Single-Element Ultrasound Imaging with Compressed Sensing

William Meng

Department of Electrical Engineering
Stanford University
wlmeng@stanford.edu

Abstract

Compressed sensing offers a way to perform 2D or 3D ultrasound imaging with a single ultrasound transducer. A coded aperture in the form of a pseudo-random delay mask can encode lateral spatial information as temporal variations in the recorded signal. A simplified model for the delay mask is presented in order to enable computationally-efficient simulation of the system. Synthetic data, corresponding to a 2D scene with three point targets, is generated according to a linear image formation model. The inverse problem is then solved with a Least Norm solution to achieve a faithful reconstruction of the scene with a PSNR of 19.99 dB, under the conditions of four mask rotations and an electronic SNR of 90 dB. The degradation of PSNR of the reconstruction with decreasing electronic SNR is also investigated.

Index Terms: single-element ultrasound, compressed sensing, coded aperture, Field II

1. Introduction

Conventionally, a single ultrasound transducer can only measure time-of-flight echos that correspond to depth (z) information, which is known as A-mode imaging (Fig. 1a). In order to measure lateral (x, y) information which is needed for a 2D or 3D imaging, mechanical scanning of a single transducer or an ultrasound array with many elements is required in order to adequately sample the spatial profile of the ultrasound field according to the Nyquist-Shannon sampling theorem.

Compressed sensing allows for reconstruction of a signal from sub-Nyquist sampled measurements by making use of the signal's sparsity in a transformed domain [1]. One such way to enable compressed sensing by transforming a signal in the context of single element ultrasound imaging is to introduce a coded aperture in front of the transducer, which alters the phase and/or amplitude of the transmitted and received wavefronts [2]. In particular, a pseudo-randomly coded delay mask in front of a transducer will produce a highly aberrated and non-uniform wave field which encodes spatial information from the scene as temporal variations in the received signal. The received signal can then be processed by an algorithm to yield a reconstructed image (Fig. 1b).

The goal of this project is to develop and characterize a computationally-efficient approach to simulating a single-element, compressed sensing ultrasound imaging system by utilizing a simplified model for the coded aperture delay mask. Whereas Kruizinga et al. take an empirical approach by treating the coded aperture as a black box which must be measured experimentally [2], this work instead uses a theoretical model for the coded aperture which is as simple as possible. Although this approach sacrifices some physical accuracy, it captures the essential features of the system and elucidates its basic behavior in a computationally-efficient manner.

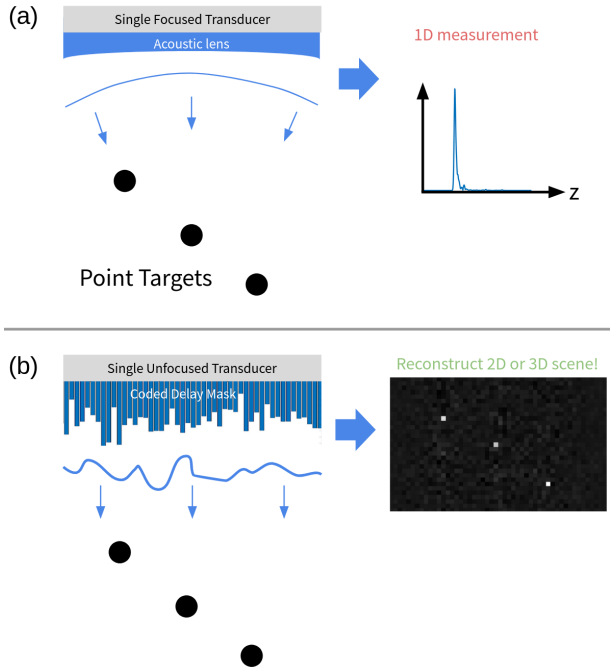

Figure 1: *Comparison of (a) A-mode imaging which can only capture 1D information, and (b) 2D compressed sensing ultrasound imaging.*

2. Related Work

Clinical ultrasound imaging systems typically use ultrasound arrays with many individually-addressable elements, thus allowing for 2D or 3D imaging by a beamforming process. These systems are typically expensive due to the electronic hardware complexity associated with a high channel count, they offer optimal image quality by satisfying the Nyquist sampling rate for the spatial ultrasound field. Although the high quality images produced by this approach are unlikely to be challenged by compressed sensing techniques in clinical settings, compressed sensing may provide tangible benefits in scenarios where hardware complexity is at a premium, such as in space-constrained or low-cost systems [2].

