Spatial compression in ultrasound imaging

Van Der Meulen, Pim; Kruizinga, Pieter; Bosch, Johan G.; Leus, Geert

DOI
10.1109/ACSSC.2017.8335502

Publication date
2018

Document Version
Final published version

Published in
Conference Record of 51st Asilomar Conference on Signals, Systems and Computers, ACSSC 2017

Spatial Compression in Ultrasound Imaging

Pim van der Meulen*, Pieter Kruizinga*†, Johannes G. Bosch†, Geert Leus*

*Delft University of Technology, Delft, Netherlands
†Erasmus Medical Center, Rotterdam, Netherlands

***Abstract*—High quality three dimensional ultrasound imaging is typically attained by increasing the amount of sensors, resulting in complex hardware. Compressing measurements before sensing addresses this problem, and could enable new clinical applications. We have developed an analogue compression technique, by positioning a plastic coding mask in front of the aperture, which distorts the ultrasound field by inducing varying local echo delays. This results in a compression of the spatial ultrasound field across the sensor surface, while retaining sufficient information for 3D imaging. Using only a single sensor, complementary measurements can be obtained by rotation of the sensor and the mask to increase the conditioning of the reconstruction problem. In this work, we study a method to optimize the shape of the coding mask. To this end, we define an approximate signal model that captures the ultrasound response of the mask, and use it to pose mask shape optimization as a sensor selection problem. We solve it by relaxing it to a convex problem, as well as by using a greedy selection method. Our simulation results show that these approaches are able to outperform the random design strategy, in particular when mask rotations are included in the problem.**

I. Introduction

Medical ultrasound imaging relies on the transmission of short ultrasonic waves into the tissue and the reception of the reflected echoes resulting from acoustic impedance contrasts between different tissues. This type of imaging is conventionally done using an array of sensors which are lined up in either one or two dimensions to provide a 2D or 3D image of the object of interest. For 2D imaging, arrays consist of 64 up to 256 individual sensors. For 3D imaging, these numbers are squared, resulting in very large complex arrays with integrated electronics and signal reduction in order to facilitate easy integration with existing ultrasound acquisition systems that have limited numbers (typically 128) of channels available. This technological difficulty is one the prominent reasons why 3D ultrasound has not yet been widely adopted throughout the clinical arena.

A possible solution to make 3D ultrasound more accessible is to take advantage of the recent advances in the field of compressive imaging [1–4] where the data compression takes place before the sampling, thereby lowering the need for fully populated sensory arrays. This compressive sampling or imaging approach and its use for ultrasound imaging has been explored by several research groups [5–9]. Most of these approaches use temporal compression of the array measurements or select fewer channels from a full array. Just recently, we successfully showed that 3D compressive ultrasound imaging is also possible using only one sensor and an irregular aperture coding mask that is placed in front of the sensor [10]. This aperture mask ensures that every pixel in the image is uniquely contained in the compressed measurement by locally adding temporal delays to the transmitted and received wavefield. These temporal delays result from the local thickness variations of the mask which is made from a material with a different acoustic impedance than the surrounding medium. Additional measurements obtained by rotating this mask allow for new measurements that can be used to reconstruct the object of interest.

In this paper, we explore optimal mask design for single-sensor imaging. To approximate the influence of the mask on the received ultrasound field, we discretize the mask surface, and regard each discretized point as an independent sensor that measures the incident ultrasound field and delays it according to the mask thickness at that point on the mask. We are then concerned with the following questions: *(i)* how does the mask discretization affect the imaging performance, *(ii)* how does the mask thickness affect the imaging performance, and *(iii)*, given a mask discretization and maximum mask thickness, how do we optimize the shape of the mask? The second question was investigated in [10], where it is shown how the mask thickness (and consequently, the maximum locally induced delays) can improve the imaging performance, although a random mask design strategy is used. In this work we make a first attempt to answer the third question. To this end, we will pose mask optimization as a sensor selection problem. This is accomplished by defining a number of sensors per mask surface point, corresponding to different mask thickness levels. The mask shape algorithm should then try to select one such a sensor for each surface point according to a given performance criterium.

