Optimal Beamforming in Ultrasound Using the Ideal Observer

Craig K. Abbey, Nghia Q. Nguyen, and Michael F. Insana, *Senior Member, IEEE*

***Abstract*—Beamforming of received pulse-echo data generally involves the compression of signals from multiple channels within an aperture. This compression is irreversible, and therefore allows the possibility that information relevant for performing a diagnostic task is irretrievably lost. The purpose of this study was to evaluate information transfer in beamforming using a previously developed ideal observer model to quantify diagnostic information relevant to performing a task. We describe an elaborated statistical model of image formation for fixed-focus transmission and single-channel reception within a moving aperture, and we use this model on a panel of tasks related to breast sonography to evaluate receive-beamforming approaches that optimize the transfer of information. Under the assumption that acquisition noise is well described as an additive wide-band Gaussian white-noise process, we show that signal compression across receive-aperture channels after a 2-D matched-filtering operation results in no loss of diagnostic information. Across tasks, the matched-filter beamformer results in more information than standard delay-and-sum beamforming in the subsequent radio-frequency signal by a factor of two. We also show that for this matched filter, 68% of the information gain can be attributed to the phase of the matched-filter and 21% can be attributed to the amplitude. A 1-D matched filtering along axial lines shows no advantage over delay-and-sum, suggesting an important role for incorporating correlations across different aperture windows in beamforming. We also show that a post-compression processing before the computation of an envelope is necessary to pass the diagnostic information in the beamformed radio-frequency signal to the final envelope image.**

I. Introduction

Backscattered ultrasound images rely on the process of beamforming to combine responses from an array of acoustic transducers into a single RF signal for a subsequent envelope computation [1], [2]. The implementation of beamforming is challenging on a practical level because the process occurs in real time at video-rate speeds. However, beamforming is also a theoretical challenge because it is not clear how the many received signals should be combined to optimally preserve the diagnostic information they contain. Irreversible steps in ultrasonic signal acquisition and processing leave open the possibility that information may be lost at various points on the way to producing a final displayed image. This latter problem, finding optimal transforms for processing the received echo data, is the subject of this work.

In principle, beamforming encompasses the entire process of signal transmission, echo reception, and subsequent processing steps up to the computation of an envelope image [3]. Here, we will concentrate on the latter stage of this process under the assumption of a linear array with a single, fixed-focus transmit beam that is received over the same aperture at each of the individual receive channels. Hence, for the purposes of this study, beamforming refers to the processing of received signals into a format appropriate for envelope computation. As such, beamforming can be thought of as a transformation of the received data that inherently employs an irreversible compression step in which the many channels of a receive aperture are combined into a single scan-line signal. The standard approach, commonly referred to as delay-and-sum (DS) beamforming [2], [4], accomplishes this by summation across channels after delay correction for the path-length to the detector. We consider a more general class of beamforming transformations that consist of 1) any pre-compression processing (such as delay correction); 2) the compression step consisting of a sum across receive channels; and 3) any post-compression processing of the signal applied before envelope detection.

In previous publications [5], [6], we have developed a statistical model of ultrasonic image formation that is based on the use of a Gaussian stochastic process to represent acoustic reflectivity in the object. A linear transformation of the object reflectivity, based on the pulse pressure profile and subsequent additive acquisition noise, models pulse-echo image formation [6]. We restrict attention to the focal region of a fixed-focus pulse, which allows us to approximate the system as a noisy convolution with a pulse-echo point-spread function (PSF). This model of RF acquisition is combined with a panel of simple visual tasks related to breast cancer detection and diagnosis to evaluate and optimize signal processing. We approach this optimization by considering an optimal discriminant function, the (log) likelihood ratio, from signal detection theory [7]. In imaging tasks, this optimal discriminant is referred to as the ideal observer because it performs the task and achieves the best possible performance by several reasonable measures [8], [9].

Here, we generalize our model of image formation to incorporate the formation of individual signals in various

Manuscript received January 3, 2010; accepted April 26, 2010. This work was supported in part by the U.S. National Institutes of Health Grant R01- CA118294.

C. K. Abbey is with the Department of Psychology, University of California, Santa Barbara, Santa Barbara, CA, and the Department of Biomedical Engineering, University of California, Davis, Davis, CA (e-mail: abbey@Psych.UCSB.edu).

N. Q. Nguyen and M. F. Insana are with the Departments of Bioengineering and Electrical and Computer Engineering, and Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL.

Digital Object Identifier 10.1109/TUFFC.2010.1616

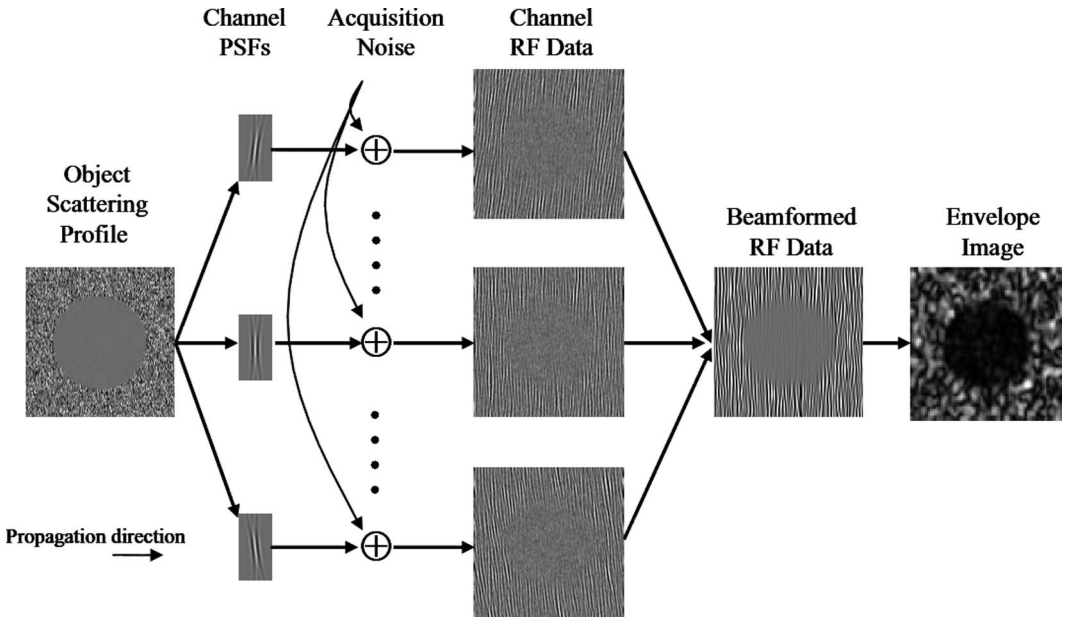


Fig. 1. Model of image formation. The diagram shows the model of image formation from the scattering object to the envelope image. Three element channels are shown from a total of 96 with ellipses (·) indicating the remaining elements. Channel point spread functions (PSFs) interact with the object scattering function with additive acquisition noise to produce element RF data shown in scan-converted format as images. These are combined into a single RF image through the process of beamforming, and then transformed to a final image via Hilbert transform and envelope detection.

receiver channels. This allows us to consider the effect of signal processing in beamforming on task performance and to analyze the ideal observer for processing strategies that achieve optimal performance. The result is a block-matrix system that takes into account the axial sampling, lateral position of the aperture, and position of the channel within the aperture. One direct consequence of this analysis is the derivation of an optimal approach to compressing received signals in a moving aperture system. We show that the use of a 2-D matched filter in each channel before summation across channels is sufficient for computing the ideal observer test statistic, and therefore retains all the diagnostic information of the received data.

To evaluate the effect of signal processing in beamforming, we compare the 2-D matched filter (2DMF) beamformer to standard delay-and-sum beamforming. We also consider the role that phase and amplitude play in transferring information in the matched filter beamformer, and this leads us to consider a phase-only matched filter (PMF) and an amplitude-only matched filter (AMF). We also consider the role of axial processing of scan data compared with processing across different apertures by implementing a 1-D matched filter (1DMF). We evaluate performance of these different beamforming approaches in five simple visual tasks related to breast cancer imaging using a system model based on Field II simulation code for acoustic pulse propagation in an incoherent scattering medium [10], [11]. This analysis of information transfer relies heavily on the notion of statistical efficiency in each task as a measure of how much diagnostic information is lost from the use of suboptimal processing.

II. Theory

In this section, we expand on the ideal observer framework for ultrasonic signal processing described previously. We consider the compression of RF signals from individual receive channels that are summed over a moving aperture window, a fundamental component of linear-array beamforming.

