CAPON BEAMFORMING FOR ACTIVE ULTRASOUND IMAGING SYSTEMS

Sverre Holm, Johan-Fredrik Synnevåg, and Andreas Austeng*

Department of Informatics, University of Oslo, Norway
* 2008/2009: Waves and Acoustics Laboratory, ESPCI, Paris, France

ABSTRACT

Medical ultrasound imaging has unique requirements regarding spatial and amplitude resolution, near-field focusing, wide bandwidth, and real-time operation. Only recently has Capon beamforming been adapted to this. We give examples of images of point targets, cysts, and regions dominated by speckle and discuss how subaperture smoothing, diagonal loading and range/time averaging must be balanced for best performance.

1. INTRODUCTION

Capon and other high resolution beamforming methods were first introduced in passive direction finding [1]. But more recently they have been applied to active imaging systems. Right now there is a large interest in the field of medical ultrasound imaging [2–5], but also in synthetic aperture radar [6] and sonar [7]. Our group has applied high resolution methods to active sonar [8] and medical ultrasound [9–14] with significant improvement in image resolution.

In medical ultrasound, images must be formed at a frequency of up to 10-15 MHz from hundreds of elements at frame rates up to several hundred per second, and with several hundred lines per image. Since the early 90's there has been a transition from analog to digital beamforming, and more recently also to 3-D systems that acquire data from 2-D arrays with thousands of elements. Still delay-and-sum beamforming is the method of choice, and some of the reasons why are:

- Ultrasound imaging has much more complex scenarios than the typical research paper's simulation. The medium will reflect at all angles and at all depths at the same time, and there are hardly any non-reflecting regions at all as in e.g. underwater sonar. Therefore it is customary to use both point target scenarios as well as cyst phantoms ("inverse point targets") for testing of systems and new methods. It is also important to preserve correctly the speckle characteristics of an image for a new beamforming method to be accepted.
- Medical ultrasound imaging needs both spatial and amplitude resolution. Many high-resolution imaging methods have a non-linear amplitude mapping as opposed to delay-and-sum. But e.g. the minimum variance method has a fairly well understood amplitude mapping, and in recent years new methods with more control of the amplitude have been introduced, such as APES (amplitude and phase estimation) [15]. As it has reduced spatial resolution, only the minimum variance method has been investigated here.
- Medical ultrasound operates in the near-field and requires both steering and focusing. The delays follow the pulse and adapt dynamically as the pulse travels in the medium. In addition, there is also dynamic variation of the aperture and the weighting (see Fig 1, explained later).
- Medical ultrasound is real-time. The high resolution beamforming methods may often require orders of magnitude more computations than the delay-and-sum method. Therefore, reduced complexity (see e.g. [14]) as well as improved processor capabilities are required.

Here we present the minimum variance method in this context, and discuss the considerations that are required in order to adapt its parameters to this application for best performance and robustness.

2. METHODS

2.1. Minimum variance beamformer

We assume a transducer consisting of M elements, and a beamformer operating on the sampled output of a subaperture of $L \leq M$ channels, $x_m[n]$. The output is:

$$z[n] = \sum_{m=0}^{L-1} w_m[n] x_m[n - \Delta_m[n]], \qquad (1)$$

where $w_m[n]$ is a time-varying, complex weight and $\Delta_m[n]$ is the delay for element m at sample n.

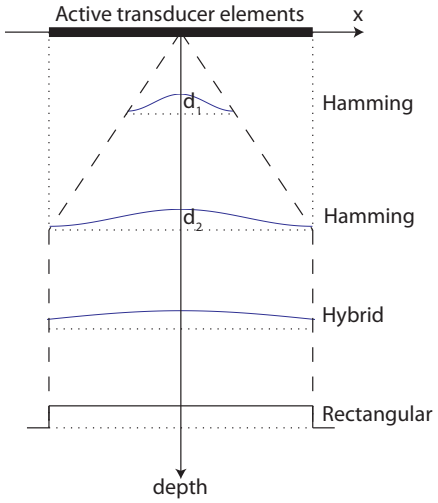

Fig. 1. Dynamic aperture and dynamic weighting as typically used in medical ultrasound

Usually the signal processing literature assumes static weights $w_m[n]$ in delay-and-sum beamforming. Thus the trade-off between mainlobe width and sidelobe level is done once and for all. However, in medical ultrasound, subjective image quality is improved by trading them differently at different depths. Near the transducer a smooth window (e.g. Hamming) is used for sidelobe suppression, while a rectangular window is often used at greater depths for maximum sensitivity and maximum penetration. In between, a transitional, hybrid window may be used as shown in Fig. 1. The figure also shows the concept of dynamic aperture, where only a subaperture of size $L < M$ is used near the transducer. This is based on the property that resolution in mm is proportional to range divided by active aperture (=f-number). In order to get an image with as uniform resolution as possible, the subaperture is allowed to grow with depth. Beyond a certain depth, d_2, all of the aperture is in use, so in this region resolution will deteriorate with depth.