Other approaches to implementing a coded aperture in compressed sensing ultrasound imaging include an amplitude mask with Hadamard-encoding [3] or a scattering layer [4].

3. Methods

3.1. Near-Field Phase Screen Approximation

In this work, we consider a coded aperture in the form of a phase or delay mask, made of a plastic material with a greater speed of sound (c_{plastic} = 2750 m/s) than the surrounding medium (c_{tissue} = 1540 m/s). The width of the transducer is divided into a set of virtual elements. By varying the thickness of the mask in front of each virtual element, a different phase shift or delay will be applied to the wavefront incident at each virtual element. Additionally, the wavefront may undergo partial reflection, refraction, and/or diffraction when crossing the interface between the mask and the surrounding medium. To fully account for these effects in simulation, a wave simulation in heterogeneous media would be required.

In order to reduce the computation required, we make the Near-Field Phase Screen approximation, which assumes that the mask only causes a phase shift or delay on the wavefront, with no amplitude changes or any other complicated wave phenomena [5]. This approximation is reasonable when the mask is near the transducer, there is not too much impedance mismatch between the two materials, and there is not too much variation in phase across each individual virtual element.

As a result of this approximation, the physical delay mask can be modeled as a delay profile which is applied to the virtual elements in the transducer. The simulation can then be performed in a homogeneous medium. During pulse transmission, a shifted version of the transducer's impulse response is generated at each virtual element, which when combined together form the transmitted wave. After this wave propagates through the medium and scatters off of targets in the scene, the received echo signals at each virtual element are shifted according to the delay profile, and are then summed together to form a single received signal (Fig. 2).

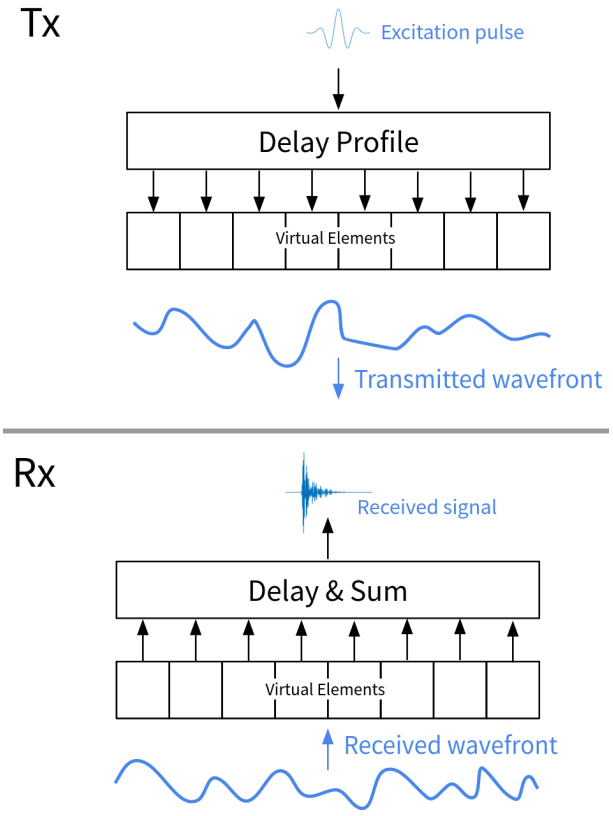

Figure 2: *Diagram indicating how the delay mask is modeled as a delay profile applied to virtual elements in the transducer during pulse transmission (Tx) and receiving echoes (Rx).*

