 50 mm.

University of Michigan.* The data set was recorded using a 64 element, 3.33 MHz transducer, and sampled at 17.76 MHz. We upsampled the recorded signals by a factor eight before applying delays. Transmit focus was 50 mm. Fig. 10 shows the images obtained using DAS with rectangular window, LCA, and MV using subarray length $L = 32$. Both adaptive beamformers used $K = 11$, meaning that we have averaged over 23 samples when estimating the variance and covariance matrices. We see that both adaptive beamformers give better definition of the ventricular walls and similar speckle as the DAS image.

4 Discussion

We see from the images in Fig. 6 that both LCA and MV gives significantly better resolution than DAS. These are relatively simple images, but it gives an indication of the achievable resolution of bright scatterers. From Fig. 8 we see that the LCA and MV beamformers performs significantly better than DAS on more realistic images of a cyst and bright point targets in speckle as well. Compared to the DAS images, the cyst is closer to its actual size (6 mm diameter), and the point targets are better resolved. For LCA, the near targets are better resolved than the deeper targets. For the deeper targets the MV beamformer performs better. This shows some of differences between the full adaptive beamformer, and the approximation we get using LCA. The MV beamformer adapts to any scenario and can place zeros at the exact locations of off-axis interference. LCA has only a limited number of windows to choose from, and will not adapt equally well. We also see that due to the discrete solution space, some image artifacts are present between the point targets in the form of sidelobes in Fig. 8(e). These are not present in the MV image in Fig. 8(f). As the MV beamformer has more freedom, the beampatterns change

*Ultrasound RF data-set 'heart' from the Biomedical Ultrasound Laboratory, University of Michigan. Available at http://bul.eecs.umich.edu/, August 2008.

more smoothly when the image changes. From the results in Fig. 10 we see that the LCA beamformer improves the image quality on experimental data as well. The ventricular walls are better defined compared to DAS, and we see in particular that the brighter points are less smeared.

The proposed method may save a considerable number of computations compared to a full implementation of the MV beamformer. The number of operations required to find the MV weights is approximately $2L^3/3$ using Gaussian elimination [16], and hence is cubic with the subarray size. The complexity of the proposed method is linear with the number of elements, and the number of operations is approximately $2PM$. The method should also be easy to implement on existing imaging systems. For $K = 0$, we just need to apply $P - 1$ more windows to the data compared to DAS. We then compare the solutions for the different apodizations and select the one with the smallest absolute value. For $K > 0$, we must store $2K + 1$ values for each window and estimate the variances, for instance by (15). To achieve windows with asymmetric responses we have used complex weights, and we hence need complex data as input to the beamformer. To obtain this we computed the analytic signals by applying the Hilbert transform to the recorded channels. We could also have used real weights and delays to achieve the asymmetric responses. The weights would have to be scaled such that the response in the angle of the transmit beam is 0 dB. Such an implementation would fit well in the context of parallel beamforming. We could use several receive beams per transmission, and scale the off axis beams such that they have 0 dB gain in the direction of the transmit beam.

The performance of the proposed method depends very much on the window design. The best set of windows may be different for different scenarios, such that care must be taken to meet different criteria. Using many windows that are slightly different is desirable, as we get smooth transitions when the image changes, *e.g.* over a edge. Applying many windows will, however, come at a computational cost. We may increase the number of windows without a linear increase in cost by using classes of windows, where each class is represented by a "parent" window. We can first apply the parent windows, and classify the point based on the selected window. If for instance a window with asymmetric response is chosen, we classify the point as an edge and then try several windows that are shifted in the same direction, but have different slopes. In that way we do not have to test all windows, but can refine our search based on the primary selection.

From the image in Fig. 9 we see that the chosen window may assist in classifying the points in the image. Windows with asymmetric responses are chosen at the edges of the cyst and the edges of the point targets. However, we also see that in the speckle regions, asymmetric responses are chosen over large areas, so we cannot depend on the window selection only. However, it may be a tool to help defining boundaries between different tissue.

In this paper we have used both classical windows and windows that were

designed from our experience with the MV beamformer in the proposed method. Future research should further explore the window design, in particular directed at different applications within ultrasound imaging.

5 Conclusions

We have applied a data-dependent beamformer with low complexity and demonstrated improved image quality compared to delay-and-sum, both on simulated and experimental RF data. The method is robust, can handle coherent targets, and is easy to implement. The method requires only a fraction of the computations of a full adaptive beamformer like minimum variance, and may be a realistic candidate for real-time implementation.

References

[1] F. Vignon and M. Burcher, "Capon beamforming in medical ultrasound imaging with focused beams," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 55, no. 3, pp. 619–628, March 2008.

[2] J. A. Mann and W. F. Walker, "A constrained adaptive beamformer for medical ultrasound: Initial results," *Ultrasonics Symposium, 2002. Proceedings. 2002 IEEE*, vol. 2, pp. 1807–1810, October 2002.

[3] M. Sasso and C. Cohen-Bacrie, "Medical ultrasound imaging using the fully adaptive beamformer," *Acoustics, Speech and Signal Processing, 2005. Proceedings (ICASSP '05). IEEE International Conference on*, vol. 2, pp. 489–492, March 2005.

[4] J.-F. Synnevåg, A. Austeng, and S. Holm, "Minimum variance adaptive beamforming applied to medical ultrasound imaging," *Proc. IEEE Ultrasonics Symposium*, vol. 2, pp. 1199–1202, Sept. 2005.

[5] F. Viola and W. F. Walker, "Adaptive signal processing in medical ultrasound beamforming," *Proc. IEEE Ultrasonics Symposium*, vol. 4, pp. 1980–1983, Sept. 2005.

[6] Z. Wang, J. Li, and R. Wu, "Time-delay- and time-reversal-based robust Capon beamformers for ultrasound imaging," *IEEE Transactions on Medical Imaging*, vol. 24, pp. 1308–1322, Oct. 2005.

[7] J.-F. Synnevåg, A. Austeng, and S. Holm, "Adaptive beamforming applied to medical ultrasound imaging," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 54, no. 8, pp. 1606–1613, Aug. 2007.

[8] I. K. Holfort, F. Gran, and J. A. Jensen, "Minimum variance beamforming for high frame-rate ultrasound imaging," *Proc. IEEE Ultrasonics Symposium*, pp. 1541–1544, Oct. 2007.