In matrix form, (1) can be written as

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

where $\mathbf{w}[n] = [w_0^*[n] \cdots w_{L-1}^*[n]]^T$ and

$$\mathbf{X}[n] = [x_0[n - \Delta_0[n]] \cdots x_{L-1}[n - \Delta_{L-1}[n]]^T. \tag{3}$$

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:

$$\min_{\mathbf{w}[n]} E\left[|z[n]|^2\right] = \min_{\mathbf{w}[n]} \mathbf{w}[n]^H \mathbf{R}[n] \mathbf{w}[n]$$
$$\text{subject to } \mathbf{w}[n]^H \mathbf{a} = 1, \tag{4}$$

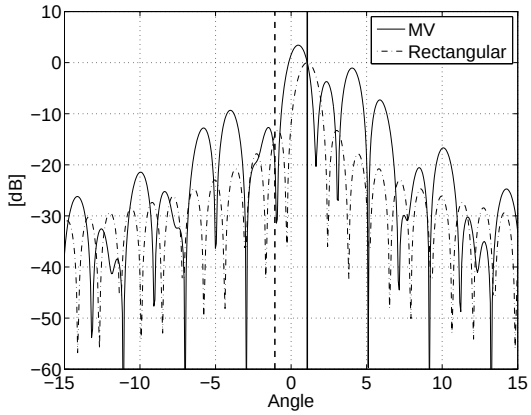

Fig. 2. Beampattern of the MV beamformer for two wire targets. The dashed line is the beampattern of a rectangular window. The vertical lines indicate the wire targets, and the solid line is the steering angle.

where $E[\cdot]$ denotes the expectation operator, $\mathbf{R}[n] = E\left[\mathbf{X}[n]\mathbf{X}[n]^H\right]$ is the spatial covariance matrix, and $\mathbf{a}$ is the *steering vector*. As we will describe later, the data has already been pre-beamformed, so $\mathbf{a}$ is simply a vector of ones. Eq. (4) has an analytical solution given by [1]:

$$\mathbf{w}[n] = \frac{\mathbf{R}[n]^{-1}\mathbf{a}}{\mathbf{a}^H \mathbf{R}[n]^{-1}\mathbf{a}}. \tag{5}$$

In order to get more degrees of freedom, it is an advantage to use complex data which can be formed by applying a Hilbert transform in the depth dimension or by down-mixing. This will allow beampatterns that are asymmetric around the steering angle.

The MV beamformer will allow large sidelobes in directions where there is little energy, and place zeros in directions where there is interference. This is illustrated in Fig. 2 which compares a MV beampattern with that of a rectangular window. We image two wire targets at $\pm$ 1 degrees, and steer towards the one at +1 degree (see Fig. 3 for examples of images). The solid vertical line indicates the angle of steering, and here the gain is unity. The dotted vertical line indicates the incoming angle of the reflection from the interfering scatterer. Here the response of the MV weights is close to zero. On the other hand, in regions where there is no signal energy, the beamformer allows large sidelobes (e.g. from 2-7 degrees). It should also be remarked that the mainlobe is asymmetric around the steering angle.

3. ESTIMATION OF THE SPATIAL COVARIANCE MATRIX

Estimation of the spatial covariance matrix, $\mathbf{R}[n]$ in (4), requires averaging of the observed signals as well as

regularization. This is done in order to deal with realistic signals and also to ensure that the rank of the covariance estimate is such that it can be inverted.

3.1. Subarray Averaging and Coherent Echoes

A system with active transmission will be dominated by coherent echoes, but the standard high resolution methods assume incoherency. Without any modification, the minimum variance (MV) method may be useful for beamwidth reduction and sidelobe suppression in simple single-target scenarios like isolated point targets, as in e.g. [2]. But for more complex scenarios, one needs to deal with target coherence.

