A comparison between temporal and subband minimum variance adaptive beamforming

Diamantis, Konstantinos; Voxen, Iben Holfort; Greenaway, Alan H.; Anderson, Tom; Jensen, Jørgen Arendt; Sboros, Vassilis

Published in:
Proceedings of the SPIE - Progress in Biomedical Optics and Imaging

Link to article, DOI:
10.1117/12.2043602

Publication date:
2014

Document Version
Early version, also known as pre-print

Citation (APA):
A comparison between temporal and subband minimum variance adaptive beamforming

Konstantinos Diamantis¹, Iben H. Voxen²,³, Alan H. Greenaway¹, Tom Anderson⁴, Jørgen A. Jensen² and Vassilis Sboros¹

¹School of Engineering & Physical Sciences, Heriot-Watt University, Edinburgh EH14 4AS, UK.
²Center for Fast Ultrasound Imaging, Dept. of Electrical Engineering Technical University of Denmark, 2800 Lyngby, Denmark.
³Greenwood Engineering, 2605 Brøndby, Denmark.
⁴Medical Physics, Centre for Cardiovascular Science, The University of Edinburgh, Edinburgh EH16 4TJ, UK.

ABSTRACT

This paper compares the performance between temporal and subband Minimum Variance (MV) beamformers for medical ultrasound imaging. Both adaptive methods provide an optimized set of apodization weights but are implemented in the time and frequency domains respectively. Their performance is evaluated with simulated synthetic aperture data obtained from Field II and is quantified by the Full-Width-Half-Maximum (FWHM), the Peak-Side-Lobe level (PSL) and the contrast level. From a point phantom, a full sequence of 128 emissions with one transducer element transmitting and all 128 elements receiving each time, provides a FWHM of 0.03 mm (0.14λ) for both implementations at a depth of 40 mm. This value is more than 20 times lower than the one achieved by conventional beamforming. The corresponding values of PSL are -58 dB and -63 dB for time and frequency domain MV beamformers, while a value no lower than -50 dB can be obtained from either Boxcar or Hanning weights. Interestingly, a single emission with central element #64 as the transmitting aperture provides results comparable to the full sequence. The values of FWHM are 0.04 mm and 0.03 mm and those of PSL are -42 dB and -46 dB for temporal and subband approaches. From a cyst phantom and for 128 emissions, the contrast level is calculated at -54 dB and -63 dB respectively at the same depth, with the initial shape of the cyst being preserved in contrast to conventional beamforming. The difference between the two adaptive beamformers is less significant in the case of a single emission, with the contrast level being estimated at -42 dB for the time domain and -43 dB for the frequency domain implementation. For the estimation of a single MV weight of a low resolution image formed by a single emission, 0.44 * 10⁹ calculations per second are required for the temporal approach. The same numbers for the subband approach are 0.62 * 10⁹ for the point and 1.33 * 10⁹ for the cyst phantom. The comparison demonstrates similar resolution but slightly lower side-lobes and higher contrast for the subband approach at the expense of increased computation time.

Keywords: adaptive beamforming, minimum variance, covariance matrix, medical ultrasound

1. INTRODUCTION

Adaptive beamforming results in a higher image quality both in terms of resolution and contrast, when compared to the conventional Delay-and-Sum (DAS) beamformer that is widely used in ultrasound imaging. Adaptive processes have been initially used in various fields of array signal processing,¹⁻³ for instance sonar and radar. One of the most common adaptive methods, which was introduced by Capon⁴ in 1969 in an attempt to improve localization of earthquakes using seismic arrays, is the minimum variance or Capon beamformer.⁵ It was not until the early 2000 that adaptive beamforming has been applied to both experimental and simulated ultrasound data, such as single point targets and cyst phantoms. With a few exceptions like the linearly constrained adaptive beamformer⁶ suggested by Mann and Walker⁷ in 2002 or the adaptive beamformers introduced by Viola and Walker⁸ in 2005, the majority of the proposed adaptive methods were based on the Capon beamformer⁹⁻¹³