A. System Model for Receive Channel Signals

Fig. 1 is a graphical diagram of the beamforming process as used in this work. We consider a single, fixed-focus transmit beam that is received over the same aperture at each of the individual channels. This process can be modeled as an array of PSFs that interact with a scattering object to produce the individual receive channel signals (shown in scan converted format). Different PSFs are needed to account for different relative positions of the channels in the receive aperture. Beamforming consists of combining the information from these individual channels into a single beamformed data set (also shown scan converted) for subsequent envelope detection.

In previous work [5], [12], we have modeled the formation of RF data as a linear system, $\mathbf{H}$, acting on a vector, $\mathbf{f}$, representing the scattering profile of the object. In principle, $\mathbf{f}$ is a continuously defined function of 2 or (more appropriately) 3 dimensions. However, because the system that acts on it is inherently band limited, it is well approximated for our purposes by a discrete set of sample points. We restrict attention to the focal zone of the transmitted pulse, where the linear system can be efficiently imple-

mented as a convolution with the pulse profile, but we note that the approach could be extended to shift-varying linear system models. The resulting product of the system matrix with the object vector is corrupted by noise, $\mathbf{n}$, to form the RF data vector, $\mathbf{g}$.

We conceive of a similar process for generating the signal for each receive channel in an aperture as shown in Fig. 1. Let a be the index of a channel within the received aperture ($a = 1, \ldots, A$), then a linear system model for the received signal of this channel is given by

$$\mathbf{g}_a = \mathbf{H}_a\mathbf{f} + \mathbf{n}_a. \tag{1}$$

Note that in this work, boldface variables represent matrix (uppercase) and vector (lowercase) quantities. We see from (1) that each receive channel has its own system matrix, $\mathbf{H}_a$, and its own (independent) vector of acquisition noise, $\mathbf{n}_a$, but the object is common to all channels.

We can assemble the linear systems specified in (1) into a single receiver system equation describing all channels simultaneously by using block-matrix equations:

$$\mathbf{g}_{\mathrm{R}} = \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \\ \vdots \\ \mathbf{g}_A \end{bmatrix} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \vdots \\ \mathbf{H}_A \end{bmatrix} \mathbf{f} + \begin{bmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \\ \vdots \\ \mathbf{n}_A \end{bmatrix} = \mathbf{H}\mathbf{f} + \mathbf{n}. \tag{2}$$

It is important for practical reasons to be aware of the size of the system, because this will limit what sorts of operations can reasonably be used in later steps. The number of vector elements in each $\mathbf{g}_a$ is the number of scan lines times the number of samples per scan line, and therefore the number of elements in $\mathbf{g}_{\mathrm{R}}$ is that number times the number of channels. For the system we use in the simulation experiments, there are 128 scan lines, 256 samples per scan line (concentrated in the focal zone), and 96 receive aperture channels for a total vector dimension of 3 145 728 elements in $\mathbf{g}_{\mathrm{R}}$. Because of the inherent band limit of the system, we assume that the object can be adequately represented by the sampling rate of the system, and therefore $\mathbf{f}$ has a total vector dimension of 32 768. Thus, the block system matrix, $\mathbf{H}$, is highly rectangular with almost 2 orders of magnitude more rows than columns.

For the purpose of evaluating different beamforming approaches, we utilize a simplified shift-invariant system model intended to approximate the workings of a focused pulse echo system. The model uses a convolution with a system point spread function (PSF) to implement multiplication by the system matrix or receive channel submatrices. This allows for rapid computations of products involving the system matrix, and a compact representation of the system by the PSF that makes for modest memory requirements.

B. Beamforming in the System Model

One role of beamforming is to compress the various channel signals into a single RF signal, $\mathbf{g}_{\mathrm{BF}}$. If we consider a linear signal compression across the channels in (2), we can write this as another linear transform

$$\mathbf{g}_{\mathrm{BF}} = \mathbf{B}\mathbf{g}_{\mathrm{R}}, \tag{3}$$

where $\mathbf{B}$ is another rectangular matrix that reduces the dimension of the result to the number in each of the individual channels. For now this compression transform will remain somewhat generic. We consider specific approaches to defining $\mathbf{B}$ in Section II-D.

Because the beamformed RF signal is defined at the same sampling rates as the channel signals, the matrix $\mathbf{B}$ can be thought of as partitioned into square blocks,

$$\mathbf{B} = [\mathbf{B}_1 \mid \mathbf{B}_2 \mid \cdots \mid \mathbf{B}_A]. \tag{4}$$

Eqs. (2)–(4) allow the beamformed data to be written as a noisy linear system,

$$\mathbf{g}_{\mathrm{BF}} = \mathbf{H}_{\mathrm{BF}}\mathbf{f} + \mathbf{n}_{\mathrm{BF}}, \tag{5}$$

where the beamformed system and noise are defined by

$$\mathbf{H}_{\mathrm{BF}} = \sum_{a=1}^{A} \mathbf{B}_a\mathbf{H}_a, \text{ and } \mathbf{n}_{\mathrm{BF}} = \sum_{a=1}^{A} \mathbf{B}_a\mathbf{n}_a, \tag{6}$$

respectively. Fig. 2 shows the equivalent one-step model of image formation whereby the beamformed RF data are generated directly, and then transformed into a final envelope image. Note that for a given pulse acquisition system, the beamformed system matrix and noise will change based on the choice of $\mathbf{B}$.

C. Beamforming and the Ideal Observer

The ideal observer is a well-established component of signal detection theory [7], [13]. It constructs a decision variable that achieves the highest possible level of task performance on whatever data it accesses. As such, it serves a benchmark for identifying where information is lost. For example, consider the transformation from the receiver signals, $\mathbf{g}_{\mathrm{R}}$, to the beamformed RF, $\mathbf{g}_{\mathrm{BF}}$. This transformation is irreversible in that $\mathbf{g}_{\mathrm{R}}$ cannot be recovered from $\mathbf{g}_{\mathrm{BF}}$ after compression. If the ideal observer acting on $\mathbf{g}_{\mathrm{BF}}$ has a performance substantially less than that of the ideal observer acting on $\mathbf{g}_{\mathrm{R}}$, then signal compression in the beamformer is losing information relevant to task performance. Note that if the transformation is invertible (i.e., if $\mathbf{g}_{\mathrm{R}}$ is recoverable from $\mathbf{g}_{\mathrm{BF}}$), then any decision variable that can be formulated in $\mathbf{g}_{\mathrm{R}}$ can also be formulated on $\mathbf{g}_{\mathrm{BF}}$, and hence no information is lost.

1) The Ideal Observer Test Statistic: For simple detection and discrimination tasks, the ideal observer test statistic is given by a likelihood ratio. Let c represent one of two image classes, where $c = 1$ consists of data from patients with malignancies, and $c = 0$ consists of data from

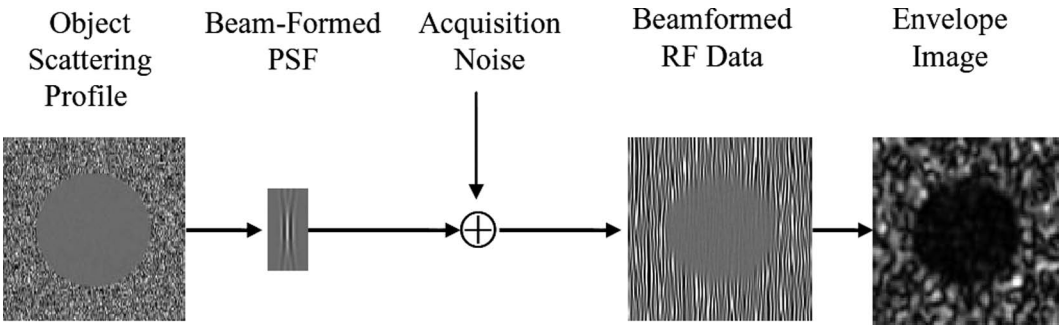


Fig. 2. An equivalent image formation model for linear beamformers. Here a single beamformed PSF interacts with the scattering object and produces beamformed RF data after a composite acquisition noise term is added. The subsequent transformation to an envelope image is identical.

normal patients who have either no disease or benign processes (cysts, fibroadenoma, etc.). We will refer to the latter class generically as benign. The two image classes can be used to define a conditional probability at any stage of the image formation process including the received signals, $p(\mathbf{g}_{\mathrm{R}} \mid c)$, the beamformed RF signal, $p(\mathbf{g}_{\mathrm{BF}} \mid c)$, and subsequent envelope images (with some issues described subsequently). The ideal observer test statistic is simply the ratio of the conditional probabilities

$$\Lambda(\mathbf{g}) = \frac{p(\mathbf{g} \mid c=1)}{p(\mathbf{g} \mid c=0)}. \tag{7}$$

This equation uses a generic data set, $\mathbf{g}$, indicating that it can be applied at any stage of image formation. However, because the conditional distributions of the data will be different at different stages, the ideal observer changes at each stage.