Subaperture averaging is the primary method for this. It was developed in the context of passive systems with coherent sources [16, 17], but it is even more useful in an active system. The imaging scenario that best demonstrates the need is two very close targets. Using this, we have demonstrated that better resolution than DAS was possible even with coherent targets [9].

The processing takes place by shifting by one sample for each subaperture, resulting in a total of $M - L + 1$ subapertures, where the size of the subaperture L determines the degrees of freedom. A smaller L gives a more robust estimate at the expense of resolution.

There is an interesting result concerning the optimal subarray size in [18]. They found that for two coherent equipower closely spaced sources, a subarray size of $L = 0.6(M + 1)$ will maximize the distance between the signal and noise subspaces of the correlation matrix. It is our experience in medical ultrasound, that a subarray size in the order of 20-50% of M is a better value. This should indicate, not unexpectedly, that the imaging scenario is more complex than the model with two sources in white noise of [18]. This is also more in agreement with [19] which states that the value should be $L \leq M/2$ in order for the covariance matrix estimate to be invertible – without diagonal loading.

3.2. Diagonal Loading and Robustness

In addition to subaperture averaging, a second method for adapting the method to a practical imaging system is diagonal loading [20]. This is a regularization method that adds a constant to the diagonal of the covariance matrix estimate, i.e. $\mathbf{R}[n]$ is replaced by $\mathbf{R}[n] + \epsilon\mathbf{I}$ in (5). There exists several methods for calculating the value of Δ based on the uncertainty in the model parameters [20]. We use a simple approach, in which the amount of diagonal loading is proportional to the power in the received signals [10]:

$$\epsilon = \Delta \cdot \mathrm{tr}\{\mathbf{R}[n]\}, \tag{6}$$

where $\mathrm{tr}\{\cdot\}$ is the trace operator. It gives robustness to element amplitude errors [4]. We have shown that it also makes the method more robust to an incorrect assumed velocity of sound [9] and reduces sensitivity to phase aberrations [12].

Δ determines the relative weight given to the solution in Eq. 5 and the DAS solution. $\Delta = 1/\mathrm{L}$ corresponds to equal weighting of the two, in the sense that the traces of $\hat{\mathbf{R}}$ and $\epsilon\mathbf{I}$ are the same. We can view diagonal loading as addition of 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. Therefore 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 DAS beamforming.

3.3. Temporal Smoothing and Speckle Statistics

We have also found that it is an advantage to do smoothing in the time (depth) dimension. In a pulsed system, it is important to maintain the range resolution, so as little averaging as possible should be done in range. In [3] it was therefore proposed instead 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. In [10] we estimated the covariance matrix by averaging in the spatial domain only, and this works fine for isolated point targets. However, for objects consisting of unresolvable micro-structures (leading to speckle as in the later cyst examples), averaging solely in the spatial dimension results in different image statistics compared to images formed with the delay-and-sum beamformer. Speckle regions may appear less homogeneous, which is undesirable. In [11] we showed experimentally that averaging in range, as well as subaperture averaging, is required to retain similar speckle statistics. This would be a very interesting result to show theoretically as well.

The general estimate of the covariance matrix with subarrays of length L and temporal averaging over $Q = 2K + 1$ samples then becomes:

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

where

$$\bar{\mathbf{X}}_l[n] = [x_l[n] \cdots x_{l+L-1}[n]]^T. \tag{8}$$

3.4. Amplitude Spectral Capon Estimator

By using $\mathbf{R}[n] = \hat{\mathbf{R}}[n]$ in (5) we find the aperture shading and get the amplitude estimate for the total array as the average over the subarrays:

$$\hat{z}[n] = \frac{1}{M-L+1} \sum_{l=0}^{M-L} \mathbf{w}[n]^H \bar{\mathbf{X}}_l[n]. \quad (9)$$

This is the Amplitude Spectral Capon (ASC) estimate from [15], but there it was formulated in the time domain. An alternative is to average the individually beamformed subapertures on a power basis, but [15] says: *Note that the ASC ... in general will yield a different, and often preferable, spectral estimate ....* The same estimator was also used in [3].

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 the instantaneous observations at sample n. Hence, there is no smoothing in the time-domain and we retain the temporal resolution.

4. OTHER MEDICAL ULTRASOUND CONSIDERATIONS

4.1. Pre-beamforming

Due to the near-field focusing in an ultrasound system, there is a need to transform the data to the far field. This can be done by a transformation, but here we simply apply time-delays. Instead of just transforming the data to the far field with delays, we apply both focusing and steering delays, i.e. perform pre-beamforming or alignment of the data in time. Thus the full delay step of the delay-and-sum (DAS) beamformer is done. This means that the desired steering vector, $\mathbf{a}$, consists of just L ones.