The goal in sensor selection is to select a given number of sensors from a larger group of sensors, such that the desired estimation (e.g. beamforming) or detection performance is optimal. Popular sensor selection approaches in literature use convex optimization strategies [11–15], or submodular approaches [16, 17]. There is a distinct difference between our optimization problem and those in the classical sensor selection studies just mentioned. In the typical sensor selection problem, each sensor candidate's data will be fully accessible after selection. However, for sensing using a coding mask, selected sensors are summed into a single final measurement,

This work is part of the ASPIRE project (project 14926 within the STW OTP programme), and the PUMA project (project 13154 within the STW OTP programme), which are financed by the Netherlands Organisation for Scientific Research (NWO).

as will be shown later. A second difference is that we want to select one sensor from a group of sensors and this for several groups, instead of a sparse selection from a single group. This is similar to [14], except that we do not associate a cost to sensors within a candidate group.

The remainder of this work is organized as follows. In the next section, we define the approximate signal model for a sensor with an aperture mask, and derive a convex optimization approach to solve the mask optimization problem. We also propose a simpler greedy approach. In section III, we show that both the convex and greedy approaches can lead to good results when optimizing the mask for a 1D line sensor and few candidate sensors. We additionally optimize a mask for a 2D circular sensor that can be used for the imaging setup in [10] and obtain good results for various SNR scenarios. In the final section we discuss the results and conclude this work.

II. Methods

In order to analyse the role of the coding mask, we use the same approximate model as in [10], where the mask surface is discretized in the width and length (x and y) dimensions into many smaller squares. Any mask is then approximated by a collection of small pillars (channels), with the height (z dimension) of each channel equal to the mask thickness at that position. We then assume that every channel measures the pulse-echo ultrasound field at the top of the channel, independently from other channels, and delays the measured signal in time according to the mask thickness of the channel and the speed of sound inside the mask. This is illustrated in Fig. 1. The output signal of the sensor is the summation of all channel signals, since the output signal at any point in time is the integration of the ultrasound field over the sensor surface, convolved with the sensors electromechanical impulse response.

Assuming a linear scattering model, and denoting the sampled measurements $a[n]$ ($n \in \{0, 1, \ldots, N-1\}$) on channel $s \in \{0, 1, \ldots, S-1\}$ from pixel $m \in \{0, 1, \ldots, M-1\}$ by the vector $\mathbf{a}_{s,m} \in \mathbb{R}^N$, the measured echo signal $\mathbf{y}_s$ on channel s resulting from the set of M scatterers is denoted as

$$\begin{aligned}\mathbf{y}_s &= [\mathbf{a}_{s,0}\ \mathbf{a}_{s,1}\ \ldots\ \mathbf{a}_{s,M-1}]\,\mathbf{x} \\ &= \mathbf{A}_s\mathbf{x}.\end{aligned}$$

where $\mathbf{x} \in \mathbb{R}^M$ contains the scattering coefficient per pixel. Any additional measurements obtained by rotations can be stacked vertically into $\mathbf{y}_s$ and $\mathbf{A}_s$. Since the ultrasound transducer behind the mask effectively integrates the entire ultrasound field over its surface, the transducer output signal is approximated as the summation of all S channels:

$$\mathbf{y} = \sum_{s=0}^{S-1} \mathbf{y}_s = \sum_{s=0}^{S-1} \mathbf{A}_s\mathbf{x}. \tag{1}$$

Based on the proposed model, we are able to pose mask optimization as a sensor selection problem; each channel can have one of several thickness levels, where each level corresponds to a different measurement of the ultrasound field.