2) The 2-D Matched-Filter Beamformer: We will now turn to a specific choice for the beamforming matrix $\mathbf{B}$ in (3). The 2DMF beamformer is defined by performing a 2-D matched filter (across sample points and aperture positions) on the data from each channel before compression to a single signal. This is implemented mathematically by selecting the beamforming matrix to be $\mathbf{B} = \mathbf{H}^{\mathrm{T}}$, where the superscript T indicates the transpose of the matrix. Note that by (4) this is equivalent to a matched filter in each channel ($\mathbf{B}_a = \mathbf{H}_a^{\mathrm{T}}$) that is then summed across channels. Thus, the beamformer uses information from multiple aperture windows to filter a given channel signal before summation. For a highly rectangular $\mathbf{H}$, it is clear that multiplication by the transpose will involve a considerable reduction in dimension. However, we will demonstrate that this choice of beamformer involves no loss of diagnostic information from the ideal observer perspective, even though the transformation is not invertible.

The main assumption needed for our result is that acquisition noise be an independent and identically-distributed Gaussian random field. We make no assumptions on the underlying class distributions of the object. For receiver signals, the conditional probabilities in (7) can be written with explicit reference to the distribution of object in each class as

$$p(\mathbf{g}_{\mathrm{R}} \mid c) = \int \mathrm{d}\mathbf{f}\, p_{\mathrm{R}}(\mathbf{g}_{\mathrm{R}} \mid \mathbf{f}) p_{\mathrm{obj}}(\mathbf{f} \mid c), \tag{8}$$

where $\int \mathrm{d}\mathbf{f}$ indicates a multidimensional integral over all dimensions of the vector $\mathbf{f}$ in the integrand that follows. Under the assumption of white Gaussian acquisition noise in (2), with variance σ_n^2, the conditional probability of the data given $\mathbf{f}$ is multivariate Gaussian,

$$p_{\mathrm{R}}(\mathbf{g}_{\mathrm{R}} \mid \mathbf{f}) = K e^{-(1/2\sigma_n^2)(\mathbf{g}_{\mathrm{R}} - \mathbf{H}\mathbf{f})^{\mathrm{T}}(\mathbf{g}_{\mathrm{R}} - \mathbf{H}\mathbf{f})}, \tag{9}$$

where K is the normalizing constant of the distribution.

Substituting (8) and (9) into (7) allows us to write the ideal observer test statistic (for the receiver data) as

$$\Lambda_{\mathrm{R}}(\mathbf{g}_{\mathrm{R}}) = \frac{\int \mathrm{d}\mathbf{f} K e^{-(1/2\sigma_n^2)(\mathbf{g}_{\mathrm{R}} - \mathbf{H}\mathbf{f})^{\mathrm{T}}(\mathbf{g}_{\mathrm{R}} - \mathbf{H}\mathbf{f})} p_{\mathrm{obj}}(\mathbf{f} \mid 1)}{\int \mathrm{d}\mathbf{f} K e^{-(1/2\sigma_n^2)(\mathbf{g}_{\mathrm{R}} - \mathbf{H}\mathbf{f})^{\mathrm{T}}(\mathbf{g}_{\mathrm{R}} - \mathbf{H}\mathbf{f})} p_{\mathrm{obj}}(\mathbf{f} \mid 0)}. \tag{10}$$

Cancelling common terms in the numerator and denominator yields

$$\Lambda_{\mathrm{R}}(\mathbf{g}_{\mathrm{R}}) = \frac{\int \mathrm{d}\mathbf{f} e^{(1/\sigma_n^2)(\mathbf{f}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{g}_{\mathrm{R}} - (1/2)\mathbf{f}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{f})} p_{\mathrm{obj}}(\mathbf{f} \mid 1)}{\int \mathrm{d}\mathbf{f} e^{(1/\sigma_n^2)(\mathbf{f}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{g}_{\mathrm{R}} - (1/2)\mathbf{f}^{\mathrm{T}}\mathbf{H}^{\mathrm{T}}\mathbf{H}\mathbf{f})} p_{\mathrm{obj}}(\mathbf{f} \mid 0)}. \tag{11}$$

The main point of this derivation is noting that the receiver data, $\mathbf{g}_{\mathrm{R}}$, only enters the right side of the equation in the product $\mathbf{H}^{\mathrm{T}}\mathbf{g}_{\mathrm{R}}$. Therefore, the ideal observer test statistic on the receiver data can be computed with access just to data compressed using the 2DMF beamformer. The result of 2DMF beamforming is sufficient for computing the ideal observer acting on the entire receiver data set and, ignoring practical issues of implementation, involves no loss of diagnostic information. We note that this result can be easily extended to correlated Gaussian acquisition noise, which requires pre-whitening $\mathbf{g}_{\mathrm{R}}$ by multiplying with the inverse of the noise covariance matrix before multiplying by $\mathbf{H}^{\mathrm{T}}$.

Eq. (11) thus provides a rationale for the compression step inherent to beamforming, based on the idea that receive-channel data can be combined in a task-independent way that nonetheless does not lose any diagnostic task information. We can view other beamforming approaches,

such as standard delay-and-sum, as an approximation to this optimal processing.

D. Approaches to Beamforming

So far, we have considered the acquired data to be a generic column vector. For specifying the transformation used in beamforming, it is more convenient to work with the data when it is formatted as a scan-converted 2-D array. The column vector $\mathbf{g}_a$ from (1) is equivalent to the 2-D array $g_a[i_x, i_y]$, where i_x represents axial samples on the scan line, and i_y represents the lateral aperture position on the face of the linear array. We will primarily be concerned with the 2-D discrete Fourier transform of these arrays, denoted by a caret (^), and the spatial frequency indices k_x and k_y (i.e., $\hat{g}_a[k_x, k_y]$ is the 2-D Fourier transform of $g_a[i_x, i_y]$). We will also consider beamformers in which the individual blocks of the matrix $\mathbf{B}$ in (4) can be implemented by convolutions. Let $b_a[i_x, i_y]$ be the PSF (or kernel) of the convolution, then the beamformed data can be written in the Fourier domain as

$$\hat{g}_{\mathrm{BF}}[k_x, k_y] = \sum_{a=1}^{A} \hat{b}_a[k_x, k_y]\hat{g}_a[k_x, k_y]. \tag{12}$$

We will specify different beamformers by the choice of the channel transfer function $\hat{b}_a[k_x, k_y]$.

In this context, the standard delay-and-sum (DS) beamformer can be thought of as a simple phase shift in the scan-line axis to implement delay correction for the channel. Let ϕ_a be the phase shift corresponding to the delay correction needed for channel a, which is often assumed to be a quadratic function of channel position within the aperture window. The delay-and-sum former is then implemented by

$$\hat{b}_{\mathrm{DS},a}[k_x, k_y] = e^{-j\phi_a k_x}, \tag{13}$$

where j is the imaginary unit. Eq. (13) shows that the sum across channels in beamforming is weighted by a phase factor in the Fourier domain that accounts for delays in acquisition timing.

For the 2DMF beamformer, we must consider the channel system matrix in (1), which we have assumed can be implemented as a 2-D convolution. Let $\hat{h}_a[k_x, k_y]$ be the transfer function of this convolution. The 2DMF beamformer is than given by

$$\hat{b}_{\mathrm{2DMF},a}[k_x, k_y] = \overline{\hat{h}_a[k_x, k_y]}, \tag{14}$$

where the overline indicates the complex conjugate. The matched filter can be applied directly to the acquired signals if the matched filter includes the channel delays. Alternatively, delay-corrected channel signals can be used with delay-corrected matched filters. This is true of the various filters described subsequently as well.

The 2DMF beamformer in (14) utilizes both amplitude and phase information from the system transfer function. To isolate the role of phase information in the matched filter beamformer, we implement a (2-D) phase-matched filter (PMF) in which each term is a frequency dependent phase factor tuned to the system transfer function. The functional form of this beamformer is

$$\hat{b}_{\mathrm{PMF,a}}[k_x, k_y] = \frac{\overline{\hat{h}_a[k_x, k_y]}}{\left|\hat{h}_a[k_x, k_y]\right|}, \tag{15}$$

where the absolute value brackets indicate the magnitude of the complex argument. The three beamforming approaches described in (13)–(15) are compared in the simulation studies. To instead isolate the amplitude component of the matched filter independent of the phase, we consider an amplitude-matched filter (AMF) defined by

$$\hat{b}_{\mathrm{AMF,a}}[k_x, k_y] = \left|\hat{h}_a[k_x, k_y]\right|. \tag{16}$$

Note that the product of the PMF and AMF beamformers is the matched filter beamformer in (14).