3.2. Mask Rotations

Rotating the delay mask allows unique spatial information to be captured by additional measurements, thus allowing for better image reconstruction, albeit with increased measurement time. Note that this depends on the amount of correlation between the rotated mask measurements [2]. For example, a delay mask which has perfect rotational symmetry with respect to the axis of rotation would see no benefit from making measurements with additional rotations, because each set of rotated measurements would be completely correlated. Furthermore, a small rotation in a mask without rotational symmetry would likely have a high degree of correlation. As another simplification in this project, we consider independent, identically-distributed delay profiles instead of a rotating a physical mask, thus ensuring minimal correlations between sets of rotated measurements.

3.3. Ultrasound Simulation

The pulse-echo response of the system was simulated using the Field II library [6][7] in MATLAB software, release 2020b (The MathWorks). The pulse-echo response is found by first computing the transmitted field by propagating a wave outward from the transducer with the delay profile applied. Then by reciprocity, the pulse-echo response at each pixel is given by the temporal autoconvolution of the transmitted field at that pixel [2]. Note that this approach is only valid for linear propagation and backscattering without any multipath reflections. If these assumptions do not hold, such as in the cases of harmonic imaging, imaging with non-linear contrast-enhancing agents, or imaging in a strongly reverberating medium, then computing the pulse-echo response would require a separate, bi-directional wave simulation for each point in the scene.

3.4. Linear Image Formation Model

Consider a 2D scene to be imaged with N pixels. Let v be the vectorized ground truth image of the scene. Let u be the measurement vector with M samples.

The linear image formation model is given by [2]:

$$u = Hv + n \tag{1}$$

where:

- u is a M-dim column vector
- H is a block matrix with dimensions $M \times N$
- v is a N-dim column vector
- n is a M-dim column vector whose samples represent zero-mean additive white Gaussian noise (AWGN) with a noise variance of σ_e^2. Note that electronic SNR is defined here as $eSNR = \frac{\max(H)}{\sigma_e}$.

This linear image formation model is illustrated graphically in Fig. 3a.

The structure of the block matrix H is as follows:

$$H = \begin{bmatrix} H_1 \\ \vdots \\ H_R \end{bmatrix} \tag{2}$$

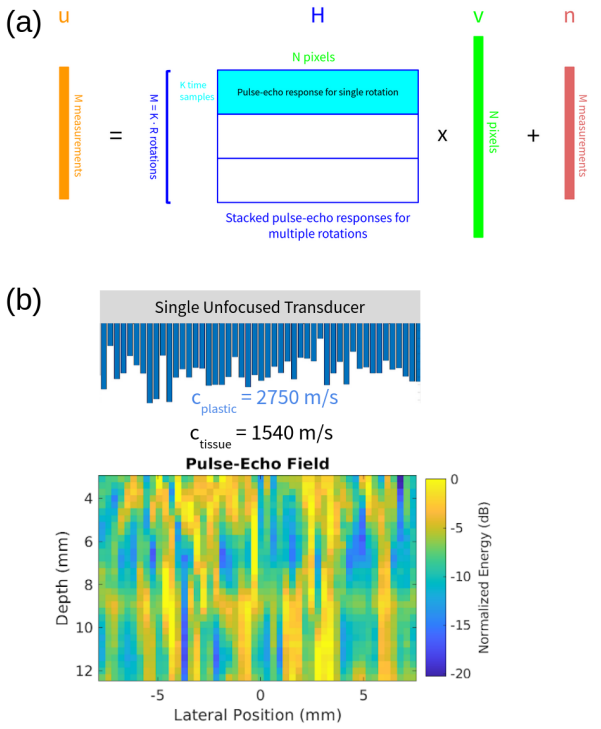

Figure 3: *Diagram illustrating (a) the linear image formation model and the structure of the H matrix, and (b) the pulse-echo field (with colormap indicating energy integrated over time at each pixel) for an example delay mask.*

where:

- $H_r = \begin{bmatrix} h_{r,1} & \dots & h_{r,N} \end{bmatrix}$ for $r \in \{1, \dots, R\}$ is a $K \times N$ matrix
- $h_{r,n}$ is a K-dim column vector representing the pulse-echo response as a time series at each pixel $n \in \{1 \dots N\}$ in the image for mask rotation r.