[9] J.-F. Synnevåg, A. Austeng, and S. Holm, "Benefits of minimum variance beamforming in medical ultrasound imaging," September 2008, submitted for publication.

[10] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, pp. 1408–1418, August 1969.

[11] B. Widrow, K. Duvall, R. Gooch, and W. Newman, "Signal cancellation phenomena in adaptive antennas: Causes and cures," *Antennas and Propagation, IEEE Transactions on [legacy, pre - 1988]*, vol. 30, no. 3, pp. 469–478, May 1982.

[12] T.-J. Shan, M. Wax, and T. Kailath, "On spatial smoothing for direction-of-arrival estimation of coherent signals," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 4, pp. 806–811, August 1985.

[13] J.-F. Synnevåg, C. I. C. Nilsen, and S. Holm, "Speckle statistics in adaptive beamforming," *Proc. IEEE Ultrasonics Symposium*, pp. 1545–1548, Oct 2007.

[14] J. A. Jensen, "Field: A program for simulating ultrasound systems," *Medical & Biological Engineering & Computing*, vol. 34, pp. 351–353, 1996.

[15] R. F. Wagner, S. W. Smith, J. M. Sandrik, and H. Lopez, "Statistics of speckle in ultrasound b-scans," *IEEE Transactions on Sonics and Ultrasonics*, vol. 30, pp. 156–163, May 1983.

[16] G. H. Golub and C. F. V. Loan, *Matrix Computations*. Baltimore, Maryland: The John Hopkins University Press, 1996.

V

On Single Snapshot Minimum Variance Beamforming

Johan-Fredrik Synnevåg and Are F. C. Jensen

Abstract

We analyze the minimum variance beamformer when only single snapshots of the wavefields are used to optimize the array weights. We present an intuitive and unifying framework for showing how a direct implementation of such beamformers may lead to signal cancellation, and how subarray averaging can mitigate this effect. A key component is interpreting subarray averaging as a series of bandpass filters which decorrelate the different sources. The framework readily makes available what is required to avoid signal cancellation altogether, and allows us to analyze the compromises between mitigating signal cancellation and maximizing angular resolution.

1 Introduction

Adaptive beamforming methods often rely on accurate estimates of the signal statistics of the recorded data. For broadband, pulsed systems, few samples are available for the estimation. In [1] they suggest a method to optimize the array weights from a single temporal snapshot of the data. In this paper we analyze the minimum variance (MV) beamformer in such a single snapshot context. This technique has recently been applied to medical ultrasound imaging [2–5]. We show that a direct implementation based on single snapshots may lead to signal cancellation, and failure to estimate the amplitude of the propagating signals. In [3, 5] signal cancellation was avoided by subarray averaging. We show more formally how subarray averaging mitigates signal cancellation in the single snapshot case. Following from the derivation, we present an alternative optimization, in which signal cancellation is avoided.

2 Theory

2.1 Minimum variance beamformer

We consider a uniform linear array of M sensors, where each sensor records a signal, $y_m(t)$. The output of a beamformer operating on this array can be written as:

$$z(t) = \sum_{m=0}^{M-1} w_m^* y_m(t - \Delta_m) \tag{1}$$

$$= \mathbf{w}^H \mathbf{y}(t), \tag{2}$$

where w_m is a (complex) weight, Δ_m is the delay applied to channel m, $\mathbf{w} = [\, w_0 \;\; w_1 \;\; \cdots \;\; w_{M-1} \,]^T$, and

$$\mathbf{y}(t) = \begin{bmatrix} y_0(t - \Delta_0) \\ y_1(t - \Delta_1) \\ \vdots \\ y_{M-1}(t - \Delta_{M-1}) \end{bmatrix}. \tag{3}$$

By assuming that the signals in $\mathbf{y}(t)$ have been delayed to steer in a particular direction prior to further processing, the following derivations are valid for both narrowband and broadband signals. The weights may be used to control the mainlobe width and sidelobe level of the beamformer. The minimum variance (MV) beamformer finds the array weights from the recorded data by minimizing

the variance of $z(t)$ while keeping unit gain in the steering direction. The optimization problem becomes, assuming zero degree steering angle, [6]:

$$\begin{aligned} \min_{\mathbf{w}} \quad & E\left\{|z(t)|^2\right\} = \mathbf{w}^H \mathbf{R} \mathbf{w} \\ \text{subject to} \quad & \mathbf{w}^H \mathbf{1} = 1, \end{aligned} \tag{4}$$

where $\mathbf{R} = E\{\mathbf{y}\mathbf{y}^H\}$ is the spatial covariance matrix, and $\mathbf{1}$ is a vector of ones. The solution to (4) is:

$$\mathbf{w} = \frac{\mathbf{R}^{-1}\mathbf{1}}{\mathbf{1}^T \mathbf{R}^{-1} \mathbf{1}}. \tag{5}$$

2.2 Single Snapshot Minimum Variance Beamforming

We consider the problem of finding the MV weights from a single snapshot of the data. For notational simplicity we drop the time dependence of the output, but introduce a location dependence which will be useful for alternative formulations of the optimization. The nth sample of the output of the beamformer can be written as:

$$z_0 = \sum_{m=0}^{M-1} w_m^* x_m \tag{6}$$

$$= \mathbf{w}^H \mathbf{x}, \tag{7}$$

where $\mathbf{x} = [\, x_0 \;\; x_1 \cdots \;\; x_{M-1} \,]^T = \mathbf{y}(nT)$ (T is the sampling interval). We assume that $x_m = 0$ for $m < 0$ and $m > M-1$. By using Parseval's relation we can express (6) as:

$$z_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} W^*(k) X(k) dk, \tag{8}$$

and the square of the output, which can be considered as the single-sample variance, as:

$$|z_0|^2 = \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} W^*(k) X(k) dk \right|^2, \tag{9}$$

where $W(k)$ and $X(k)$ are the discrete-time Fourier transforms of w_m and x_m, and hence the wavenumber response of the weight function and the sampled wavefield, respectively. Note that the constraint in (4) forces $W^*(0) = 1$. The optimization problem based on a single temporal sample may be expressed as:

$$\begin{aligned} \min_{\mathbf{w}} \quad & \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} W^*(k) X(k) dk \right|^2 \\ \text{subject to} \quad & W^*(0) = 1. \end{aligned} \tag{10}$$

The solution to (10) is given by (5), where $\mathbf{R} = \mathbf{x}\mathbf{x}^H$. Note that $\mathbf{R}$ will be of rank one, and hence not invertible. By constraining the norm of $\mathbf{w}$, which

corresponds to adding a small regularization to $\mathbf{R}$, we can find a solution. From the optimization problem in (10) we can readily see that the one-sample variance estimate may lead to *signal cancellation*, a well known problem with the minimum variance beamformer when sources are coherent [7]. The constraint in (10) assures that the signal of interest is passed with gain equal to one. However, with more than one signal present, there may exist a solution in which the product $W^*(k_1)X(k_1)$ has the same value, but opposite sign, as $W^*(0)X(0)$. The constraint is then fulfilled, but the signal we steer towards is canceled by another signal propagating along the array with a different wavenumber. Such a solution would be smaller than if $W^*(k) = 0$ for all k where interfering signals are present, and hence the preferred solution given the cost function.