With the exception of delay-and-sum beamforming, all of the approaches we consider require axial and lateral processing across aperture positions. However, lateral processing is practically difficult because of more extensive memory requirements. Axial-only processing along each scan line is considerably simpler and easier to implement. We evaluate the effect of matched filtering without lateral processing using a 1DMF defined by

$$\hat{b}_{\mathrm{1DMF},a}[k_x, k_y] = \overline{\hat{h}_a[k_x, 0]}, \tag{17}$$

where fixing the y-frequency constrains the spatial convolution kernel to a single scan line.

E. System Model for Receive Channel Signals

1) Modeling Beamformed RF Data: We evaluate the different beamformers on a panel of simple tasks described in Section III-B. Each task is defined by the conditional densities of the two object classes, $p_{\mathrm{obj}}(\mathbf{f} \mid c)$. For the purpose of the simulation studies, the object density function is parameterized by the variance profile of the object in the benign and malignant classes. The object vector is defined as a multivariate Gaussian random variable with each sample point being independent from all others, and with variance specified by the variance profile of the class. We use zero-mean random fields for the object reflectivity, because we are not considering specular reflections. Thus the statistical differences between classes are entirely contained in the variance of the object. Let us denote objects from the benign class by $\mathbf{f}^-$ and those from the malignant class by $\mathbf{f}^+$. We can then describe the distributions for each class as

$$\begin{aligned} \text{Benign: } \mathbf{f}^- &\sim \text{MVN}(\mathbf{0}, \sigma^2_{\text{obj}}(\mathbf{I} + \mathbf{S}_0)) \\ \text{Malignant: } \mathbf{f}^+ &\sim \text{MVN}(\mathbf{0}, \sigma^2_{\text{obj}}(\mathbf{I} + \mathbf{S}_1)), \end{aligned} \tag{18}$$

where σ^2_{obj} is the generic variance of the object, and the diagonal matrices $\mathbf{S}_0$ and $\mathbf{S}_1$ represent the task- and class-dependent deviation from white noise in each class.

The distribution of the object propagates through the system matrix as indicated in (2), and then the beamforming transform in (3) along with acquisition noise. The result is a multivariate Gaussian distribution with correlations across the components of $\mathbf{g}_{\text{BF}}$,

$$p(\mathbf{g}_{\text{BF}} \mid c) = \text{MVN}(\mathbf{0} \mid \mathbf{\Sigma}_{\text{BF},c}), \tag{19}$$

where $\mathbf{\Sigma}_{\text{BF},c}$ is a multivariate covariance matrix describing the second-order moments of $\mathbf{g}_{\text{BF}}$ under each hypothesis. These covariance matrices are related to the task, system model, and beamformer by propagation of variance,

$$\mathbf{\Sigma}_{\text{BF},c} = \sigma^2_{\text{obj}}\mathbf{B}\mathbf{H}\big(\mathbf{I} + \mathbf{S}_c\big)\mathbf{H}^{\text{T}}\mathbf{B}^{\text{T}} + \sigma^2_n\mathbf{B}\mathbf{B}^{\text{T}}. \tag{20}$$

2) Evaluating the Ideal Observer in Simulation Studies: In previous work [5], we have shown that for a general Gaussian model of this sort, the ideal observer can be implemented by the test statistic,

$$\lambda_{\text{BF}}(\mathbf{g}_{\text{BF}}) = -\frac{1}{2}\mathbf{g}^{\text{T}}_{\text{BF}}(\mathbf{\Sigma}^{-1}_{\text{BF},1} - \mathbf{\Sigma}^{-1}_{\text{BF},0})\mathbf{g}_{\text{BF}}. \tag{21}$$

We have also developed a power-series approach for iteratively computing the necessary matrix inverses in (19). Thus, given a realization of beamformed RF data, we can compute a realization of the ideal observer decision variable.

3) Estimating Figures of Merit: By considering many realizations of the decision variable we can estimate figures of merit for task performance. For a given task, we generate a sample of RF data in each of the two classes and then use (19) to transform these into realizations of the ideal observer decision variable. The measure of performance used in the simulation studies is statistical efficiency with respect to the ideal observer. This metric is based on the proportion of correct responses in a forced-choice experiment, P_C, which is the probability that a decision variable randomly drawn from the malignant class is greater than an independently drawn decision variable from the benign class [8]. P_C is mathematically equivalent to the area under a receiver operating characteristic (ROC) curve.

Let $\lambda^m_{\text{BF},c}$ be samples of the decision variable ($m = 1, \ldots, M$) in each class ($c = 0, 1$). The proportion correct is estimated by the average

$$\hat{P}_C = \frac{1}{M^2}\sum_{m=1}^{M}\sum_{m'=1}^{M}\text{step}(\lambda^m_{\text{BF},1} - \lambda^{m'}_{\text{BF},0}), \tag{22}$$

where the step function is zero for negative arguments and one for positive arguments. In principle, the step function attains a value of 1/2 when the argument is zero, but this should never occur for continuously distributed variables. Standard errors on this estimate are readily obtained by resampling the decision variables.

As defined here, P_C applies to fixed choices of the task-dependent variance maps, and hence to a fixed level of task difficulty. To compute the statistical efficiency of a beamformer, we seek to find the level of task difficulty that produced equivalent performance. Each of the tasks used in the study focuses on a feature of interest in breast sonography (lesion contrast, eccentricity, spiculation, etc.). To put these different features on a common scale, we define an object contrast [5], [12] as the absolute value of the difference in normalized variance maps,

$$C = \sum_i \big|[\mathbf{S}_1 - \mathbf{S}_0]_{ii}\big|\Delta x \Delta y, \tag{23}$$

where Δx and Δy are the axial and lateral sampling intervals. The threshold contrast is the object contrast needed for an observer to achieve a targeted level of performance, 80% correct in this work.

Our simulation studies follow the approach used in prior studies by evaluating each task at 4 to 5 different contrasts, and using these evaluations to build a lookup table, from which threshold contrast can be interpolated. The efficiency η is defined from the ratio of threshold contrasts squared. The numerator of the ratio is given by the 2DMF beamformer, reflecting the fact that it retains all of the task relevant information in the receiver data before beamforming. The threshold contrast in the denominator, C_{BF}, is the threshold contrast of the beamformer of interest,

$$\eta = \left(\frac{C_{\text{MF}}}{C_{\text{BF}}}\right)^2. \tag{24}$$

4) Analysis of Envelope Images: The final step of the acquisition and processing diagram in Fig. 1 is computation of an envelope image [7]. In this step, variability in the RF signal is converted into a direct measure of signal magnitude. Computation of the envelope is irreversible [5], and therefore has the possibility of losing information in the RF signal. Thus, it is important that we include an assessment of the effect of an envelope computation, even though this is not part of the beamforming process as we have defined it. In the end, the goal is to retain as much task-relevant information as possible, and to format this information so that it will be passed efficiently to the observer via the envelope image.

We will consider two approaches to computing the envelope image; both based on the Hilbert transform [14]. The standard B-Mode envelope consists of adding a Hilbert transform to the imaginary component along each scan line to find the analytic signal, followed by the mag-

nitude of the complex result. The analytic signal, denoted $g_{\mathrm{AS}}[i_x, i_y]$, can be implemented in the frequency domain by a filtering operation on the beamformed RF data,

$$\hat{g}_{\mathrm{AS}}[k_x, k_y] = \hat{g}_{\mathrm{RF}}[k_x, k_y]\hat{T}_{\mathrm{AS}}[k_x], \tag{25}$$

where the filter transfer function is dependent only on the axial frequency. For an even number of axial samples, N_x, the form of the filter is

$$\hat{T}_{\mathrm{AS}}[k_x] = \begin{cases} 0.5 & k_x = 0, N_x/2 \\ 1 & 1 \le k_x \le N_x/2 - 1 \\ 0 & \text{otherwise,} \end{cases} \tag{26}$$

which selects the positive frequency indices. After the Hilbert transform, the envelope image, $e_{\mathrm{BM}}[i_x, i_y]$, is computed by a complex magnitude at each sample point

$$e_{\mathrm{BM}}[i_x, i_y] = |g_{\mathrm{AS}}[i_x, i_y]|. \tag{27}$$

The complex magnitude is what makes the envelope computation irreversible. Note that practical implementations of envelope computations may also include downsampling. For simplicity, we retain the RF sampling rate for all envelope images.