MV based adaptive beamforming can be implemented in the frequency domain,¹⁴ where the broad-band, ultrasound sensor signals are divided into frequency bands. This is to ensure that the original narrow-band condition of the adaptive beamformer is met. Each subband is processed separately and for each one, an optimized set of complex apodization weights is provided based on the sample covariance matrix estimated from the data. The time-domain implementation¹⁵ calculates the sample covariance matrix directly from the data.
The objective of the MV beamformer is to minimize the variance of $B(t)$ where $E$ is the expectation value and without the previous requirement of dividing them into separate frequency bands. Therefore, a single set of data-dependent apodization weights is provided instead of one for each segment. Among others, Synnevåg et al.\textsuperscript{16} that implemented the MV beamformer in the time-domain, have shown that the method can be made robust by controlling a number of parameters and that subarray averaging greatly affects the beamformer performance. In 2008\textsuperscript{17} they proposed an easy to implement, low-complexity approach, which might suit the real-time requirements. Holfort et al.\textsuperscript{18, 19} that implemented the frequency domain MV beamformer, have applied it to different imaging sequences such as synthetic aperture and plane wave imaging. They have proposed the application of minimum variance apodization weights both in the transmitting and receiving apertures\textsuperscript{20} and have investigated the influence of sound speed errors on the adaptive beamformer.\textsuperscript{21}

In the time-domain approach the MV response is derived from a single beamformer used for all data. In the frequency-domain, a number of narrow-band beamformers equal to the number of frequency subbands is needed for the processing, and then the outputs of all beamformers are summed to provide the final response. As a consequence, implementation in the frequency domain is a more complicated problem, and the total number of calculations is significantly higher. The purpose of this paper is to compare the temporal and subband MV beamformers, and examine the benefit of the subband approach. The comparison is accomplished by evaluating the performance of the two different approaches using Field II\textsuperscript{22, 23} simulated data.

2. METHODS

2.1 Temporal minimum variance beamformer

The DAS beamformer is the standard method to process the signals received by an array of sensors in ultrasound imaging. In DAS beamforming, the sensor signals are time-delayed, weighted, and subsequently summed to form a maximized beamformer output. The process is not different in the case of the MV adaptive beamformer apart from the calculation of the apodization weights. Thus, the output $B(t)$ of the MV beamformer for a transducer array with $M$ active transmit and receive elements is given by:

$$B(t) = \sum_{m=0}^{M-1} w_m(t) x_m(t - \tau_m) = w(t)^H X(t) ,$$  \hspace{1cm} (1)

where $t$ is the time index, $w(t) = [w_0(t), w_1(t), ..., w_{M-1}(t)]^H$ is the vector of apodization weights, $X(t) = [x_0(t - \tau_0), x_1(t - \tau_1), ..., x_{M-1}(t - \tau_{M-1})]^H$ is the array of the sensor signals, and $\tau_m$ is the time delay, depending on the distance between the position of the $m$th element and the focus point. In MV beamforming, the power $P(t)$ of the beamformer output is given by:

$$P(t) = E[|B(t)|^2] = E[|w(t)^H X(t)|^2] = E[w(t)^H X(t) X(t)^H w(t)] = w(t)^H R(t) w(t) ,$$  \hspace{1cm} (2)

where $E[.]$ denotes the expectation value and $R(t)$ is the covariance matrix given by:

$$R(t) = E[X(t) X(t)^H] .$$  \hspace{1cm} (3)

The objective of the MV beamformer is to minimize the variance of $B(t)$, thus power, while maintaining the response of the focal point. For this reason, the apodization weights are not pre-defined and independent from the ultrasound sensor signals as in DAS beamforming, but are being constantly updated. This objective, by omitting the dependency on time, can be expressed as:

$$\min w^H R w , \; \text{subject to} \; w^H e = 1 ,$$  \hspace{1cm} (4)

Here $e$ is the steering vector that is only a vector of ones, since the time delays have been already applied to the signals. An analytical solution to this constrained optimization problem can be given by the use of Lagrangian multiplier theory.\textsuperscript{24} Under the assumption that $R^{-1}$ exists, the MV weights can be extracted by:

$$w = \frac{R^{-1} e}{e^H R^{-1} e} .$$  \hspace{1cm} (5)
2.2 Subband minimum variance beamformer