2.3 Subarray Averaging

Subarray averaging is a common method to avoid signal cancellation in the presence of correlated sources [8]. The aperture is divided into (overlapping) subarrays of length L and the output from each subarray is averaged:

$$z_{SA}(nT) = \frac{1}{M-L+1} \sum_{m=0}^{M-L} \mathbf{w}_L^H \mathbf{x}_m \tag{11}$$

$$= \frac{1}{M-L+1} \sum_{m=0}^{M-L} z_m \tag{12}$$

where $\mathbf{w}_L$ is a vector containing L weights, $\mathbf{x}_m = \begin{bmatrix} x_m & x_{m+1} & \cdots & x_{m+L-1} \end{bmatrix}$ is a vector containing the values of the mth subarray, and z_m is the output of the mth subarray beamformer.

The sum of the squares of the output of each subarray beamformer, which can be considered as an estimate of the subarray variances, becomes:

$$\sigma_{SA}^2 \propto \sum_{m=0}^{M-L} |z_m|^2 = \sum_{m=0}^{M-L} \left| \mathbf{w}_L^H \mathbf{x}_m \right|^2. \tag{13}$$

Minimizing (13) subject to $\mathbf{w}^H \mathbf{1} = 1$ yields the solution in (5), where

$$\mathbf{R} = \sum_{m=0}^{M-L} \mathbf{x}_m \mathbf{x}_m^H, \tag{14}$$

and $\mathbf{w} = \mathbf{w}_L$.

We consider $w_m^{(L)}$ to be an FIR filter of length L whose coefficients are given by the vector $\mathbf{w}_L$. We note that the sum in (11) is exactly the sum of a windowed convolution of $w_{-m}^{(L)}$ with x_m. Let

$$h_m = \begin{cases} 1 & \text{if } 0 \leq m \leq M-L \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

and

$$q_m = h_m \cdot (w_{-m} * x_m), \tag{16}$$

with Fourier transform:

$$Q(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H(k-\kappa) \left(W_L^*(\kappa) \cdot X(\kappa)\right) d\kappa, \tag{17}$$

where $H(k)$, $W_L(k)$ and $X(k)$ are the Fourier transforms of h_m, $w_{-m}^{(L)}$ and x_m, respectively. We can hence write

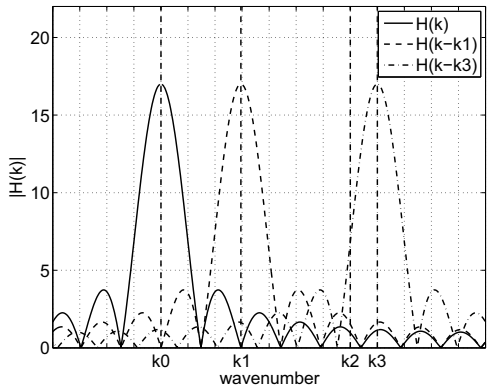
$$z_{SA} = \frac{1}{M-L+1} \sum_{m=-\infty}^{\infty} q_m, . \tag{18}$$

and, using Parseval's relation, we can write the estimate of the variance as:

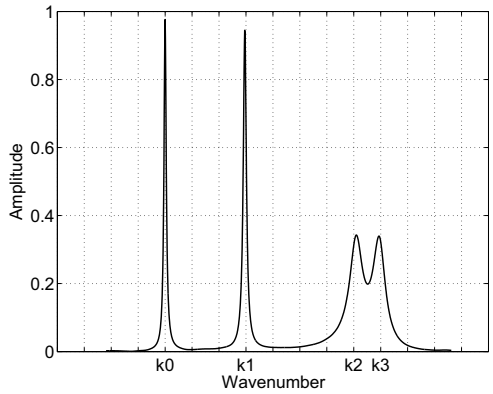
$$\begin{aligned} \sigma_{SA}^2 &\propto \sum_{m=-\infty}^{\infty} |q_m|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |Q(k)|^2 dk \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} H(k-\kappa) \left(W_L^*(\kappa) \cdot X(\kappa)\right) d\kappa \right|^2 dk. \end{aligned} \tag{19}$$

Note that for $L = M$, which corresponds to $h_m = \delta_m$ and $H(k) = 1$, we get the expression for $|z_0|^2$ in (9).

We can interpret (19) as an infinite sum of the powers in the signals $W_L^*(k) \cdot X_\kappa(k)$ (over all k), where $X_\kappa(k)$ is bandpass filtered version of $X(k)$ (filtered by $H(k-\kappa)$). $H(k)$ acts as a spatial bandpass filter which "decorrelates" signals at different wavenumbers. If two signals appear within the passband of the filter, they may cancel each other. If $H(k) = 1$, as in the expression for $|z_0|^2$, no decorrelation occurs. For subarray averaging, smaller L, i.e., shorter subarrays, better decorrelates closely spaced sources. As h_m gets longer, $H(k)$ becomes a narrower bandpass filter. Fig. 1(a) shows an example of $H(k-\kappa)$ for different values of κ using subarray averaging in which $M = 32$ and $L = 16$. The vertical dotted lines indicate possible sources at different wavenumbers. If sources are sufficiently separated such that only one source is within the passband of $H(k-\kappa)$ (as for the sources located at k_0 and k_1), signal cancellation is avoided. If two or more sources are spaced less apart than the width of the mainlobe of $H(k-\kappa)$, signal cancellation may occur (as for the sources at k_2 and k_3). Fig. 1(b) shows the steered response of an MV beamformer applied to the scenario in Fig. 1(a). All propagating signals have amplitude equal to one. We see that signal cancellation occurs when trying to estimate the sources located at k_2 and k_3.