A second approach to envelope computation involves filtering the RF data with a Weiner filter before the Hilbert transform. Previous results have shown that this can reduce the loss of information in the envelope image. The Wiener-filtered RF data are computed by considering the beamformed system in (6) and (12). The Wiener filter requires knowing the beamformed system transfer function and noise power spectrum. The resulting Wiener filter is given by

$$\hat{w}[k_x, k_y] = \frac{\sum_a \hat{b}_a[k_x, k_y]\hat{h}_a[k_x, k_y]}{\sigma_{\mathrm{obj}}^2 \sum_e \left|\hat{b}_a[k_x, k_y]\hat{h}_a[k_x, k_y]\right|^2 + \sigma_n^2 \sum_a \left|\hat{b}_a[k_x, k_y]\right|^2}. \tag{28}$$

Note that this filter assumes that the object consists of bandpass white noise with variance σ_{obj}^2. The Weiner filter can be applied along with the Hilbert transform,

$$\hat{g}_{\text{AS-WF}}[k_x, k_y] = \overline{\hat{w}}[k_x, k_y]\hat{g}_{\mathrm{RF}}[k_x, k_y]\hat{T}_{\mathrm{AS}}[k_x]. \tag{29}$$

The resulting envelope image is computed as in (25).

To evaluate the envelope images, we rely on an approximation also used in previous work. Smith and Wagner [15] have shown that for low-contrast detection, the ideal observer acting on envelope images effectively integrates the envelope squared over the lesion area. We use this general approach as an approximation to ideal observer performance on our panel of detection and discrimination tasks although there is some evidence that these do not necessarily achieve optimal performance [12]. The Smith-Wagner test statistic is given by the sum

$$\lambda_{\mathrm{SW}} = -\frac{1}{2}\sum_{i_x, i_y} (S_1[i_x, i_y] - S_0[i_x, i_y])e[i_x, i_y]^2, \tag{30}$$

where S_1 and S_0 are the task variance maps arranged as a 2-D array. The test statistic is used in (20) and (22) to get measures of task performance and efficiency. However, in this case the numerator of the efficiency computation is still the ideal observer acting on beamformed RF data. Thus, efficiency is still relative to the information contained in the received signals.

In addition to the efficiency measure given in (24), we are also interested in the effect of the envelope computation on information content for each beamforming approach. To isolate this component, we consider a relative efficiency, in which the numerator is the object contrast threshold for the beamformed RF, and the denominator is the object contrast threshold for the same beamforming system derived from envelope images using the Smith-Wagner test statistic,

$$\eta_{\mathrm{rel}} = \left(\frac{C_{\mathrm{BF}}}{C_{\mathrm{SW}}}\right)^2. \tag{31}$$

We consider both the standard B-Mode envelope image as well as Weiner-filtered envelope images.

III. Methods

In this section we describe the methods used to evaluate ideal observer performance on a battery of simple tasks related to breast sonography.

A. System Model Using Field II

We have used the Field II code [10], [11] to simulate the 2-D pulse-echo impulse response from a standard linear array transducer. We assumed a 44.3-mm, physical, in-plan aperture with 224 elements, each element 0.19 × 1.0 mm in size. The element pitch and kerf were 0.2 mm and 0.01 mm, respectively. We transmitted two-cycles of a 5.0-MHz sinusoidal voltage from a 19-mm (96 element) active transmit aperture that was circularly focused with a 40-mm electronic radius of curvature. We then received echoes individually on each channel in the same active aperture. Images consisted of 128 scan lines formed by indexing the 96-channel active aperture along the 224-element array. To minimize weak artifacts generated during pulse simulation, echoes were created at 320 Msamples/s, low-pass filtered (Hamming), and downsampled to 15 Msamples/s before simulating the echo fields in $\mathbf{g}_{\mathrm{R}}$. For simulating properties of the Siemens Sono-line Antares system with a VF10–5 linear array, the echo SNR of individual channels averaged 11.2 dB with less than 1% variability. After DS beamforming, the echo SNR increased to 29.4 dB, which is consistent with system measurements made from phantoms (ATS Laboratories, Bridgeport, CT). We have

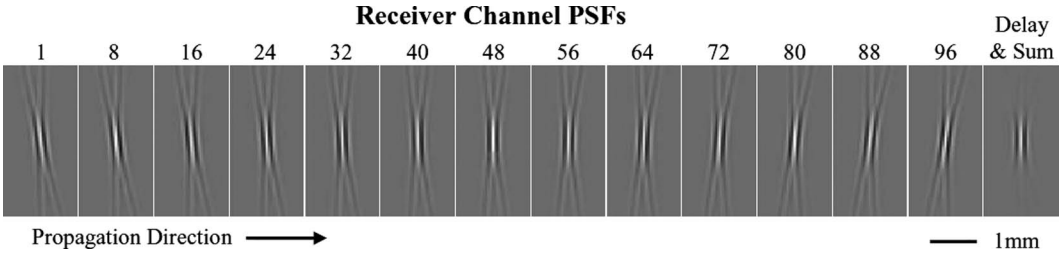


Fig. 3. Receive element pulse profiles. The panel shows various receive channel pulse PSFs. The pulse profile rotates because of the relative position of each channel in the aperture.

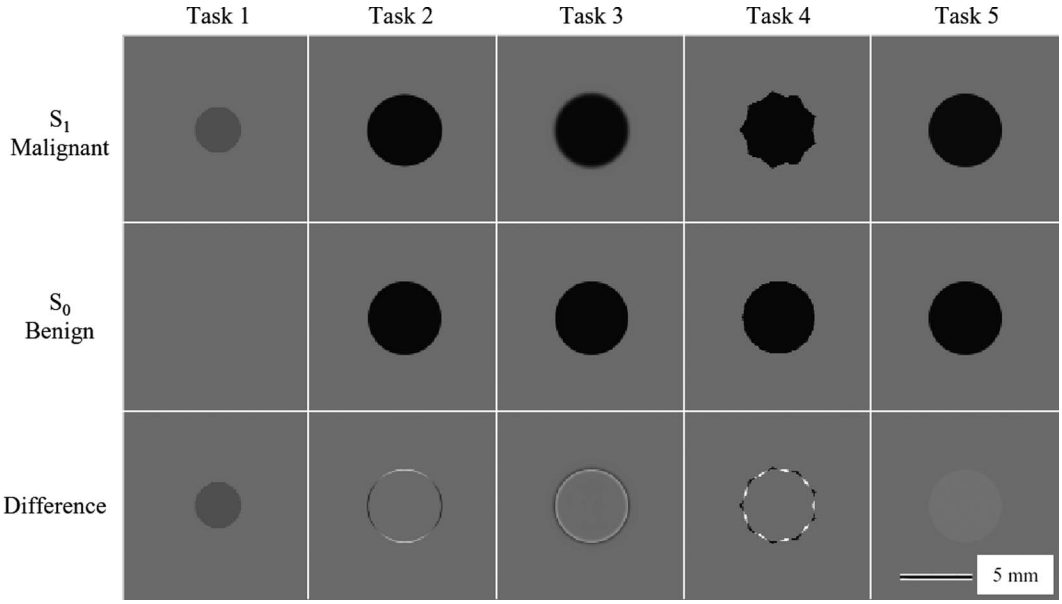


Fig. 4. Object variance profiles used to evaluate beamforming. Malignant (top) and benign (middle) profiles are shown along with difference between the two profiles (bottom).

also measured the noise power spectrum for the Siemens Antares system over the 20 dB signal bandwidth, and we found it was flat within 2 dB, provided no time-gain compensation (TGC) was applied. With TGC, there was systematic variability in the noise power as much as 6 dB. Of course, these echo SNR values vary depending on the system parameters and the scattering strength. However, we would expect our efficiency results to be somewhat more stable because the figure of merit is a ratio of contrast thresholds. Any change in contrast threshold caused by a different SNR should be somewhat compensated by a change in the contrast threshold of the ideal observer, resulting in a similar level of efficiency. This process will eventually break down in the limit of no noise (i.e., infinite SNR), where all reasonable beamformers (i.e., beamformers that do not zero frequencies with positive magnitude) are equivalent.

Examples of the system PSF for various receiver channels within the moving aperture are shown in Fig. 3. As expected, PSFs near the edge of the aperture window are more tilted than those in the center, and all have extending lateral edge waves. The PSF after delay-and-sum beamforming is shown on the far right side of Fig. 3. The effect of beamforming here is to concentrate energy in the central region of the pulse and to modulate the effects of the edge waves.