In the frequency-domain implementation of the MV beamformer, the time delayed sensor signals are divided into frequency subbands with the use of the short-time Fourier transform (STFT) and are processed individually. This increases the computational burden in relation to the temporal approach. For each focus point, \( \vec{r}_p \), the Discrete Fourier Transform (DFT) is applied on a segment, hence STFT, of the aligned sensor signals. The \( n \)th segmented, sensor signal \( y_m(t) \) is given for \( t \in [-t_d/2, t_d/2] \) where \( t_d \) is the time duration of the segment size, as the response from the focus point will be centered around \( t = 0 \). The inverse Fourier Transform is then employed for the summed individual subband responses to extract the MV response. The beamformer output for a specific focus point, and for each frequency subband, \( \omega \), is given by:

\[
B(\omega) = \sum_{m=0}^{M-1} w_m(\omega) Y_m(\omega) = w(\omega)^H Y(\omega),
\]

where \( w(\omega) = [w_0(\omega), w_1(\omega), ..., w_{M-1}(\omega)]^T \) is the vector of the complex apodization weights, and \( Y(\omega) = [Y_0(\omega), Y_1(\omega), ..., Y_{M-1}(\omega)]^T \) is the vector of the Fourier Transform of the segmented sensor signals. Hence, the power of the beamformer output is given by:

\[
P = E\{|B(\omega)|^2\} = E\{|w(\omega)^H Y(\omega)|^2\} = E\{|w(\omega)^H Y(\omega)Y(\omega)^H w(\omega)|\} = w(\omega)^H R(\omega) w(\omega),
\]

where \( R(\omega) \) is the covariance matrix given by:

\[
R(\omega) = E\{Y(\omega)Y(\omega)^H\}.
\]

By omitting again the dependency on \( \omega \) this time, and by assuming that \( R^{-1} \) exists, the goal of the subband MV beamformer remains unvaried and the optimized apodization weights are given by (5). An important advantage of the subband MV beamformer is that due to the subband division, it is possible to weight both each subband and each point differently.

2.3 Subarray averaging

Simply by applying \( w \) to (1) or (6) the output of the MV beamformer can be calculated for both cases. However, due to the lack of sufficient statistical information in near-field applications, the covariance matrix is unknown and has to be replaced by the sample covariance matrix \( \hat{R} \), which is estimated from the data. In this paper, a spatial smoothing technique is employed to obtain a good estimate. The array is divided into several overlapping sub-arrays, from the front to the end, and then the covariance matrix is averaged across the array. The estimate of the covariance matrix can be expressed as:

\[
\hat{R} = \frac{1}{M - L + 1} \sum_{l=0}^{M-L} G_l G_l^H,
\]

where \( L \) is the sub-array length, and \( G_l \) is the set of signals from the \( l \)th sub-array that may take the form of \( G_l(t) = [x_l(t), x_{l+1}(t), ..., x_{l+L-1}(t)]^T \) for the temporal, or \( G_l(\omega) = [Y_l(\omega), Y_{l+1}(\omega), ..., Y_{l+L-1}(\omega)]^T \) for the subband MV beamformer. The size of the subarray is strongly related to the accuracy of the covariance matrix estimate, with smaller ones providing more accurate estimates, but resulting in reduced resolution at the same time. After the calculation of the optimized apodization weights, \( \hat{w} \), with the aid of the covariance matrix estimate, the beamformer output can be given for both approaches by:

\[
B(\vec{r}_p) = \hat{w}^H \frac{1}{M - L + 1} \sum_{l=0}^{M-L} G_l.
\]
3. DATA ACQUISITION

Simulated synthetic aperture data are used for the comparison of the temporal and subband MV beamformers. Synthetic aperture (SA) is the imaging sequence in which the transmitting aperture consists of only one transducer element, while all elements are used as the receiving aperture. The single active element is emitting a spherical wave in every direction covering the whole image region, and is moved across the array. The following transducer element then becomes active, with a full SA sequence consisting of as many emissions as the total number of transducer elements. From each emission a low resolution image can be formed. Those single emission images can be then combined so as to provide a final image of higher resolution and contrast.