Therefore, $M = KR$. Each column n of H can be thought of as a set of vertically-stacked vectors corresponding to the pulse-echo response at pixel n as measured with each mask rotation r. The effect of increasing R is to acquire more measurements which encode unique spatial information about the scene, thus improving the degree to which the reconstruction problem is ill or well-posed.

Synthetic data is generated directly from this image formation model in order to avoid issues with spatial and temporal offsets that may arise in simulated or experimental measurements.

3.5. Image Reconstruction

The image reconstruction problem requires inverting a possibly underdetermined system, depending on how many measurements M there are compared to the number of pixels N in the image to reconstruct. If the system is underdetermined ($M < N$), there are infinitely many possible solutions $\hat{v}$ satisfying $u = H\hat{v} + n$. In order to choose just one solution for $\hat{v}$, a criterion for the merit of a particular solution $\hat{v}$ is needed. An optimal solution for $\hat{v}$ can then be found by maximizing the merit function, or alternatively by minimizing the cost function. The problem of finding this maximum or minimum can be formulated as an optimization problem.

In this project, we consider two approaches to finding an optimal solution. The first is Least Norm, which solves the problem:

$$\min_{\hat{v}} ||H\hat{v} - u||_2^2 \tag{3}$$

with the solution:

$$\hat{v} = H^T(HH^T)^{-1}u \tag{4}$$

The matrix inversion is computed by either the Preconditioned Conjugate Gradients (PCG) method or with the Moore-Penrose Pseudoinverse.

The second approach is to impose a prior assumption of sparse gradients on the image and solve a constrained optimization problem with the Alternating Direction Method of Multipliers (ADMM) algorithm using a Total Variation (TV) regularization term [8]. This approach solves the problem:

$$\min_{\hat{v}} \frac{1}{2}||H\hat{v} - u||_2^2 + \lambda||z||_1 \tag{5}$$

- λ is the regularization parameter
- $z = \begin{bmatrix} \nabla_x \\ \nabla_y \end{bmatrix} \hat{v}$ is the anisotropic gradient of the reconstructed image

The ADMM method solves this problem with an iterative approach, which is described by Boyd et al. [8].

4. Analysis

The need to take multiple measurements while rotating a mask in order to achieve a well-conditioned inverse problem limits the practicality of this approach. As such, this approach does not appear to offer not much benefit over a mechanically-scanned transducer for 2D imaging, except in the case of space-constrained or cost-constrained applications. However, for 3D imaging it can drastically cut down the amount of physical movement required, since the compressed sensing mask only needs to be rotated about a single axis, whereas the mechanical scanning transducer would have to be rotated about or translated across two axes. The reduced motion required for acquiring each image means that a series of images can be acquired with a faster frame rate.

Another limitation of the compressed sensing approach to ultrasound imaging is its dependence on sparsity in the image. The more complex the scene is, the more measurements and thus mask rotations are required to yield comparable image quality to traditional methods of ultrasound imaging.

5. Results

A scene with dimensions of 31×50 pixels containing 3 point targets was generated (Fig. 4a). The $N = 2050$ pixel image was vectorized to yield the column-vector v.

The generated delay masks were pseudorandomly sampled from a uniform distribution with a seed number supplied to the random number generator in order to ensure repeatability. A constant offset was added to the thickness of all elements, but this has no effect on the results because only the relative delays between elements matters. The mask thicknesses were then converted into delay profiles according to the speed of sound in the material.

Simulations were performed with no mask, a single rotation ($R = 1$), and multiple rotations ($R = 4$). The case with no mask can be considered as a special case of $R = 1$ with a poorly-designed mask. The number of measurements K in each case varied slightly due to timing offsets in Field II. However, in all cases K was approximately 480. Recall that $M = KR$ whereas N is held constant in all of the simulations performed, so the conditioning of the inverse problem strongly depends on the number of rotations R.