B. Panel of Tasks

We have developed a panel of simple detection and discrimination tasks for evaluating observer performance on various features of interest related to detection and diagnosis of breast cancer by ultrasound. The five tasks were selected in consultation with a practicing mammographer who routinely uses ultrasound as an adjunct to X-ray mammography. These are simple classification tasks with two possible classes corresponding to our notions of benign and malignant.

Fig. 4 shows variance profiles for the five tasks used in this work. For clarity, the difference between the malignant and benign variance profiles is also shown in the bottom row of the panel. Task 1 represents detection of a 3-mm low-contrast hypoechoic region. Tasks 2–4 all involve discriminating features from the boundary of a suspicious region. This corresponds to the frequent use of ultrasound for differential diagnosis of breast lesions in adjunct examinations. Task 2 considers lesion eccentricity with increased eccentricity being a sign of malignancy. Task 3

considers a soft margin as opposed to a well circumscribed boundary. Here, a soft margin is indicative of malignancy. Task 4 considers the presence of spiculations in the border as a malignant feature as compared with a smooth border, which tends to indicate a benign lesion. Finally, Task 5 considers the presence of hypoechoic material in the lesion interior as a feature of malignant lesions as opposed to nearly anechoic interiors (i.e., scattering variance is 5% of surrounding region), which might indicate a cyst or other benign finding.

C. Monte-Carlo Simulations

Each simulation was based on a total of 8000 sample objects (4000 malignant and 4000 benign). These initial simulation data were used to evaluate all three types of images (RF, Wiener-filtered envelope, and B-mode envelope), all five beamforming systems (2DMF, PMF, AMF, 1DMF, and DS), and at a set of 4 or 5 object contrasts to build up a lookup table for evaluating efficiency. In this way, we ensured that all methods were compared on all the same underlying set of objects and realizations of acquisition noise. The raw data for the simulation was obtained by multiplying the object by the square root of the appropriate variance map, and then passing it through the transfer functions of the echo-forming and beamforming systems before adding the acquisition noise filtered by the square root of the system noise power spectrum. In this way, the resulting data represented beamformed RF signals.

The ideal observer was evaluated by computing the decision variable in (21). Inverse covariance products needed in the ideal observer computation were implemented iteratively through a power series [5], [16]. The Weiner-filtered envelope observer was implemented by Wiener filtering the (scan-converted) RF data and then computing the magnitude of the analytic signal as shown in (29). B-mode envelopes were computed similarly without the Wiener filtering step. Thus for a given image type, beamforming system, and task-dependent object contrast, there were a total of 4000 malignant decision variables and 4000 benign decision variables. Performance was estimated using (22). Standard errors were obtained by bootstrap resampling of the decision variables.

In this study, we used two measures of convergence for computing the ideal observer test statistic. The first measure was that the last iterative update should have at most only a very small improvement in performance. Let PC_k be the performance at iteration k. The criterion was implemented by requiring that $\mathrm{PC}_k - \mathrm{PC}_{k-1} < 0.001$ at the final k. This ensured that iterative updates were small at convergence. The second criterion was meant to ensure that the difference between the current value and the fully converged value is small. Because the power-method for computing an inverse covariance matrix product is known to have approximately linear convergence, the difference between the fully converged performance and the performance at iteration k should decrease at a rate proportional to $1/k$. To implement this convergence criterion, we fit the function $\mathrm{PC}_k = \mathrm{PC}_\infty - \alpha/k$ over a range of the final iterations, where PC_∞ is the fully converged performance, and α is the linear rate of convergence. At the final iteration, we required that $\mathrm{PC}_\infty - \mathrm{PC}_k < 0.001$ at all contrasts. We found that Task 1 converged very quickly, meeting all criteria by the 2nd iteration, although a total of 5 were used. For Tasks 2 and 4, the criteria were achieved after 50 iterations. Task 3 achieved the convergence criteria after 80 iterations and Task 5 required 100 iterations.

As seen in (22), efficiency is computed as a ratio of threshold contrasts. We use multiple contrasts to find the threshold contrast by interpolation. In previous work [5] we have found that simple linear interpolation is most effective when the performance measure is the detectability index d' rather than PC. We converted each PC to a detectability using the standard formula [8] for 2AFC experiments, $d' = \sqrt{2}\Phi^{-1}(\mathrm{PC})$, with Φ being the cumulative normal distribution function. Linear interpolation was then used to find the contrast where PC = 80% (i.e., $d' = 1.19$). Standard errors on thresholds were determined by propagation of error from the Monte-Carlo simulation. Once thresholds were computed, efficiency was computed by taking the appropriate ratio.

IV. Results and Discussion

A. Rationale for Signal Compression

A critical component of beamforming, as we have defined it, is a compression step. Signals from all receiver channels in a moving aperture are combined irreversibly into a single echo signal whose envelope represents the B-mode image line for that active aperture location. But if compression inherently involves a loss of information, then we might conclude that the only justification for signal compression in beamforming is the practical requirement of producing an image.

The ideal observer bases its decision on a scalar test statistic, which can be regarded as the ultimate level of data compression. However, it does so using a computation that is highly task specific. We use the array of tasks in this work as simple examples of the many possible purposes for acquiring an ultrasound image of the breast. Processing the image for any one specific task would likely be detrimental to the performance of the many others. Hence, in using the ideal observer to guide image processing, we seek to find steps that are independent of the task.

Our derivation in (11) shows that compression can, in principle, be achieved without loss of diagnostic information. The fact that the ideal observer decision variable for any simple detection or discrimination task can be as effectively implemented on compressed data as on the original data demonstrates the high degree of statistical redundancy in the acquired data, which can be reduced without penalty by the 2DMF beamformer. We can then

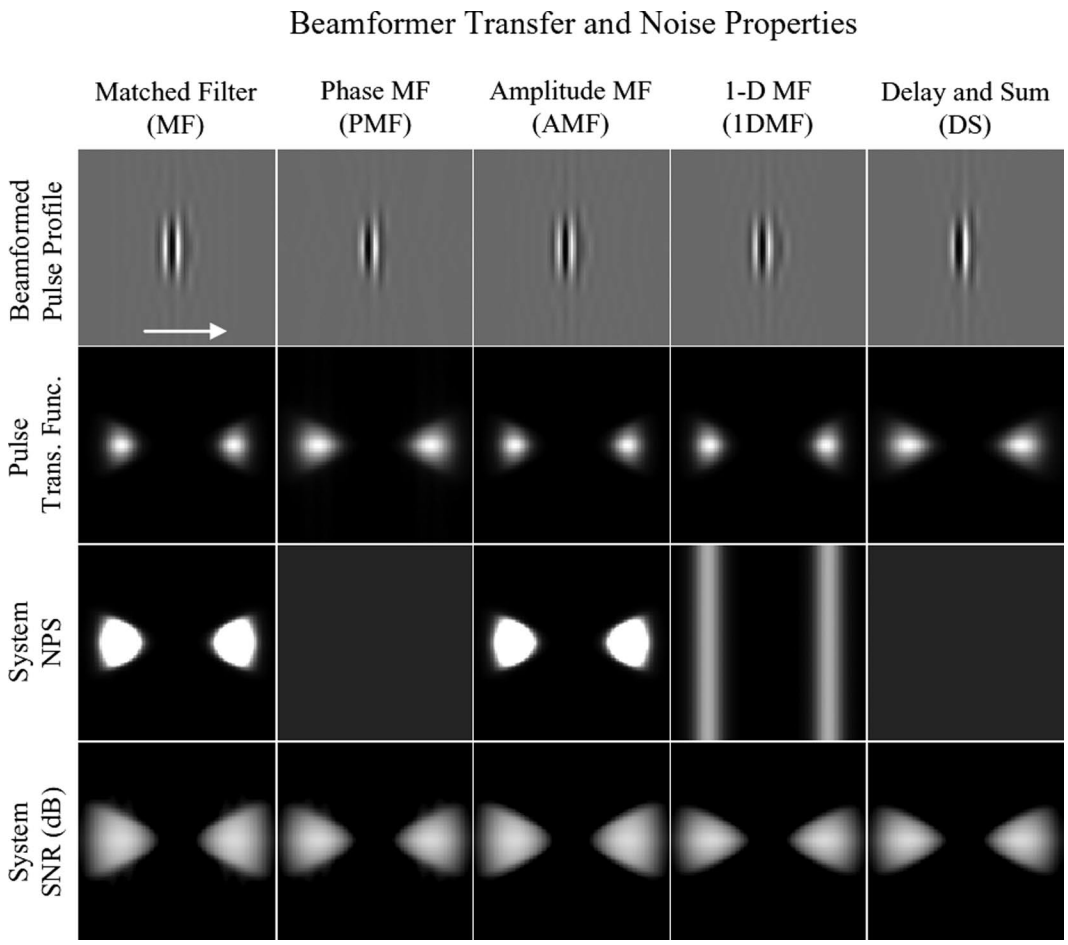


Fig. 5. Statistical properties of beamformers. The five beamforming approaches used in this study are analyzed here for their transfer and noise properties. The top row gives the beamformed pulse profile (horizontal arrow at left shows direction of axial propagation), which is effectively the PSF of the beamformed RF data. The second row of the panel is the amplitude of the transfer function of this pulse (origin is at the center of each image). The third row of the panel shows the noise power spectrum of the composite acquisition noise term. The fourth row of the panel displays the frequency SNR with intensity log-compressed to a decibel scale. Note that the top row is magnified by a factor of 4 relative to the lower rows.

interpret other beamformers, such as delay-and-sum, as approximations to this optimal transform. Our results on the efficiency of delay-and-sum relative to the 2DMF beamformer will give a sense of the price of this approximation.