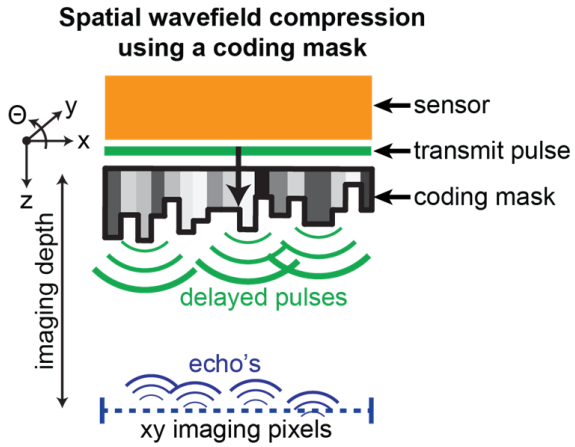

Fig. 1: Single sensor imaging setup. The coding mask distorts measurements in such a way that previously ambiguous image pixels become distinguishable based on (multiple) temporal pulse-echo measurements. For the approximate model used in this study, the mask is discretized into several channels, indicated by different shaded of grey. We assume that each channel measures the incident wavefield independently, and that the spatial integration of the ultrasound field on the sensor surface can be approximated by summing the outputs of all channel signals.

In other words, the measurement of each thickness level can be represented by a candidate sensor. The goal is then to select one sensor per channel in such a way that the best possible imaging performance is obtained.

Suppose there are R potential sensors (or: R potential thickness levels) for each channel s, and $w_{r+Rs} \in \{0, 1\}$ indicates whether sensor candidate $r \in \{0, 1, \ldots, R-1\}$ is used for channel s or not. Then the output signal is written as

$$\begin{aligned}\mathbf{y} &= \sum_{t=0}^{RS-1} w_t\mathbf{A}_t\mathbf{x} \\ &= \mathbf{A}(\mathbf{w})\mathbf{x},\end{aligned} \tag{2}$$

where $\mathbf{w} = [w_0, w_1, \ldots, w_{RS-1}]^\mathsf{T}$. Note that in this formulation only one sensor should be selected per channel, resulting in only S non-zero values for $\mathbf{w}$. More specifically, if $\mathbf{w}_s$ only contains the entries of $\mathbf{w}$ that correspond to the R selection coefficients for channel s, i.e., $\mathbf{w}_s = [w_{Rs}, w_{Rs+1}, \ldots, w_{R(s+1)-1}]^\mathsf{T}$, then we want $\|\mathbf{w}_s\|_0 = 1$ for all channels s.

Our goal is to optimally select sensors resulting in the best performance. Many different performance criteria can be considered, such as the largest or smallest eigenvalue of $\mathbf{A}(\mathbf{w})$, or the determinant of $\mathbf{A}(\mathbf{w})$. A first approach to find an optimal mask is to solve the problem by minimizing the mean squared error (MSE) under additive zero-mean i.i.d. white Gaussian measurement noise. More specifically, we minimize

$$f(\mathbf{w}) = \text{trace}\left[\left(\mathbf{A}(\mathbf{w})^\mathsf{H}\mathbf{A}(\mathbf{w})\right)^{-1}\right]. \tag{3}$$

Using (2), $\mathbf{A}(\mathbf{w})^{\mathsf{H}}\mathbf{A}(\mathbf{w})$ is equal to

$$\mathbf{A}(\mathbf{w})^{\mathsf{H}}\mathbf{A}(\mathbf{w}) = \sum_{t=0}^{RS-1}\sum_{u=0}^{RS-1} w_t w_u \mathbf{A}_t^{\mathsf{H}}\mathbf{A}_u. \tag{4}$$

Ideally, the optimization problem can then be posed as

$$\begin{aligned} \min_{\mathbf{w}} \quad & f(\mathbf{w}) \\ \text{s.t.} \quad & \|\mathbf{w}_s\|_0 = 1, \quad s = 0, 1, ..., S-1 \\ & [\mathbf{w}]_t \in \{0,1\}, \quad t = 0, 1, \ldots, RS-1 \end{aligned} \tag{5}$$

Unfortunately, both the objective function and the constraints are not convex in $\mathbf{w}$, so we need to apply some relaxation techniques in order to solve (5).