Images displayed in this paper correspond to combined responses from a full synthetic aperture sequence, consisting of 128 emissions, for a point and a cyst phantom obtained both from Field II. However, results from one single emission with the central element, number #64, as the transmitting aperture could be considered comparable to those obtained from 128 averaged emissions. Therefore they are also being presented here for both temporal and subband MV beamformers as they come at a computational time many times lower than for the total number of emissions. The point phantom consists of 12 point targets in pairs of two, separated by 4 mm laterally and located at axial depths of 30 to 80 mm. The purpose is to evaluate the point spread function (PSF) at different depths and compare the beamformers in respect to resolution and side lobes. The second phantom contains a circular cyst with a radius of 5 mm and centered at $(x, z) = (0, 40)$ mm, in a speckle pattern. This phantom is used for the evaluation of contrast of the different beamformers. The cyst phantom has 10 randomly placed scatterers within a resolution cell of $\lambda \times \lambda \times \lambda$ to ensure fully developed speckle. All the Field II simulation parameters can be found in Table 1.

<table>
<thead>
<tr>
<th>Transducer</th>
<th>Linear array</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transducer element pitch</td>
<td>110 µm</td>
</tr>
<tr>
<td>Transducer element kerf</td>
<td>35 µm</td>
</tr>
<tr>
<td>Transducer element height</td>
<td>6 mm</td>
</tr>
<tr>
<td>Center frequency, $f_0$</td>
<td>7 MHz</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>60% fractional</td>
</tr>
<tr>
<td>Speed of sound, $c$</td>
<td>1540 m/s</td>
</tr>
<tr>
<td>Wavelength, $\lambda = c/f_0$</td>
<td>220 µm</td>
</tr>
<tr>
<td>Excitation pulse</td>
<td>Two-cycle sinusoid at $f_0$</td>
</tr>
</tbody>
</table>

Table 1: Parameters of Field II simulations

4. SIMULATION RESULTS

The beamformed responses of the point targets are shown in Fig. 1 with a dynamic range of 60 dB, and a subarray length of $L = 32$ for the MV beamformers. Responses from the conventional beamformer using boxcar and Hanning apodization weights are also displayed as a standard of comparison. Fig. 2, in which the $x$-axis represents the lateral distance and the $y$-axis the power in dB, shows the lateral variations at 40 mm depth for the different methods and for both cases of a single and 128 emissions. Lateral variations are also shown at a depth of 80 mm in Fig. 3. Lateral Full-Width-Half-Maximum (FWHM) and Peak-Side-Lobe level (PSL) are the two criteria adopted for the performance evaluation and their values are shown in Table 2. From images and lateral variations, it can be noticed that the performance of all beamformers is affected by the location of the point targets. This variation of FWHM in respect to depth can be found in Fig. 4 for the MV beamformers and in Fig. 5 for the conventional ones. FWHM values between the two different MV approaches differ only in the fourth decimal digit at a depth of 40 mm and this difference becomes slightly more significant as the depth increases. The behavior of the PSL values is similar in favor of the frequency domain implementation.
Figure 1: Beamformed responses of 12 point targets (a) DAS Boxcar, (b) DAS Hanning, (c) MV Temporal, and (d) MV Subband. Images are shown with a dynamic range of 60 dB.

Figure 2: Lateral variations at 40 mm of the beamformed responses for (a) single emission (element #64), and (b) 128 emissions.
Figure 3: Lateral variations at 80 mm of the beamformed responses for (a) single emission (element #64), and (b) 128 emissions.

Table 2: Peak-side-lobe level (PSL) and lateral Full Width at Half Maximum (FWHM) for the beamformed responses at $z = 40$ mm, where $\lambda = c/f_0 = 220 \mu m$.

<table>
<thead>
<tr>
<th></th>
<th>PSL</th>
<th>FWHM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Single emission (element #64)</td>
<td></td>
</tr>
<tr>
<td>DAS Boxcar</td>
<td>$-16$ dB</td>
<td>0.78 mm</td>
</tr>
<tr>
<td>DAS Hanning</td>
<td>$-40$ dB</td>
<td>1.26 mm</td>
</tr>
<tr>
<td>MV Temporal</td>
<td>$-42$ dB</td>
<td>0.04 mm</td>
</tr>
<tr>
<td>MV Subband</td>
<td>$-46$ dB</td>
<td>0.03 mm</td>
</tr>
<tr>
<td></td>
<td>Full sequence (averaged over 128 emissions)</td>
<td></td>
</tr>
<tr>
<td>DAS Boxcar</td>
<td>$-31$ dB</td>
<td>0.65 mm</td>
</tr>
<tr>
<td>DAS Hanning</td>
<td>$-50$ dB</td>
<td>0.82 mm</td>
</tr>
<tr>
<td>MV Temporal</td>
<td>$-58$ dB</td>
<td>0.03 mm</td>
</tr>
<tr>
<td>MV Subband</td>
<td>$-63$ dB</td>
<td>0.03 mm</td>
</tr>
</tbody>
</table>