B. Beamformer Properties

We have analyzed beamforming with various forms of matched filtering and by standard delay-and-sum transformations. Each pre-compression process results in somewhat different statistical properties in the compressed RF data they produce. In Fig. 5, we characterize these statistical properties as images, by considering the scan-converted RF data. We utilize stationary measures such as the transfer function, $\left|\sum_a \hat{b}_a[k_x,k_y]\hat{h}_a[k_x,k_y]\right|$, and noise power spectrum (NPS), $\sigma_n^2\sum_a \left|\hat{b}_a[k_x,k_y]\right|^2$, because the system is assumed to be locally shift invariant in the focal region. The top panel of Fig. 5 shows the pulse profile after signal compression, which can be thought of as the point-spread function (PSF) of the system. The transfer function of each PSF in the next panel shows how the bandwidth depends on the beamforming approach. The 2DMF beamformer and AMF beamformers appear to be somewhat more narrow-band than either the DS or the PMF beamformers. The 1DMF agrees with the 2DMF in the axial direction but has a slightly broader bandwidth in the lateral direction. Noise power in the 2DMF and AMF beamformers is also concentrated around the transfer function bandwidth. By contrast, the phase-only DS and PMF beamformers leave the noise white with a flat power spectrum over the Nyquist band. Noise power in the 1DMF is limited to the band of the 2DMF on the axial direction and is white to the Nyquist frequency in the lateral direction. As a way to normalize the different transfer functions and power spectra, we compute a signal-to-noise ratio defined as the ratio of the transfer function squared to the noise power spectrum at each frequency. The ratio is then log-compressed to make it equivalent to units of decibels. In these images of frequency SNR, the various systems generally look similar with a slightly larger region of high SNR in the 2DMF beamformer.

Fig. 6 plots echo and peak SNRs of each beamforming system. The echo SNR is defined as receiver variance (in-

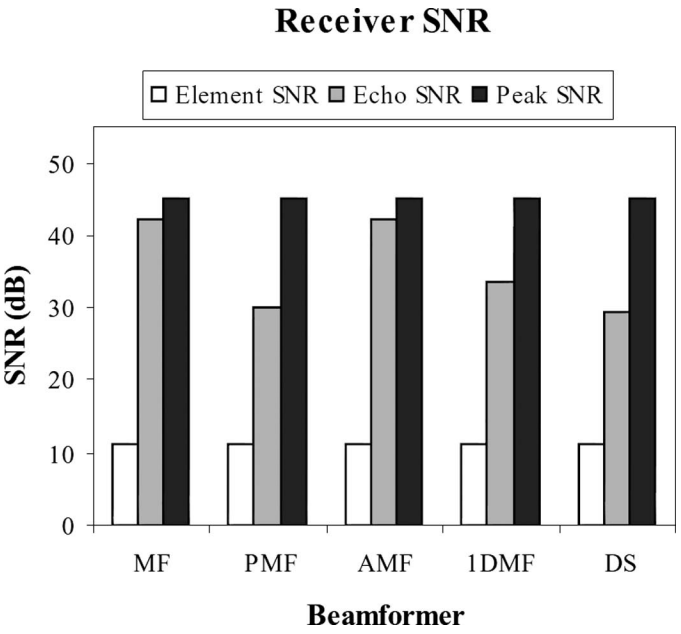


Fig. 6. Receiver SNRs. Various SNRs are plotted as a function of the beamforming approach used (acronyms defined in Fig. 5). Element SNR is the average echo SNR across all 96 elements and is the same for all approaches because it precedes the matched filtering operation. Echo and peak SNRs give different assessments of the beamforming approaches.

tegrated PS) caused by a uniform scattering field divided by the variance of acquisition noise. The peak SNR is the largest value of SNR for a single frequency in the signal band. Across the individual receiver channels, echo SNR averages 11.2 dB with a standard deviation of 0.1 dB. Because this occurs before any beamforming is done, it has no dependence on the beamformer used. After beamforming, the echo SNR is substantially higher for the 2DMF and AMF beamformers than either the PMF or DS beamformer. The 1DMF falls between the two. This appears to reflect the reduction in acquisition noise variance from the amplitude modulation in the beamformer. However, when peak SNR is used, there appears to be little dependence on the beamformer. Our echo SNR results are consistent with the findings of Guenther and Walker [17] using the related measure of cystic resolution [18], [19], although further investigation is needed to see how task performance and cystic resolution are related.

Fig. 7 gives examples of envelope images derived from different beamformers. The object being simulated is a 5-mm hypoechoic region, and all envelope images were computed by taking the magnitude of the analytic signal. The upper row of the panel shows envelope images computed directly from the simulated RF data for each of the five beamformers evaluated. The lower row shows the effect of Wiener filtering the RF data before computing the envelope image. Careful inspection of the images reveals some differences in speckle size between beamformers in both rows. However the more striking difference is between the standard B-mode envelope and the Weiner filtered envelope, where speckle size is reduced considerably. It is not clear that this reduction in speckle size will necessarily result in better performance because the contrast of the simulated lesion appears to be reduced somewhat as well.

C. Efficiency of Beamforming Systems in RF Frames

Fig. 8 plots the efficiency of beamforming in the RF data as well as Wiener-filtered and B-mode envelope images. From the derivation in Section II-C-2, we know that the ideal observer acting on RF data that has been compressed using the 2DMF beamformer is equivalent to the ideal observer acting before compression takes place, and therefore sets an upper bound on performance. All efficiencies are computed using the contrast threshold of the 2DMF beamformer in the numerator of (22). The result of this is that the 2DMF acting on RF data has an efficiency of 100% in all tasks.

For the RF data, efficiency drops consistently across the different beamformers arranged from the 2DMF to DS beamformers. For standard delay-and-sum beamforming, efficiency values range from 24% to 75% with an average across tasks of 45%. Therefore, taking the efficiency as

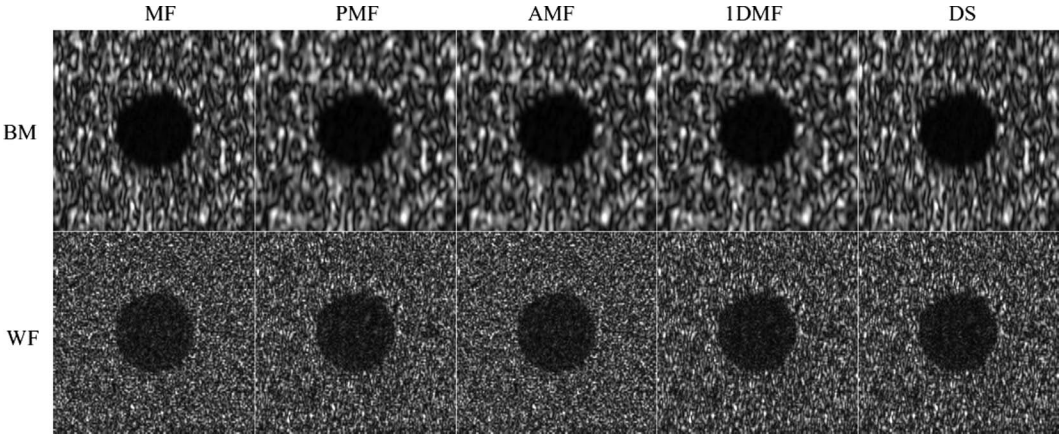


Fig. 7. Example envelope images. B-mode and Wiener-filtered envelope images for the five beamforming approaches evaluated in this study. All images are generated from the same simulated RF element data. Both the choice of beamformer and post-compression processing influence speckle sizes and textures.