The reconstructed images were scaled to have pixel values ranging from 0 to 1 before computing the Peak Signal-to-Noise Ratio (PSNR) as:

$$PSNR = 10 \log_{10} \left(\frac{\max(v)}{MSE} \right) \tag{6}$$

where:

$$MSE = \frac{1}{N} \sum_{n=1}^{N} (v_n - \hat{v}_n)^2 \tag{7}$$

The effect of different mask configurations was investigated with $eSNR = 90dB$ (Fig. 4b). First, a measurement without any delay mask is considered. Even without an intentionally-designed coded aperture, the pulse-echo field has some spatiotemporal diversity due to the near-field interference pattern of an unfocused transducer, but there is still significant symmetry and therefore correlation in the structure of the pulse echo response. The resulting reconstructed image demonstrates some ability to resolve the 3 point targets. However, there are substantial serious artifacts present, such as double targets, fictitious targets, and a high level of background noise.

Next, a measurement with a single rotation ($R = 1$) of the delay mask is considered. The pulse-echo field is shown to have much less symmetry and more spatial diversity than the case with no mask. The resulting reconstructed image accurately resolves the presence and location of the 3 point targets, without introducing any fictitious targets. However, there is still a high level of background noise, which degrades the image PSNR.

Finally, a measurement with multiple rotations ($R = 4$) of the delay mask is considered. Each rotation introduces a unique and diverse pulse-echo field. Therefore, it is unsurprising that the reconstructed image looks much better than the no mask and single rotation cases. The 3 point targets are clearly resolved with a low level of background noise, thus achieving a high PSNR value. In fact, the pseudoinverse solution looks essentially perfect.

The effect of electronic SNR on reconstruction quality was also investigated (Fig. 4c). The PCG solution degrades gradually as the electronic SNR is decreased, whereas the pseudoinverse solution fails completely for any case that doesn't have extremely high electronic SNR.

ADMM results with $eSNR = 90dB$ and various numbers of mask rotations are shown in (Fig. 4b). The ADMM parameters were 25 iterations, $\lambda = 0.01$, and $\rho = 10$. The results are substantially worse than the Least Norm solutions, both qualitatively and quantitatively.

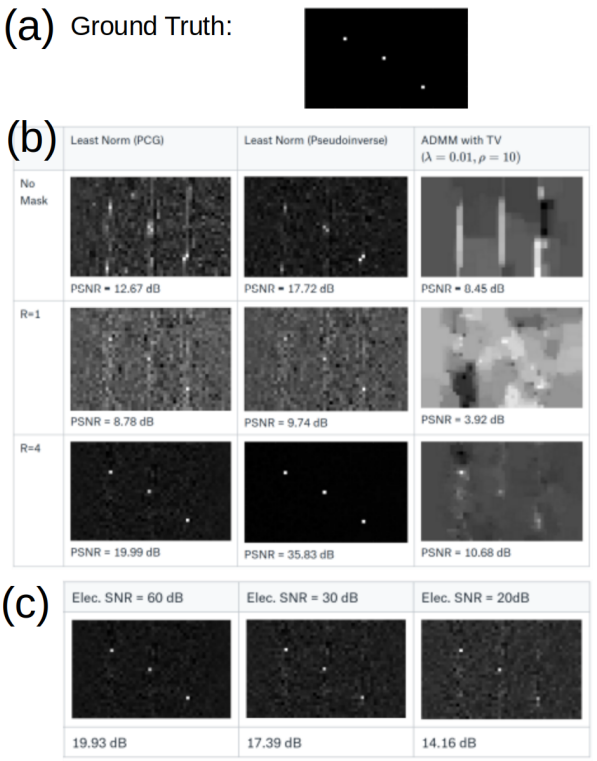

Figure 4: *(a) Ground truth image with 3 point targets (b) Results with various mask configurations (c) Results with varying electronic SNR.*