Figure 4: FWHM variation in respect to depth for (a) single emission (element #64), and (b) 128 emissions.
Figure 5: FWHM variation in respect to depth for (a) single emission (element #64), and (b) 128 emissions.

Figure 6: Circular cyst with radius of 5 mm and center at \((x, z) = (0, 40)\) mm. (a) DAS Boxcar, (b) DAS Hanning, (c) MV Temporal, and (d) MV Subband. Images are shown with a dynamic range of 60 dB.

Figure 7: Lateral variations at 40 mm of the beamformed responses for the cyst phantom, for (a) single emission (element #64), and (b) 128 emissions.
Images of the cyst phantom and the lateral variations, for the different methods can be found in Fig. 6 and Fig. 7. The contrast levels for the cyst responses are -63 and -54 dB for subband and temporal MV Beamformers. The value from boxcar weights is -53 dB, and -60 dB in the case of Hanning apodization. Although the contrast level between the 4 beamformers does not differ greatly in value, it is obvious from the images that the shape of the cyst has been distorted in the case of conventional beamforming. The corresponding values for the case of a single emission are -22, -44, -42, and -43 dB for DAS Boxcar, DAS Hanning, MV Temporal and MV Subband. For the subband approach, up to 41 different subbands are needed, depending on the values of the segmented sensor signals. Those subbands correspond to up to 41 narrow-band beamformers and thus, to up to 41 sets of apodization weights, subsequently combined to form the final MV response.

5. NUMBER OF CALCULATIONS

A comparison between the two different approaches shows slight improvement in FWHM and PSL from the point targets and in contrast from the cyst phantom, for the subband MV beamformer. Apart from the overall performance, the number of total calculations required for the estimation of MV weights, is an equally important characteristic of the beamformers. In general the high computation time is the major disadvantage of adaptive beamforming, and the purpose of the current Section is to analyze this aspect of the MV beamformers as well. Both beamformers have been implemented in Matlab and a description of calculations taking place follows below.

For the estimation of apodization weights in the case of the temporal MV beamformer, a number of matrix multiplications are needed. They are determined by (5) in Section 2.1, and (9) in Section 2.3. The computations for the estimation of a single MV weight, for a single emission will be considered here. The exact same steps are followed for every MV weight, in order to create low resolution image from a single emission. Then the same process is repeated for the total number of emissions that will result into the high resolution images displayed in the previous Section. The first step is the estimation of the sample covariance matrix, $\hat{R}$, that substitutes the covariance matrix, $R(t)$, in (3). The matrix $G_t$ has as many rows as the selected subsize, $L$, and as many columns as $l$, which is equal to $M - L + 1$, and the inverse dimensions stand for matrix $G_t^H$. Therefore $(M - L + 1) \times L \times L$ multiplications and $(M - L) \times L \times L$ additions are required up to this stage. In (5) there are 3 more matrix multiplications until the extraction of weights. The first one corresponds to the multiplication of the inverse matrix, $\hat{R}$ as calculated above, with the steering vector $e$, resulting in $L \times L$ multiplications and $(L - 1) \times L$ additions. The second one is the product of the quantity estimated in the previous step with the inverse steering vector. This output is just a scalar, adding $L$ more multiplications and $L - 1$ more additions to the total. The third one is the ratio of the two last quantities resulting in $L$ last multiplications. With all the multiplications and additions summed, $N_{\text{mult}}$ and $N_{\text{add}}$, the total number of calculations is summarized below:

$$N_{\text{mult}} = L^2(M - L + 2 + 2/L) . \quad (11)$$

$$N_{\text{add}} = L^2(M - L + 1 - 2/L^2) . \quad (12)$$

The same stages are followed for the subband MV beamformer as (5) and (9) indicate. This time $\hat{R}$ is the substitute of $R(\omega)$ in (8). However, there are important differences between the two beamformers, originating from the required frequency division. The $M$ sensor signals are divided into segments with the segment size, $N$, depending on the excitation pulse length. If $N$ is not equal to a power of 2 then sensor signals are zero-padded, thus here $N = 128$. They are then passed to the frequency domain with the use of Fast Fourier Transform (FFT) adding another $M(N/2)\log_2 N$ floating point operations to the total. Therefore each multiplication and addition between matrices encompasses complex numbers, and multiplication is equivalent to 6 operations and each addition to 2. Finally, the whole process is repeated for the estimation of a single weight as many times as the number of subbands, $N_{\text{sub}}$. As a consequence the final number of complex, operations is $N_{\text{sub}} \times N_{\text{mult}}$ for the multiplications and $N_{\text{sub}} \times N_{\text{add}}$ for the additions.

A total of approximately $2 \times 10^5$, calculations results from the sum of (11) and (12), in the case of the temporal MV beamformer. This number corresponds to calculations needed for a single MV weight and for a single emission. All the values of the parameters appearing in the above equations are given throughout the paper. The equivalent total for the subband approach is $5.6 \times 10^6$ for the point phantom and $30.5 \times 10^6$ for the
cyst phantom. Those results come up after taking into consideration the FFT computations, the number of subbands, and the fact that operations are between complex numbers. The difference between the two phantoms relies on the different number of subbands. As a consequence, the extraction of a single MV weight may require up to 152 times more calculations in the case of the subband approach.

Furthermore, Matlab functions of both MV methods have been timed while executed. The average time of the temporal MV function is measured at 0.45 ms, for the estimation of a single MV weight and for a single emission (element #64). The corresponding time of the frequency domain MV function is 9 ms for the point phantom and 23 ms for the cyst phantom. With execution time and number of calculations given, it is possible to extract the required floating-point operations per second (FLOPS). Resulting values are \(0.44 \times 10^9\) FLOPS for the temporal and \(0.62 \times 10^9\) and \(1.33 \times 10^9\) for the subband approach in the case of point and cyst phantom respectively.

![Figure 8: FWHM variation in respect to the total number of MV weights](image)

To the calculations described above there are two more factors that need to be considered. The first one is the number of weights, hence the number of image points, as an MV weight is estimated for each point. The total number of points depends on the sampling frequency and the bandwidth during the data acquisition. However, in MV adaptive beamforming, it is possible to apply further lateral division to the image and divide it into a larger number of parts. In Fig. 8 it can be found the effect of lateral points number, to the FWHM in the case of the point targets located at a depth of 40 mm and for a single emission (element #64). The higher this number is, the lower is the FWHM but the number of calculations is multiplied by the number of image points and thus weights. Images displayed in the current paper correspond to a lateral division of the image at 665 parts. The second factor is the number of emissions as it has already been mentioned. The number of numerical operations as described above should be multiplied by the number of emissions as the same process is repeated for each emission. For this reason results and images acquired from a single emission are also of a great importance.

6. CONCLUSION AND PERSPECTIVES

Temporal and subband minimum variance beamformers for medical ultrasound imaging, have been compared. The ultimate objective of both is to determine optimal apodization weights, but time-domain implementation is faster and less complex as the weight calculation is done directly from the data without the preprocessing that takes place in the subband approach. The methods are validated with the use of Field II simulated point and cyst phantoms. A full synthetic aperture sequence provides a lateral FWHM of 0.03 mm for both methods for a point target located at a depth of 40 mm. Furthermore, a -5 dB lower PSL value for the frequency domain implementation is noticed at the same depth. The contrast level from the cyst phantom is calculated at -54 dB and -63 dB for temporal and subband approach respectively. The number of calculations per second, required for all the above results is up to 3 times higher for the subband implementation, depending on the number of
subbands. A comparison between the two approaches has shown no significant differences in achieved resolution and contrast. This is an important observation for future attempts to use the MV adaptive beamformer for real time applications, where the complexity of the beamformer should be kept at the lowest possible level.

ACKNOWLEDGMENTS

This work was financially supported by Heriot-Watt University, by grant 024-2008-3 from the Danish Advanced Technology Foundation, and by BK Medical (BK Medical ApS, Herlev, Denmark).

REFERENCES