(*a*)



(*b*)

Fig. 1: (a) Examples of $H(k-\kappa)$ for three different values of κ. (b) Steered response of a subarray averaged MV beamformer. The wavefield consists of four signals (corresponding to the vertical dashed lines in (a)).

2.4 Spatial Convolution

From the optimization in (10) we see that to avoid signal cancellation we would rather optimize:

$$\begin{aligned} \min_{\mathbf{w}} \quad & \frac{1}{2\pi}\int_{-\pi}^{\pi} |W^*(k)X(k)dk|^2 \\ \text{subject to} \quad & W^*(0) = 1, \end{aligned} \tag{20}$$

as any non-zero value of $W^*(k)X(k)$ then gives a positive contribution to the cost function. This is achieved by setting $L = M$ and $H(k) = \delta(k)$ in (19) (with corresponding $h_m = 1$). In the spatial domain the optimization becomes:

$$\begin{aligned} \min_{\mathbf{w}} \quad & \sum_{m=-M+1}^{M-1} \left|\mathbf{w}_L^H \mathbf{x}_m\right|^2 = \sum_{m=-M+1}^{M-1} |z_m|^2 \\ \text{subject to} \quad & \mathbf{w}^H \mathbf{1} = 1. \end{aligned} \tag{21}$$

The solution to this minimization can be found by constructing a zero-padded, $3M-2$ length measurement vector:

$$\bar{\mathbf{x}} = \begin{bmatrix} 0 & 0 & \cdots & x_0 & \cdots & x_{M-1} & \cdots & 0 & 0 \end{bmatrix}^T, \tag{22}$$

divide it into subarrays of length M, and form the spatial covariance matrix by subarray averaging. Note that although the calculation of the weights involves a sum over $|\mathbf{w}_L^H \mathbf{x}_m|^2$, the weights are applied to the original measurement vector.

Fig. 2 shows delay-and-sum (DAS) beamforming ($\mathbf{w} = 1/M \cdot \mathbf{1}$) compared to the subarray-averaged MV solution and the solution found from (21). Four signals, indicated by the vertical lines, were present in the wavefield, all with amplitude equal to one. We see that the conventional beamformer is unable to resolve closely spaced targets within the two groups. The MV beamformer in (21) resolves the two signals at $k = k_0$ and $k = k_1$, but is unable to resolve the signals at k_2 and k_3. However, this solution gives better estimates of the amplitude at $k = k_2$ and $k = k_3$ than both the conventional and subarray-averaged MV beamformer. The subarray-averaged MV beamformer provides better separation of the signals at $k = k_0$ and $k = k_1$, but gives inaccurate amplitude estimates of the signals at $k = k_2$ and $k = k_3$, due to signal cancellation.

3 Discussion and Conclusion

The single sample minimum variance beamformer given by the optimization in (10) has rather limited use in practice, unless the wavefield consists of a single signal source. With more than one source present, signal cancellation is likely to occur. Computing the signal statistics from a single temporal sample leads to an estimate of the spatial covariance in which all signals are correlated. The

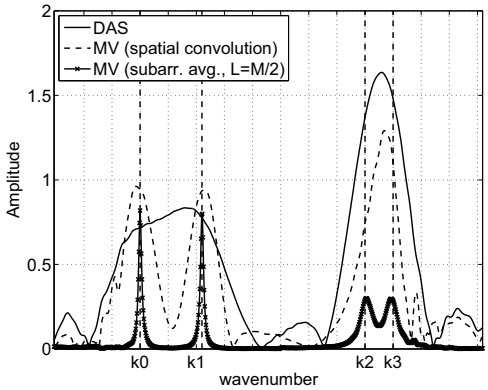


Fig. 2: Steered response of DAS (solid), MV in (21) (dashed) and subarray averaged MV (solid with x-marks) from wavefield consisting of four signals.

correlation is given by the instantaneous value of the source signals at sample n.

We have interpreted subarray averaging as a bandpass filter which decorrelates signals at different wavenumbers. Smaller L, i. e. smaller subarrays, gives a narrower bandpass filter and a better decorrelation of closely spaced sources. However, smaller L gives fewer weights, and hence fewer degrees of freedom to suppress interfering signals and noise. The choice of L will be a trade off between the performance of the beamformer and the robustness towards signal cancellation. Note that the solution given by (21), which corresponds to $H(k) = \delta(k)$, can be seen as a "perfect" decorrelation of the signals at different wavenumbers. Hence, this solution will not suffer from signal cancellation. However, this solution gives poorer resolution than using subarray averaging.

Note that in our simulation results we have used narrowband, far-field signal sources which only occupy a single wavenumber in the Fourier domain. We acknowledge that short, non-stationary sources will have a certain spatial bandwidth. However, the conclusions regarding the different optimization problems are valid for broadband sources as well.

References

[1] M. Ali and F. Schreib, "Adaptive single snapshot beamforming: a new concept for the rejection of nonstationary and coherent interferers," *Signal Processing, IEEE Transactions on*, vol. 40, no. 12, pp. 3055–3058, Dec 1992.

[2] J. A. Mann and W. F. Walker, "A constrained adaptive beamformer for medical ultrasound: Initial results," *Ultrasonics Symposium, 2002. Proceedings. 2002 IEEE*, vol. 2, pp. 1807–1810, October 2002.

[3] M. Sasso and C. Cohen-Bacrie, "Medical ultrasound imaging using the fully adaptive beamformer," *Acoustics, Speech and Signal Processing, 2005. Proceedings (ICASSP '05). IEEE International Conference on*, vol. 2, pp. 489–492, March 2005.

[4] F. Viola and W. F. Walker, "Adaptive signal processing in medical ultrasound beamforming," *Proc. IEEE Ultrasonics Symposium*, vol. 4, pp. 1980–1983, Sept. 2005.

[5] J.-F. Synnevåg, A. Austeng, and S. Holm, "Adaptive beamforming applied to medical ultrasound imaging," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 54, no. 8, pp. 1606–1613, Aug. 2007.

[6] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proc. IEEE*, vol. 57, pp. 1408–1418, August 1969.

[7] B. Widrow, K. Duvall, R. Gooch, and W. Newman, "Signal cancellation phenomena in adaptive antennas: Causes and cures," *Antennas and Propagation, IEEE Transactions on [legacy, pre - 1988]*, vol. 30, no. 3, pp. 469–478, May 1982.