Using a common technique to make the objective function convex, $\mathbf{w} \in \{0,1\}^{RS}$ is lifted to obtain the matrix $\mathbf{W} = \mathbf{w}\mathbf{w}^{\mathsf{T}}$, with $\mathbf{W} \in \{0,1\}^{RS\times RS}$. Consequently, (4) is equivalent to

$$\mathbf{A}(\mathbf{w})^{\mathsf{H}}\mathbf{A}(\mathbf{w}) = \sum_{t=0}^{RS-1}\sum_{u=0}^{RS-1} [\mathbf{W}]_{t,u} \mathbf{A}_t^{\mathsf{H}}\mathbf{A}_u. \tag{6}$$

The non-convex equality $\mathbf{W} = \mathbf{w}\mathbf{w}^{\mathsf{T}}$ together with $\mathbf{W} \in \{0,1\}^{RS\times RS}$ is relaxed to $\mathbf{W} - \mathbf{w}\mathbf{w}^{\mathsf{T}} \succeq 0$ and $\text{diag}(\mathbf{W}) = \mathbf{w}$. Since all coefficients of $\mathbf{w}$ will be non-negative using this relaxation, we also drop the Boolean constraint $\mathbf{W} \in \{0,1\}^{RS\times RS}$. The l_0 norm is finally relaxed to the l_1 norm, which due to the positivity of the elements of $\mathbf{w}$, is simply the sum of the components of $\mathbf{w}_s$. The relaxed problem is thus stated as

$$\begin{aligned} \min_{\mathbf{w},\mathbf{W}} \quad & f(\mathbf{W}) \\ \text{s.t.} \quad & \mathbf{1}^{\mathsf{T}}\mathbf{w}_s = 1, \quad s = 0, 1, ..., S-1 \\ & \mathbf{W} - \mathbf{w}\mathbf{w}^{\mathsf{T}} \succeq 0 \\ & \text{diag}(\mathbf{W}) = \mathbf{w} \end{aligned} \tag{7}$$

A simple scheme to obtain a discrete solution from the relaxed problem, is by taking the sensor corresponding to the maximum of the solution $\hat{\mathbf{w}}_s$ per channel. We denote this solution by $\hat{\mathbf{w}}^{\text{direct}}$:

$$\left[\hat{\mathbf{w}}_s^{\text{direct}}\right]_r = \begin{cases} 1, & \text{if} \quad r = \arg\max_v \left[\hat{\mathbf{w}}_s\right]_v \\ 0, & \text{otherwise} \end{cases}, \tag{8}$$

assuming each $\hat{\mathbf{w}}_s$ has a unique maximum.

Greedy optimization alternative

As an alternative rounding method to further optimize the solution obtained from the relaxed problem (7), we can use a greedy optimization scheme, by iteratively (re)selecting the sensor on the channel that most increases the overall MSE, until a (local) optimum is reached. In other words, we use the algorithm as specified in Alg. 1, using $\mathbf{w}^{\text{in}} = \hat{\mathbf{w}}^{\text{direct}}$.

Alternatively, we could use the greedy optimization scheme by using a flat mask as a starting point (i.e., $\forall s \in \{0, 1, \ldots, S-1\}, \left[\mathbf{w}_s^{\text{in}}\right]_0 = 1$), instead of the convex solution.

Algorithm 1 Greedy optimization scheme

Input: $\mathbf{w}^{\text{in}}$
Output: $\hat{\mathbf{w}}$
$\alpha = f(\mathbf{w}_{in}), it = 0$, $\hat{\mathbf{w}}^{(0)} = \mathbf{w}^{\text{in}}$
while $\alpha > 0$ **do**
  $it = it + 1$
  for s in S **do**
    for r in R **do**
      $\hat{\mathbf{w}}^{temp} = \hat{\mathbf{w}}^{(it-1)}$
      $\hat{\mathbf{w}}_s^{temp} = \mathbf{0}$
      $\left[\hat{\mathbf{w}}_s^{temp}\right]_r = 1$
      $\epsilon_{s,r} = f(\hat{\mathbf{w}}^{temp})$
    end for
  end for
  $\hat{\mathbf{w}}^{(it)}$is the $\hat{\mathbf{w}}^{temp}$ corresponding to the minimum of $\epsilon_{s,r}$
  $\alpha = f(\hat{\mathbf{w}}^{(it-1)}) - f(\hat{\mathbf{w}}^{(it)})$
end while
$\hat{\mathbf{w}} = \hat{\mathbf{w}}^{(it)}$