4.2. Narrow vs. wide-band formulation

The high-resolution methods were originally formulated for narrow-band signals. Therefore, for a wide bandwidth signal such as medical ultrasound with relative bandwidths in the 50-100 % range, it is more correct to split the wide bandwidth into subbands and do independent or coupled beamforming per subband. But we have found that the performance is very good even without subband beamforming. This may partially explained by the fact that the pre-beamforming, which is a proper broadband, time-delay operation, already may have captured most of the required beamforming. What is left is only a slight correction term, represented by the weight vector, $\mathbf{w}$. A full broadband approach has however been used in the same application by others, see e.g. [5], and it is of interest to pursue this approach further.

4.3. Spatial Sampling

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. Sonars often use omnidirectional transmission, but in medical ultrasound the transmitter beam needs to be almost as sharp as the receiver beam to get acceptable contrast resolution. As the transmitter beam gets sharper, more beams are required in order to cover the full image, so the frame rate will drop. Also the need for closer sampling in angle grows, and with too sharp beams it is easy to miss small point targets. Therefore angular sampling needs to be balanced with resolution. A controlled degradation of resolution, e.g. with diagonal loading, can be used to balance the sampling density and the resolution so that all point targets will be properly sampled even under non-ideal conditions such as incorrect propagation speed or phase aberrations.

4.4. Dynamic Aperture Considerations

The minimum variance method also has to deal with dynamic apertures as shown in Fig. 1. There can be several strategies for doing this.

In our preferred scheme, the number of subapertures $N_{sub} = M - L + 1$ is to be maintained at all depths. This will give as even robustness as possible while resolution will vary with depth. Let the number of available elements be $M[n] \leq M$, then the new subaperture length should be $L[n] = M[n] - N_{sub} + 1$. However, $L[n]$ may become negative, so the value is restricted to a minimal value L_{min}. Therefore the dynamic aperture region will be characterized by a constant subaperture length, L_{min}, closest to the transducer, depth $< d_1$ where also the number of subapertures will shrink the closer one gets to the transducer. Then comes a region with a constant number of subapertures and increasing subaperture size ($d_1 <$ depth $< d_2$), and finally as the maximum aperture is reached at depth d_2, both L and N_{sub} will be constant until the maximum depth of interest is reached.

Another possibility is to try to keep the subaperture as close to the desired size as possible so the number of subapertures will vary more than in the first scheme. At small enough depths, there will be only a single subaperture left, so robustness may suffer as well, as the rank of the covariance estimate. In order to avoid possible numerical problems, a third scheme somewhat in between the two former ones, is to try to maintain the rank of the covariance estimator as depth varies.

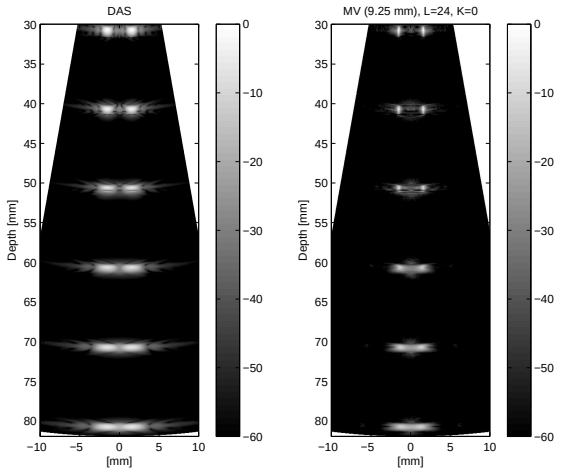

Fig. 3. Simulated wire targets, 3.5 MHz, dynamic Tx and Rx focus. Left: DAS using 96 element, 18.5 mm transducer, rectangular window. Right: MV using 48 element, 9.25 mm transducer.