[8] T.-J. Shan, M. Wax, and T. Kailath, "On spatial smoothing for direction-of-arrival estimation of coherent signals," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 4, pp. 806–811, August 1985.

VI

Blind Source Separation for Convolutive Mixtures Using Spatially Resampled Observations

Johan-Fredrik Synnevåg and Tobias Dahl

Abstract

We propose a new technique for separation of sources from convolutive mixtures based on independent component analysis (ICA). The method allows coherent processing of all frequencies, in contrast to the traditional treatment of individual frequency bands. The use of an array enables resampling of the signals in such a way that all frequency bands are effectively transformed onto the centre frequency. Subsequent separation is performed "all-bands-in-one". After resampling, a single matrix describes the mixture, allowing use of standard ICA algorithms for source separation.

The technique is applied to the cocktail-party problem to obtain an initial estimate of the separating parameters, which may further be processed using cross-talk removal or filtering. Experiments with two sources of speech and a four element microphone array show that the mixing matrix found by ICA is close to the theoretically predicted, and that 15 dB separation of the sources is achieved.

1 Introduction

A vast number of techniques have been developed for blind source separation (BSS) and extraction (BSE) over the last decades, and the field of BSS and BSE spurs hundreds of paper every year (see [1] for a recent survey).

Many BSS techniques are based on the transition from time domain convolutive mixtures to frequency domain instantaneous mixtures. A well known problem with this approach is the permutation and scaling inconsistencies which leads to a "re-mixing" of the sources when the frequency separated sources are transformed to the time-domain. As a consequence, numerous papers have been published that deal with this inconsistency, for a recent overview see [2, 3].

An interesting comment about the limitation of frequency based BSS was made by Araki et al. [4], showing that such techniques are essentially limited by the performance of an adaptive beamformer. This is down to an argument that works across each and every frequency: The echoes from the echoic environment will appear as directional signals coming into the array, essentially taking the place as “new signals”. For every frequency, it holds true that the limitations in the number of zeros that can be placed over the angular directions is limited, and that a tradeoff must be struck between the strength of the signal in the direction of the desire, and the various suppression levels that can be put on other directions. A direct consequence of Araki’s statements, is that frequency-band based BSS is heavily overparamaterized: Since the optimal result is obtained by zero forcing in the same positions along all frequencies, separation could ideally be derived for all frequency bands by “extrapolating” the zero-forcing settings estimated for a single band. Assuming all sources to occupy the same frequency band, or that no source separation can be based on frequency contents alone, it follows that performing ICA on each and every bin is a potential waste of computer power, since analysis of the bands essentially outputs the same directional separation-information. Given the limitations of frequency-based BSS, post-filtering and cross-talk removal is necessary to improve separation.

Other scientists pursue the time-domain approach [5] to avoid dealing with this inconsistency, but such methods easily become very complex. Yet others use combined approaches, a recent method computes inversion filters in time domain while using a cost function in frequency domain [6]. Time-frequency signatures [7] is a powerful tool that may be used for separation even in cases where there are more sources than mixes.

We propose a technique which has the potential for utilizing advantages of both time-domain and frequency-domain BSS. By using spatial resampling along the array direction in the temporal frequency domain, every frequency band of the original signal is "forced" onto the same spatial frequency. This enables an ICA-like representation of the BSS problem, avoiding the use of multiple frequency bins and the resulting permutation inconsistency. We present real-life experiments based on this approach and discuss its limitations.

It should be noted that we are no longer attempting to solve the "original" cocktail-party problem, but rather a modified problem with a lower number of estimated separation parameters than would be required for perfect separation in an echoic envirmonent. In [8], the authors propose a technique which incorporates knowledge of the microphone setup. However, the proposed techniques do no not involve the spatial resampling key step we propose here, and is hence quite different from the material in the present paper.

2 Method

2.1 Signal model

The use of independent component analysis for blind separation of independent sources requires observations of the form

$$\mathbf{x} = \mathbf{A}\mathbf{s}, \tag{1}$$

where $\mathbf{x} \in R^M$ is a vector of observations of M mixes , $\mathbf{s} \in R^N$ is a vector containing samples from N independent sources and $\mathbf{A} \in R^{M \times N}$ is the mixing matrix describing how the sources are observed.

The time-domain model for the cocktail party is more complex. To build intuition around the problem, we first consider a simplified, anechoic model, with no distortion effects at the microphones. Each observation, $x_j(t)$, for $j = 1, \ldots, M$ can then be modeled as

$$x_j(t) = \sum_{i=1}^{N} s_i(t) * \frac{1}{r_{ij}} \delta(t - \tau_{ij}), \tag{2}$$

where $x_j(t)$ is the output of the jth microphone, s_i is independent component i (the ith speaker), $\delta(\cdot)$ is the Dirac delta-function, $*$ is the convolution operator, r_{ij} is the distance from speaker i to microphone j, τ_{ij} is the sound propagation delay from speaker i to microphone j, and N is the number of speakers. Whereas the ICA model in (1) does not capture the linear convolution required to describe the time-delays between the sources and observation points, the model (2) encompasses this possibility. The full convolutive BSS model, containing possible echoes from multiple directions as well as filtering and attenuation effects in space, time and linear equipment distortion is

$$x_j(t) = \sum_{i=1}^{N} s_i(t) * b_{ij}(t), \tag{3}$$

where $\{b_{ij}(t)\}$ is a set of FIR filters describing the contributions of each of the sources indexed by j on each of the mixes indexed by i. The traditional way to adapt the echoic cocktail party problem on to the form (1) is by transforming the

observations to the frequency domain and treat each frequency independently, avoiding the convolution. Also, methods dealing purely with time-shifting and attenuation in the echo-free scenario [9] have been proposed. However, these do not take full advantage of the possibilities of array processing.