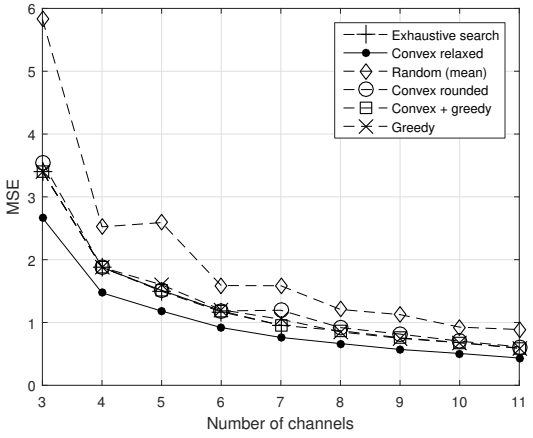

Fig. 2: Expected MSEs for varying number of channels S.

Although we have no optimality guarantees in this case, it is much more time and memory efficient than solving the convex relaxation in (7).

III. Results

We first test the various optimization schemes for a relatively small problem for $R = 5$ and varying values of S, without using rotations. We optimize delays for a linear array of 1.1 cm, with a maximum mask thickness of 2 mm, excited with a single transmitted 5 MHz Gaussian pulse, and pixels in a 5 by 3 grid in a (z, x) plane covering a 4 by 2 cm region. To prevent solutions to (7) which are symmetric after rounding, we shift the grid in the x-dimension such that it is non-symmetrically positioned with respect to the z-axis. The results are shown in Fig. 2. We take the results of the convex relaxation and obtain a discrete solution using the direct rounding scheme. We also plot the results of the MSE before rounding, which acts as a lower bound on the best obtainable MSE. When tractable, we plot the MSE of the

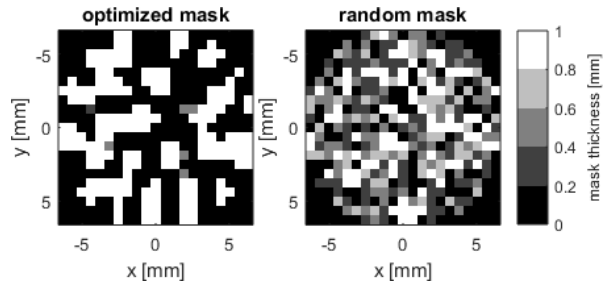

Fig. 3: Optimized and random mask used for the results in Figs. 4 and 5.

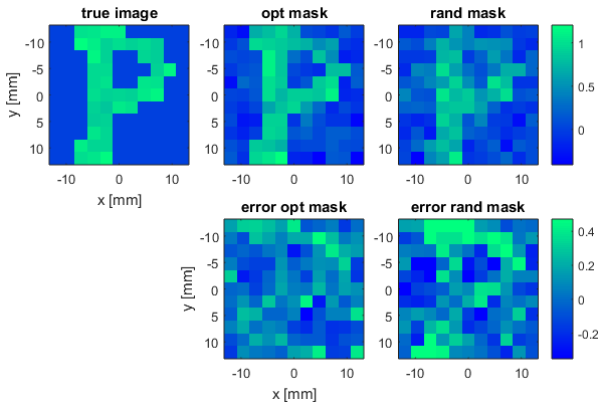

Fig. 4: Reconstruction of a letter P image. The top figures show the true image and its reconstructions using both the optimized mask and a random mask. The bottom figures show the error maps for each reconstruction result.

best discrete solution obtained by an exhaustive search. We also generated 500 random masks and plot their mean MSE. For the random masks, the probability of selection for each sensor candidate per sensor is equivalent, and independent of other channels. From Fig. 2 we conclude that for small imaging problems, all techniques can do better than a random optimization strategy, and are near-optimal.