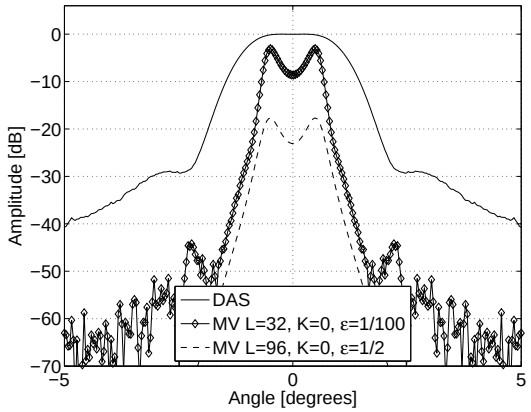

Fig. 4. Response of DAS (solid), MV with heavy diagonal loading, but no subarray averaging (dashed) and MV with subarray = 32 and a small amount of diagonal loading (solid, diamonds) for two wires at 70 mm.

5. EXAMPLES

We have used Field II which is a de facto standard for medical ultrasound simulations from Denmark Technical University, [21] for the simulations. All of the cases use the full aperture, i.e. depths are beyond d_2 in Fig. 1.

5.1. Point Target Performance

First we simulated a 3.5 MHz transducer (typical for adult cardiology). Fig. 3 shows results for pairs of wire targets spaced 1.5 mm apart. To the left it shows the results from a 96 element/18.5 mm probe using DAS beamforming and rectangular weighting. To the right is the same image with half the aperture running MV with subarray length $L = 24$, $K = 0$ (no time averaging), and a small amount of diagonal loading ($\Delta = 1/(100L)$). A detailed analysis shows that the MV beamwidth is less than 1/4th of that of DAS with twice the aperture, and that the sidelobes are 8 to 12 dB lower.

The pair of targets at 70 mm were then imaged by the same 96 element transducer. Full dynamic focus was applied, both for the receiver and transmitter. Fig. 4 shows results for the wire pair with DAS, and two cases of MV (heavy diagonal loading alone $\Delta = 1/(2L)$, and subarrays of size 32 + light diagonal loading, $\Delta = 1/(100L)$). Signal cancellation occurs when there is no subarray averaging and the level falls by about 15 dB. If the pair of targets at 80 mm had been replaced with a single wire and a similar anlysis done, there would not have been any signal cancellation for either set of parameters. This demonstrates that closely spaced targets in an active imaging system will generate coherent echoes.

5.2. Cyst Performance

Fig. 5 shows results for a cyst phantom. Fig. 5(a) shows a 18.5 mm probe (96 elements) using DAS and rectangular weighting. Fig. 5(b) shows an image using half the aperture and MV beamforming with $L = 24$ and $K = 11$. This corresponds to temporal averaging over approximately the length of the transmitted pulse. Even though only half the aperture is used, the contrast and point target resolution is similar or better.

Fig. 6 shows results from a 96 element, 3.5 MHz, 18.5 mm transducer for a second cyst phantom, 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 frequency response of the transducer the effective center frequency was approximately 2.5 MHz in the first case. In Fig. 6(a) we have simulated 1.75 MHz excitation and MV receive beamforming with $L = 40$ and $K = 11$. Compared to Fig. 6(b) the lateral resolution is about the same and the point scatterers are much better defined. However, we see some distortion of the edges of the cyst.

The examples demonstrate that better resolution may be obtained even when only half the aperture has been used. This may open up new imaging applications, which up to now were not possible due to lack of access of a transducer. The last example also showed how the improved resolution could be used to lower the frequency, and thus achieve higher penetration. This may also enable new imaging applications. Other examples can be found in [13] which also gives examples of how high spatial resolution can be traded

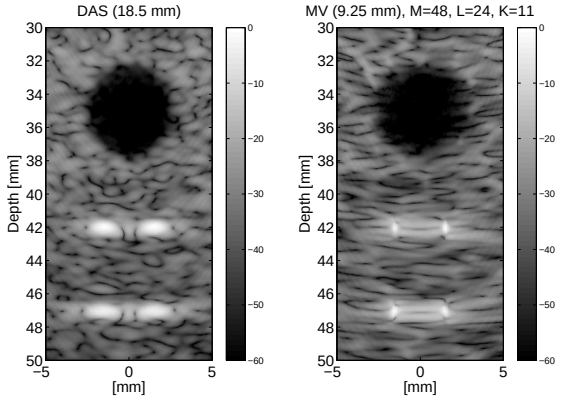

Fig. 5. Simulated cyst phantoms using 3.5 MHz transducers. (a) DAS with 96 element, 18.5 mm transducer (rectangular window) (b) MV with 48 element, 9.25 mm transducer ($L = 24$ and $K = 11$). Tx focus was 35 mm.