2.2 Frequency domain representation

Consider the model (2) in the frequency domain. The observations for one narrow frequency bin can be written as

$$x_j(\omega) = \sum_{i=1}^{N} a_{ij}(\theta) S_i(\omega), \tag{4}$$

where $S_i(\omega)$ is the amplitude of the ith source at frequency $\omega/2\pi$. Collecting terms from the M mixes into the vector $\mathbf{x}(\omega) = [x_1(\omega), x_2(\omega), ..., x_N(\omega)]^T$ and from all the mixes and the one source i into the vector $\mathbf{a}_i(\theta) = [a_{i1}(\theta), a_{i2}(\theta), \ldots, a_{iM}(\theta)]^T$, we can write

$$\mathbf{x}(\omega) = \sum_{i=1}^{N} \mathbf{a}_i(\theta) S_i(\omega) \tag{5}$$

where $\mathbf{a}_i(\theta)$ contains the time-delays between arrivals at the different sensors for the ith source. $\mathbf{a}_i(\theta)$ is known as the *steering vector* [10] in array processing. In matrix form, (5) can be written as

$$\mathbf{x}(\omega) = \mathbf{A}(\omega)\mathbf{S}(\omega), \tag{6}$$

where

$$\mathbf{A}(\omega) = [\mathbf{a}_1(\theta), \ldots, \mathbf{a}_M(\theta)], \tag{7}$$

and

$$\mathbf{S}(\omega) = [s_1(\omega), \ldots, s_N(\omega)]^T. \tag{8}$$

The problem (6) is now on the same form as (1), but each frequency has to be treated independently.

2.3 Transformation from convolutive to instantaneous mixtures

We propose to transform the convolutive model in (2) to the linear sum in (1) using a technique from array processing. The transformation requires an array of sensors with known geometry, as both the temporal and spatial frequency spectrum of the recorded wavefield must be captured. With this knowledge, spatial resampling is performed along the array direction for each temporal frequency band, effectively transforming the observations onto the ICA form.

2.3.1 Spatial frequency: The wavenumber

First, we explain the term *spatial frequency*, which is central in array signal processing. Imagine an acoustic wave measured in a single point in space and consider the *temporal* frequency of the signal. The question at hand is then how many times the signal oscillates within a given time span. Of course, a speech signal consists of many superpositioned waves oscillating with different periods, giving a whole spectrum of frequencies. Moving on to *spatial frequency*, the question is slightly different: If the wavefield is observed, not in a single point, but along a directive line segment in space, how many times does the wave oscillate within the line segment? This situation is illustrated in Fig. 1, which shows a narrowband, plane wave propagating in the xy-plane. The solid lines are the wave-fronts, meaning the lines of constant phase of the travelling wave, and the arrow indicates the direction of propagation. The number of periods of the wave fitting into a line segment of limited length in the xy-plane gives a measure of the spatial frequency in the direction of the line. Clearly, the spatial frequency will vary with the direction of the segment. In this example, if we look towards the wavefront, the spatial frequency will be higher than if the segment is placed along the x-axis. To be able to measure the frequency along a spatial dimension, we need access to data sampled along a line in space. This requires the use of an array.

The spatial frequency along the direction of propagation is called the *wavenumber* and is given by

$$k = \frac{\omega}{c}, \tag{9}$$

where c is the propagation velocity of the medium. The *wavenumber vector*, $\vec{k}$, contains the spatial frequencies along each spatial dimension of the wavefield, and satisfies the relation

$$|\vec{k}| = k. \tag{10}$$

By using a linear array of sensors, we sample one spatial dimension of the wavefield, and can estimate the wavenumber component in that direction. In Fig. 1 an array is located along the x-axis. The spatial frequency along this dimension is given by

$$k_x = k \sin(\theta), \tag{11}$$

where θ is the propagation angle, defined clockwise with respect to the y-axis. For the remainder of the paper we will refer to the wavenumber component along the array dimension simply as the wavenumber.

We denote the wavefield along the x-axis $z(x,t)$. By placing a microphone every d meters we can describe the sampled wavefield as

$$y_m(n) = z(md, nT), \tag{12}$$

where m is the sensor number, n is the temporal sample number, and T is the sampling interval. Similar to estimating the temporal frequency of the sampled

signal using the discrete Fourier transform as

$$Y_m(\omega) = \sum_{n=0}^{L-1} y_m(n) e^{-j\omega nT}, \tag{13}$$

where L is the number of temporal samples, we can estimate the wavenumbers along the array direction as

$$Y(k_x) = \sum_{m=0}^{M-1} y_m(n) e^{-jk_x md}. \tag{14}$$

The wavenumber-frequency response is a summation over both time and space,

$$Y(k_x, \omega) = \sum_{m=0}^{M-1} \sum_{n=0}^{L-1} y_m(n) e^{-j\omega nT} e^{-jk_x md}. \tag{15}$$

For narrowband waves, ω is fixed, and k_x will change as a function of propagation direction of the wave. If the temporal frequency of the signal is known, the propagation angle can be found using (11). Fig. 2(a) and (b) shows two examples of sinusoidal waves sampled in space and time by an eight channel linear array. Fig. 2(c) shows spatial samples at a selected point in time. For the wave in Fig. 2(a), which propagates in a direction orthogonal to the array, the amplitude is the same on all channels, because the signal arrives at the same time on all sensors. The wavenumber is then zero. For the signal propagating with non-zero angle, a sinusoidal pattern is evident over the array, leading to a non-zero wavenumber. Fig. 2(d) shows the estimated wavenumbers for the two examples. The corresponding angle of arrival is shown on the top axis.

Broadband waves are described by superpositioning narrowband waves with different temporal frequencies. Broadband waves correspond to lines in wavenumber-frequency space. Fig. 3 (a) shows the estimated wavenumber-frequency spectra of three broadband waves propagating across a uniform linear array. For non-zero angles of arrival, the wavenumber increases linearly with frequency, where the slope depends on the angle of arrival. All temporal frequencies of signals originating perpendicular to the array appear with zero wavenumber. Going back to the frequency domain model of the cocktail-party problem, the steering vector modeling the observed signal is simply $\mathbf{a}(\theta) = [1\ 1\ \ldots\ 1]^T$ for all frequencies, hence no time-delays are required to describe the observations. For signals arriving at an angle, the steering vector is frequency-dependent and given by $\mathbf{a}(\theta) = [1\ e^{jk_x d}\ \ldots\ e^{jk_x(M-1)d}]^T$. If the wavenumber was constant regardless of frequency for any incidence angle, the steering vector would be identical for all frequencies, and a single mixing matrix would describe the observations of several sources. The problem would then be of the form (1). Spatial resampling [11] is a means to achieve that.