Next, we optimize for a circular mask as used in [10], and include rotations in our signal model. We define a disc covering the area of a circular transducer of 12.6 mm diameter, and cover it with 373 channels. For each channel, we take five sensor candidates corresponding to five mask thickness levels uniformly distributed between 0 and 1 mm per channel. We use an image pixel grid of 2.4 by 2.4 cm at a depth of 1.2 cm, for a total of 121 pixels, and rotate the transducer 40 times, uniformly distributed over 360 degrees. We assume that the transducer simultaneously transmits on 4, 5, and 6 MHz. Typically, ultrasound is used with band limited pulses, but to reduce the problem size we only use three single frequencies from a typical ultrasound bandwidth. Together with 40 rotations, this results in 120 measurements. Since the problem size is relatively large, we only use the greedy rounding scheme initialized with a flat mask. Moreover, since $\mathbf{A}(\mathbf{w})$ is typically ill-posed due to high pulse-echo signal correlations between neighbouring pixels, we inevitably have to regularize the least squares estimate of the image. Hence, we optimize for the Bayesian MSE instead, $BMSE = \text{trace}(\mathbf{C}_x^{-1} + \mathbf{A}(\mathbf{w})^{\mathsf{H}}\mathbf{C}_n^{-1}\mathbf{A}(\mathbf{w}))^{-1}$, assuming that $\mathbf{x}$ and the measurement noise are zero-mean, i.i.d. white Gaussian, and have covariance matrices $\mathbf{C}_x = \mathbf{I}$ and $\mathbf{C}_n = \sigma_n^2\mathbf{I}$. When optimizing the mask, we set σ_n^2 such that it corresponds to an expected SNR of 20 dB for a flat mask. Fig. 3 shows the obtained optimized mask in the left pane. It shows a definite structure, and tends to select either the smallest or largest thickness levels. It also tends to select neighbouring elements, resulting in clusters of maximum and minimum thickness levels.

After optimization, we first compare the reconstruction of a fixed image with the optimized mask to the reconstruction with a random mask for a high SNR scenario (40 dB). The random masks were generated as described before. To estimate the image, we use the Bayesian linear estimator $\hat{\mathbf{x}} = (\mathbf{C}_x^{-1} + \mathbf{A}(\mathbf{w})^{\mathsf{H}}\mathbf{C}_n^{-1}\mathbf{A}(\mathbf{w}))^{-1}\mathbf{A}(\mathbf{w})^{\mathsf{H}}\mathbf{C}_n^{-1}\mathbf{y}$. The noise is zero-mean white Gaussian. The results using the masks in Fig. 3 are shown in Fig. 4. Clearly, the optimized mask is better able to exploit the additional measurements obtained by rotation. Pixels at the outer edges of the image, where the pulse-echo signals change more rapidly between rotations than pixels close to the centre, are significantly better reconstructed using the optimized mask than the random mask.

As a final experiment, we iteratively reconstruct many random images, and carry out the reconstruction as before, using both the optimized mask and a random mask. Each value of $\mathbf{x}$ is drawn independently from a zero-mean unit-variance Gaussian distribution. For each random image, a random mask is generated as described earlier. The additive noise is zero-mean white Gaussian noise. We then compare the normalized MSE between each random image and its reconstruction, $NMSE = \|\mathbf{x} - \hat{\mathbf{x}}\|_2^2/\|\mathbf{x}\|_2^2$, for many realizations and several SNR scenarios. The results are shown in Fig. 5. The results in Figs. 3-5 show that the selection algorithm is able to exploit the extra structure provided by taking rotations into account, and that the optimized mask is able to outperform the randomly generated masks.

IV. Conclusion and Discussion

In this work, we proposed a convex and greedy method to optimize the coding mask for a single-sensor imaging setup, and have compared their effectiveness to a random design strategy. For a small problem size, we tested both methods, and both show near-optimal performance, outperforming the random design strategy. For a larger problem similar to our previous work, we tested the greedy approach, and show that it performs significantly better than a randomly designed mask for various SNR scenarios. Although our cost-function is currently not sub-modular, we were able to obtain good results using the greedy selection method with the current model parameters. Solving this problem by finding a submodular surrogate for our sensor selection problem can provide us with optimality bounds, and is a topic of future research.