6. Discussion

Curiously, the reconstructions for a single rotation have a worse PSNR than with no mask. This may be a consequence of the scaling applied to the reconstructed images before computing PSNR. If the reconstructed image has very high outliers in any of its pixels, then the rest of the pixels will be scaled down proportionately. Since the ground truth image is zero everywhere except for the 3 point targets, the presence of a high outlier would cause the background noise to be suppressed almost everywhere after scaling, resulting in a lower MSE and higher PSNR. Therefore, the comparison between the no mask and single rotation cases may be unfair. It should be denoted that the PSNR value is not a complete description of the quality of a reconstructed image for this reason, as well as the fact that PSNR doesn't directly give any information about the presence of artifacts or dimensional accuracy of the reconstructed point targets.

Interestingly, the ADMM reconstruction yielded worse results than the Least Norm approach. This may be due to an issue with the chosen parameters, or perhaps the prior term is not well suited to the content of the image.

In future work, this work can be extended to reconstruct 3D scenes. Furthermore, an analysis of system resolution and a characterization of the dependence of reconstruction quality on image sparsity should be done. Additionally, there may be more effective coded aperture designs than a pseudorandom delay mask. This can be evaluated by measuring the correlation between the columns of H in a given coded aperture design. For practical implementations, the system should be modified to work with real world experimental data or simulated data with offsets and inaccuracies, rather than only the idealized, synthetic data used in this work. Furthermore, the reconstruction algorithm should be implemented with parallelization, so that it may benefit with GPU acceleration. Another extension to the system would be to implement a Kalman filter to dynamically update the reconstructed image using knowledge from both the previously reconstructed image and new measurements, thus

yielding improved results for dynamically changing scenes.

7. Conclusion

In conclusion, a compressed sensing method for 2D ultrasound imaging that uses a pseudorandom delay mask to encode lateral spatial information as temporal variations with a simplified, computationally-efficient simulation approach is presented. The treatment of the coded aperture delay mask as a delay profile applied to the virtual elements in a transducer trades off physical accuracy for computational simplicity. Additional rotations of the mask, which are modeled here as independent delay profiles, improve the conditioning of the inverse problem at the expense of increased measurement time. A scene with 3 point targets was used to generate synthetic data, which was then fed into Least Norm and ADMM reconstruction algorithms. The quality of the reconstructed images are investigated for a varying number of mask rotations R and electronic SNR. Further work can be done to improve the physical accuracy and performance of the simulation, as well as to enhance the overall utility of the system.

8. Acknowledgements

Code for the ultrasound simulation was adapted from RAD 235 workshop examples. Code for PCG and ADMM were adapted from the EE 367 homework.

9. References

[1] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, iss. 4, 2006.

[2] P. Kruizinga, et al, "Compressive 3D ultrasound imaging using a single sensor," *Science Advances*, vol. 3, no. 12, 2017.

[3] E. Hahamovich, A. Rosenthal, "Ultrasound Detection Arrays Via Coded Hadamard Apertures," *IEEE Trasactions on Ultrasounics, Ferrorelectrics, and Frequency Control*, vol. 67, iss. 10, 2020.

[4] X. Luís Deán-Ben, et al, "Acoustic Scattering Mediated Single Detector Optoacoustic Tomography," *Physical Review Letters*, vol. 123, iss. 17, 2019.

[5] Y. Li, "Correction of Phase Aberrations in Medical Ultrasound Images Using Signal Redundancy," *IntechOpen*, 2011.

[6] J.A. Jensen, "Field: A Program for Simulating Ultrasound Systems," *10th Nordic-Baltic Conference on Biomedical Imaging Published in Medical & Biological Engineering & Computing*, pp. 351-353, Volume 34, Supplement 1, Part 1, 1996.

[7] J.A. Jensen and N. B. Svendsen, "Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers," *IEEE Trans. Ultrason., Ferroelec.*, Freq. Contr., 39, pp. 262-267, 1992.

[8] S. Boyd, et al, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," *Foundations and Trends in Machine Learning*, Vol. 3, No. 1, 2010.