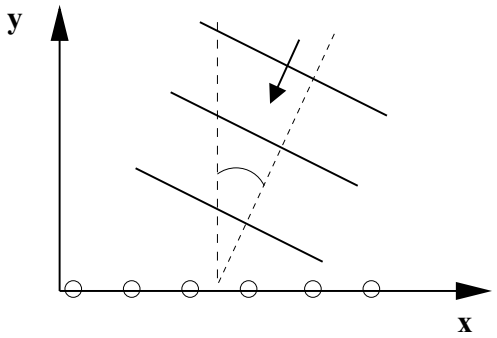


Fig. 1: Illustration of a plane wave propagating in the xy-plane.

2.3.2 Spatial resampling

The goal of spatial resampling is to force all temporal frequency components of a wave to appear with the same wavenumber. We choose a center frequency, f_c, where the corresponding wavenumber will appear for all frequencies. For $f > f_c$ the spatial "sampling rate" is increased by a factor f/f_c by interpolating between samples. Each sample corresponds to a point in space, given by the sensor location. The original number of samples are kept, effectively reducing the size of the aperture as we are throwing away samples at the edges. For frequencies $f < f_c$, we decrease the sampling rate. Since we must keep the original number of samples, the latter case introduces zeros at the edges as information is missing outside the original aperture. Fig. 3(b) shows the corresponding frequency-wavenumber plot of Fig. 3(a) after spatial resampling. We see that all frequency components appear with the same wavenumber. After the resampling step, the observations are transformed back to the time domain. Each temporal frequency component of the observations are now described by identical steering vectors, and we have a set of observations on ICA form (1).

3 Results

We have evaluated the performance of the method experimentally with a linear array of four microphones in an anechoic chamber. The experimental setup is shown in Fig. 4. The microphones were separated by 30 cm. Two speakers were placed in front of the array, each transmitting a different speech signal. The distance between the speakers was 1 m, giving propagation angles of 0° and approximately 17° for the sources, with reference to the center of the array. The recorded signals were bandpass filtered, passing frequencies between 300 Hz and 3 kHz. We resampled the observations spatially, using 500 Hz as resampling frequency. The sources were then separated with the JADE algorithm [12].

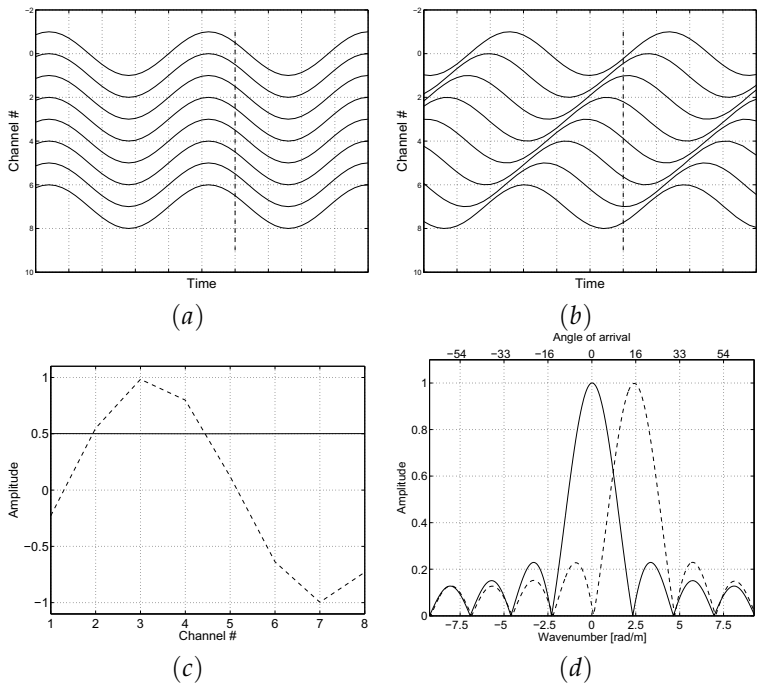


Fig. 2: Simulated time-series from an 8-channel linear array recording a plane wave propagating with (a) 0° and (b) 17° angle. (c) Spatial samples at the time instance given by the dotted lines. (d) The estimated wavenumber of the propagating waves. The top axis shows the corresponding propagation angles.

We evaluated the performance of the separation using the signal-to-interference ratio (SIR) for each of the estimated sources, given by the power in the desired signal to the power in the interfering signals. The SIR for source 1 was estimated as

$$\hat{SIR}(\hat{s}_1) = \left(\frac{E\left(\hat{s}_1(t) s_1(t)\right)}{E\left(\hat{s}_1(t) s_2(t)\right)} \right)^2 \tag{16}$$

where $\hat{s}_1$ is estimate of source 1, $s_1(t)$ is the true source 1 and $s_2(t)$ is the true "interfering" source. Both $s_1(t)$ and $s_2(t)$ were normalized to unit variance before evaluation. This measure assumes that the estimate of source 1 contains scaled versions of $s_1(t)$ and $s_2(t)$, which is a simplification as no filtering effects are taken into account. The equivalent measure was calculated for source 2.

To evaluate the performance of the method, we first looked at how well the estimated mixing matrix compared to the theoretically predicted. Note that the theoretical mixing matrix relates to the resampled observations, hence the

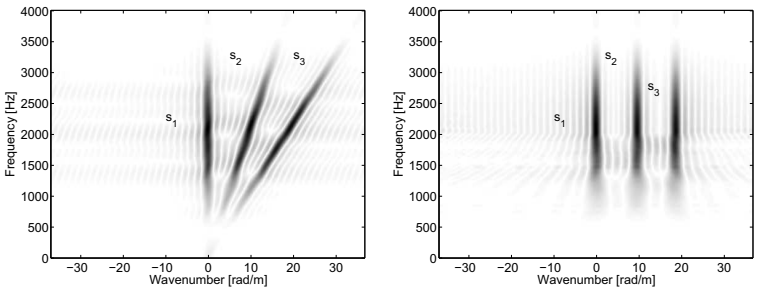


Fig. 3: Illustration of the effect of spatial resampling. The s_i represent three different broadband waves propagating across a linear array in different directions. The figures show frequency-wavenumber plots (a) before and (b) after spatial resampling.

columns of this matrix is the steering vectors in (5) with $\theta = 0^\circ$ and $\theta = 17^\circ$ for frequency $f_c = 500$ Hz. It is most instructive to look at the wavenumber response of the steering vectors, as its peak corresponds to the propagation angle. Fig. 5 shows the wavenumber response of the first column of the theoretical mixing matrix and the corresponding vector found by ICA. Note that for this, and the remaining plots, wavenumber has been translated to propagation angle using (11). The dashed vertical lines shows the true propagation angles. We see that the peak of the steering vector found by ICA corresponds to the true propagation angle for source 2.