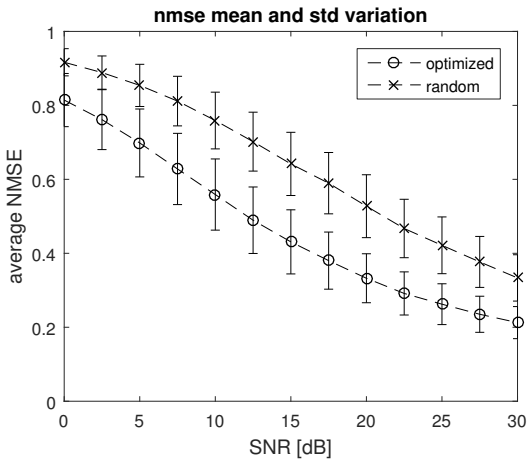

Fig. 5: NMSE spread for many image reconstructions for various SNR scenarios. The vertical bars indicate the standard deviation of the NMSE for 1000 random images per SNR scenario.

Currently, an approximate signal model is used. An extensive validation of this model has not yet been done, and may invalidate our results in real experiments, although [10] shows that correlations between pixels for the approximate model and real measurements are very similar.

REFERENCES

[1] D. L. Donoho, "Compressed sensing," *IEEE Transactions on information theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[2] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE signal processing magazine*, vol. 25, no. 2, pp. 21–30, 2008.

[3] R. G. Baraniuk, "Compressive sensing [lecture notes]," *IEEE signal processing magazine*, vol. 24, no. 4, pp. 118–121, 2007.

[4] M. F. Duarte, M. A. Davenport, D. Takbar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE signal processing magazine*, vol. 25, no. 2, pp. 83–91, 2008.

[5] T. Chernyakova and Y. Eldar, "Fourier-domain beamforming: the path to compressed ultrasound imaging," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 61, no. 8, pp. 1252–1267, 2014.

[6] M. F. Schiffner and G. Schmitz, "Fast pulse-echo ultrasound imaging employing compressive sensing," in *Ultrasonics Symposium (IUS), 2011 IEEE International*. IEEE, 2011, pp. 688–691.

[7] H. Liebgott, R. Prost, and D. Friboulet, "Pre-beamformed rf signal reconstruction in medical ultrasound using compressive sensing," *Ultrasonics*, vol. 53, no. 2, pp. 525–533, 2013.

[8] C. Quinsac, A. Basarab, and D. Kouamé, "Frequency domain compressive sampling for ultrasound imaging," *Advances in Acoustics and Vibration*, vol. 2012, 2012.

[9] G. David, J.-l. Robert, B. Zhang, and A. F. Laine, "Time domain compressive beam forming of ultrasound signals," *The Journal of the Acoustical Society of America*, vol. 137, no. 5, pp. 2773–2784, 2015.

[10] P. Kruizinga, P. van der Meulen, A. Fedjajevs, F. Mastik, G. Springeling, N. de Jong, J. G. Bosch, and G. Leus, "Compressive 3d ultrasound imaging using a single sensor," *Science Advances*, vol. 3, no. 12, Dec. 2017.

[11] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 451–462, 2009.

[12] S. P. Chepuri and G. Leus, "Sparsity-promoting sensor selection for non-linear measurement models," *IEEE Transactions on Signal Processing*, vol. 63, no. 3, pp. 684–698, 2015.

[13] ——, "Sparse sensing for estimation with correlated observations," in *Signals, Systems and Computers, 2015 49th Asilomar Conference on*. IEEE, 2015, pp. 581–585.

[14] O. M. Bushnaq, T. Y. Al-Naffouri, S. P. Chepuri, and G. Leus, "Joint sensor placement and power rating selection in energy harvesting wireless sensor networks," in *Signal Processing Conference (EUSIPCO), 2017 25th European*. IEEE, 2017, pp. 2423–2427.

[15] A. Abdi and F. Fekri, "Optimal sensor selection in the presence of noise and interference," in *Information Theory (ISIT), 2017 IEEE International Symposium on*. IEEE, 2017, pp. 2378–2382.

[16] M. Shamaiah, S. Banerjee, and H. Vikalo, "Greedy sensor selection: Leveraging submodularity," in *Decision and Control (CDC), 2010 49th IEEE Conference on*. IEEE, 2010, pp. 2572–2577.

[17] A. Krause and C. Guestrin, "Near-optimal observation selection using submodular functions," in *AAAI*, vol. 7, 2007, pp. 1650–1654.