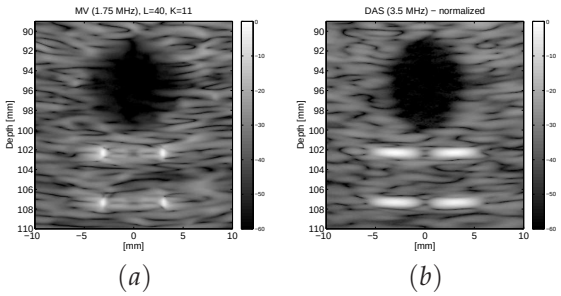

Fig. 6. Simulated cyst phantoms using a 18.5 mm, 96 element transducers. (a) MV (1.75 MHz, $L = 40$ and $K = 11$). (b) DAS (3.5 MHz). Tx focus was 100 mm.

for higher frame rates, which may be important for imaging fast-moving structures like the cardiac valves.

6. CONCLUSION

Medical ultrasound imaging, with its scenario of scatterers everywhere, a need for both spatial and amplitude resolution, near-field focusing as well as steering, and real-time requirement, presents a challenge for adaptation of Capon beamforming. In this paper we show how it can be done by using subaperture smoothing for dealing with the coherent echoes that the transmitter always will generate, diagonal loading for robustness to medium uncertainties, and averaging in range/time for improved imaging of speckle.

7. REFERENCES

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

[2] J. A. Mann and W. F. Walker, "A constrained adaptive beamformer for medical ultrasound: Initial results," *Proc. IEEE Ultrason. Symp.*, vol. 2, pp. 1807–1810, Oct. 2002.

[3] Magali Sasso and Claude Cohen-Bacrie, "Medical ultrasound imaging using the fully adaptive beamformer," *Proc. IEEE Int. Conf. Acoust. Speech, Sign. Proc.*, vol. 2, pp. 489–492, March 2005.

[4] Zhisong Wang, Jian Li, and Renbiao Wu, "Time-delay- and time-reversal-based robust Capon beamformers for ultrasound imaging," *IEEE Trans. Med. Imag.*, vol. 24, pp. 1308–1322, Oct. 2005.

[5] 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.

[6] Jian Li and P. Stoica, "An adaptive filtering approach to spectral estimation and SAR imaging," *IEEE Trans. Signal Process.*, vol. 44, no. 6, pp. 1469–1484, Jun 1996.

[7] Kam W. Lo, "Adaptive array processing for wide-band active sonar," *IEEE J. Ocean. Eng.*, vol. 29, no. 3, July 2004.

[8] A. Rønhovde, L. Yang, T Taxt, and S. Holm, "High-resolution beamforming for multibeam echo sounders using raw EM3000 data," in *Proc. IEEE Oceans*, Seattle, WA, Sept. 1999, pp. 923–930.

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

[10] J.-F. Synnevåg, A. Austeng, and S. Holm, "Adaptive beamforming applied to medical ultrasound imaging," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 54, pp. 1606–1613, Aug. 2007.

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

[12] A. Austeng, T. Bjaastad, J.-F. Synnevåg, S.-E. Masoy, H. Torp, and S. Holm, "Sensitivity of minimum variance beamforming to tissue aberrations," in *Proc. IEEE Ultrasonics Symposium*, Beijing, China, Nov 2008.

[13] J.-F. Synnevåg, A. Austeng, and S. Holm, "Benefits of high-resolution beamforming in medical ultrasound imaging," *submitted*, 2008.

[14] J.-F. Synnevåg, S. Holm, and A. Austeng, "A low complexity data-dependent beamformer," in *Proc. IEEE Ultrasonics Symposium*, Beijing, China, Nov 2008.

[15] T. Ekman, A. Jakobsson, and P. Stoica, "On efficient implementation of the Capon algorithm," in *Proc. EUSIPCO*, Finland, Sept. 2000.

[16] 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.

[17] Tie-Jun Shan, Mati Wax, and Thomas 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, Aug. 1985.

[18] A.B. Gershman and V.T. Ermolaev, "Optimal subarray size for spatial smoothing," *IEEE, Sign. Proc. Let.*, vol. 2, no. 2, pp. 28–30, 1995.

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

[20] Jian Li, P. Stoica, and Zhisong Wang, "On robust Capon beamforming and diagonal loading," *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1702–1715, July 2003.

[21] J.A. Jensen, "Field: A Program for Simulating Ultrasound Systems," *Med. Biol. Eng. and Comp.*, vol. 34, pp. 351–352, 1996.