Rather than looking at the mixing matrix, it is more instructive to study the unmixing matrix, as its rows correspond to spatial unmixing filters for each source. In the directions of interfering signals the response should be zero for perfect separation. Fig. 6 shows the wavenumber response of the second row of the unmixing matrix found by ICA, together with the response of the theoretical unmixing filter for the second source. We see that there is a positive gain in the direction of source 2, and that a zero is placed in the direction of source 1. The resulting average signal-to-interference ratio was approximately 15 dB for the two sources.

4 Discussion

The success of source separation with ICA on spatially resampled observations demands that no significant correlations is introduced between the original independent sources during resampling. The temporal frequency contents of the original signals are affected by the transformation, depending on resampling

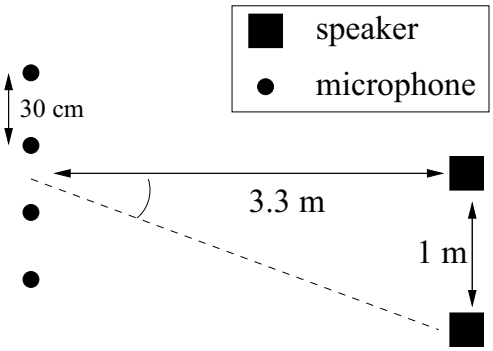


Fig. 4: Experimental setup.

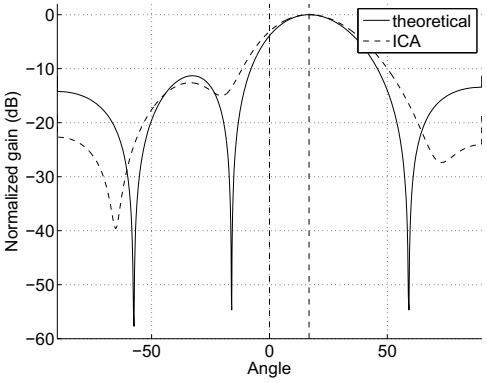


Fig. 5: Wavenumber responses of the theoretical steering vector (solid) and the steering vector found by ICA (dashed) for source 2 ($f = 500$ Hz).

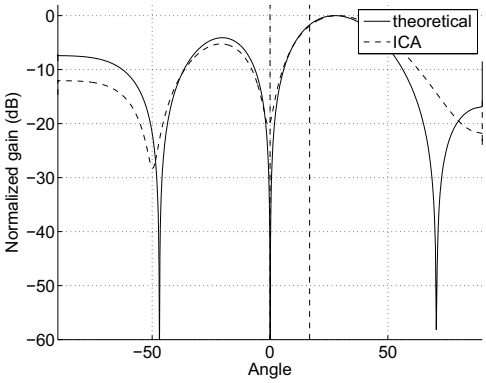


Fig. 6: Wavenumber responses of the unmixing filters for source 2, theoretical (solid) and ICA (dashed). The dashed vertical lines indicate the propagation angles of each source.

frequency and array geometry. Closely spaced sources may be correlated after the transformation and thereby indistinguishable with ICA.

The present method finds unmixing filters in the spatial domain, not exploiting temporal correlations in the observations. As a consequence, echos have to be treated as new "independent" sources, and we are limited by the number of zeros that can be forced in the wavenumber response. In the echoic scenario we need as many sensors as there are sources and echos for perfect separation.

5 Conclusion

We have presented a new method for blind source separation of convolutive mixtures based on independent component analysis. The method treats all frequency bands simultaneously and thereby avoids the permutation problem of frequency domain methods. We have demonstrated the method on experimental data from the anechoic cocktail-party problem, and shown good separation performance.

References

[1] S.T.Rickard P.D. O'Grady, B.A. Pearlmutter. Survey of sparse and non-sparse methods in source separation. *International Journal of Imaging Systems and Technology*, Special Issue on Blind Source Separation and Deconvolution in Imaging and Image Processing:18–33, July 2005.

[2] Nikolaos Mitianoudis and Mike E. Davies. Audio source separation: Solutions and problems. *International Journal of Adaptive Control and Signal Processing*, 18:299–314, 2004.

[3] Hiroshi Sawada, Ryo Mukai, Shoko Araki, and Shoji Makino. A robust and precise method for solving the permutation problem of frequency-domain blind source separation. *IEEE Transactions on Speech and Audio Processing*, 12:530–538, September 2004.

[4] Shoko Araki, Ryo Mukai, Shoji Makino, Tsuyoki Nishikawa, and Hiroshi Saruwatari. The fundamental limitation of frequency domain blind source separation for convolutive mixtures of speech. *IEEE Transactions on Speech and Audio Processing*, 11(2):109–116, March 2003.

[5] S.C.Douglas and A.Cichocki. Convergence analysis of local algorithms for blind decorrelation. *NIPS96 Workshop, Blind Signal Processing and Their Applications*, 1996.

[6] F. Yin T.Mei, J.Xi and Z.Yang. A half-frequency domain approach for convolutive source separation based on the kullback-leibler divergence. *Eight International Symposium on Signal Processing and its Applications (ISSPA-2005)*, 1:25–28, August 2005.

[7] B.Barkat and K.Abed-Meraim. Algorithms for blind components separation and extraction from the time-frequency distribution of their mixture. *EURASIP Journal on Applied Signal Processing*, 13:2025–2033, 2004.

[8] L.C.Parra and C.V. Alvino. Geometric source separation: Merging convolutive source separation with geometric beamforming. *IEEE Transactions on Speech and Audio Processing*, 10(6):352–362, September 2002.

[9] Justinian Rosca, NingPing Fan, and Radu Balan. Real-time audio source separation by delay and attenuation compensation in the time domain. *Proc. of the 3rd ICA and BSS Conference, San Diego, CA*, December 2001.

[10] Hamid Krim and Mats Viberg. Two decades of array signal processing research: the parametric approach. *IEEE Signal Processing Magazine*, 13(4):67–94, July 1996.

[11] Jeffrey Krolik and David Swingler. The performance of minimax spatial resampling filters for focusing wide-band arrays. *IEEE Transactions on Signal Processing*, 39(8):1899–1903, August 1991.

[12] Jean-François Cardoso and Antoine Souloumiac. Jacobi angles for simultaneous diagonalization. *SIAM J. Mat. Anal. Appl.*, 17(1):161–164, January 1996.