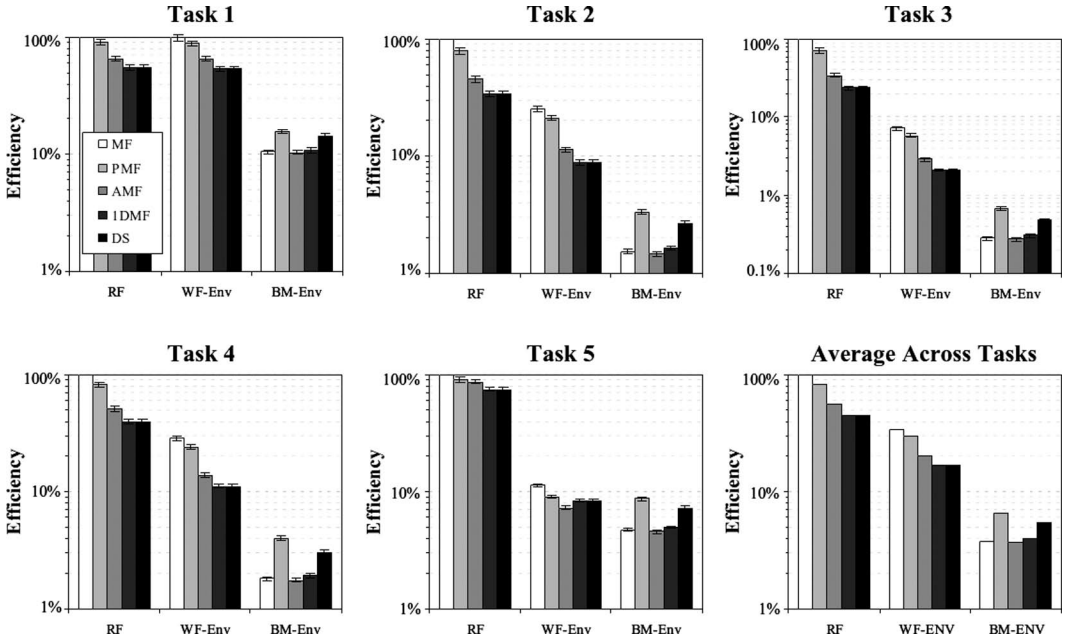


Fig. 8. Task efficiency. Efficiency of the ideal observer acting on pre-envelope RF data, and Smith-Wagner approximations to the ideal observer acting on Wiener-filtered envelope (WF-ENV) and B-mode envelope (BM-ENV) images are shown across beamforming systems for each task (the legend in the upper left plot applies to all plots). Note that a different logarithmic range is used on the y-axis for Task 3. Bootstrapped error bars in Tasks 1–5 represent $+/-$ 1 standard error due to the finite number of Monte-Carlo samples (4000) used to estimate efficiency. No error bars are plotted for the average across tasks.

a measure of diagnostic information suggests that standard delay-and-sum beamforming loses roughly half the diagnostic information in the acquired data. The phase-matched filter efficiency ranges from 71 to 90% with an average value across tasks of 83%. Thus, incorporating only the phase component of the matched filter makes up roughly 60% of the efficiency drop between the matched-filter and delay-and-sum beamformers. The amplitude matched filter efficiency ranges from 34% to 87%, and makes up an average 21% in the difference between the 2DMF and DS beamformers. This shows that amplitude has a significant, but substantially smaller role than phase in transferring information to the compressed RF signal. A surprising finding was that the 1DMF showed no improvement over the DS beamformer. Beamforming is often viewed as an operation that combines signals from within an aperture to produce a single scan line. Our findings indicate that channel correlations from different apertures play a crucial role in the transfer of diagnostic information in beamforming.

Efficiency of the Wiener filtered envelope images in Fig. 8 is very close to ideal observer performance in Task 1, as has been described previously [5]. In Tasks 2–4, efficiency is substantially less than the ideal observer acting on the RF data, but the relative ordering of the different beamformers is quite similar. Task 5 is the exception, with much less of a drop in efficiency going from the 2DMF to DS beamformers. There is also less of a relative advantage with the PMF, and an observed disadvantage with the AMF beamformers. This is also consistent with previous findings [5] showing less of a benefit to the Wiener-filtered envelope in general for Task 5, and has been explained as poor tuning of a shift-invariant Wiener filter to the lower level of noise in the interior of the lesion. Nonetheless, the 2DMF beamformer maintains roughly a 2-fold advantage over delay-and-sum for Wiener-filtered envelope images.

Somewhat surprisingly, the 2DMF beamformer loses its advantage in standard B-mode envelope images. The 2DMF and AMF beamformers have the lowest efficiency of the five beamformers tested, roughly two-thirds the efficiency of delay-and-sum, and even less when compared with the phase-matched filter, which performed the best of the methods applied to B-mode envelope images. The 1DMF also underperforms DS beamforming in this case. These results show that a substantial advantage in diagnostic information in the compressed RF data can subsequently be lost in the transformation to an envelope image. The advantage of the 2DMF beamformer is only realized with a post-processing step before the envelope image is computed.

D. Relative Efficiency of Envelope Images

There are two possible sources of inefficiency in the envelope images of Fig. 8. Information may be lost by the beamformer or by the transformation to the envelope image. Loss of information by the beamformer shows up as a reduction in efficiency in the resulting RF data. This can be determined from the RF plots in Fig. 8. For understanding of how much information is lost in the envelope transformation, we plot relative efficiencies defined in (29) for the various beamforming systems in Fig. 9. Relative efficiency uses the contrast threshold of the beamformer in the RF divided by the contrast threshold in the B-Mode

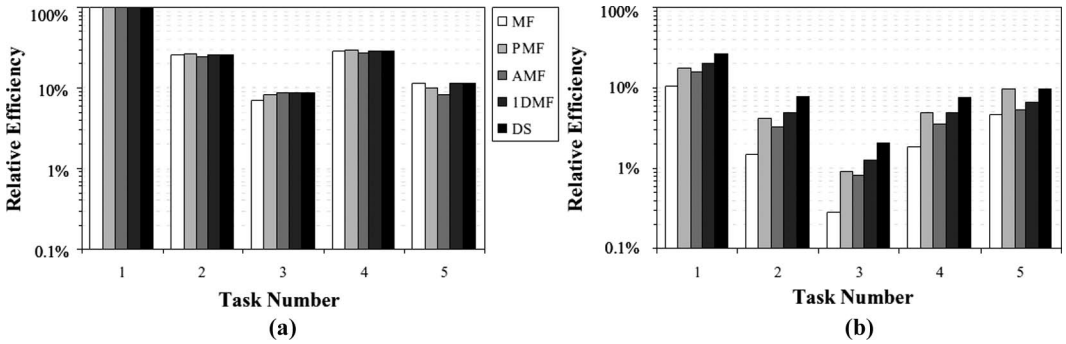


Fig. 9. Relative efficiency of envelope images. Efficiency of Smith-Wagner approximations for envelope images relative to the ideal observer acting on RF data are plotted across tasks for each beamforming system (legend applies to both plots). Although there is substantial variability across tasks in both plots, the Wiener-filtered envelope images (a) show much less variability and generally higher relative efficiency across beamforming systems than do B-mode envelope images (b).

or Wiener-filtered envelope image. This is somewhat different than Fig. 8, where the 2DMF contrast threshold was always used in the numerator. The relative efficiency uses only quantities occurring after the compression step, and therefore isolates the information lost in the envelope transformation.

The relative efficiency of the Wiener-filtered envelope has substantial variability across tasks, but somewhat surprisingly, it has relatively little change for different beamforming systems within a given task. By contrast, the B-mode envelope images show both variability across tasks and variability across beamformers within a task, with delay-and-sum achieving the highest relative efficiency (and the 2DMF the lowest) of the beamforming systems tested. For the 2DMF beamformer, the transformation to a B-mode envelope is so inefficient that its substantial advantage in the beamformed RF domain is completely negated.

E. Limitations of the Current Study

Before concluding, we review the three limiting assumptions of the approach used here. Resolving the effects of these assumptions is the subject of ongoing work by the authors.

We have evaluated beamforming in the context of simple detection and discrimination tasks. The images in each class of a given task are completely characterized by the object variance map, and hence the rather uniform textures of the sample images in Fig. 7. This is an enormous simplification of clinical breast sonography, where contrast and shape variability, as well as a variety of tissue types, lead to many different variance maps for each class. The assumption here is that performance in an array of simple tasks is indicative of performance in more realistic settings.

We make strong Gaussian assumptions in our system model. The finding that the matched-filter beamformer preserves diagnostic information required only a Gaussian assumption of the electronic noise, but our simulation results also made the assumption of Gaussian-distributed object reflectivity that is effectively independent at sampling rates relevant to ultrasound. It is possible that other distributions on reflectivity would lead to different optimal beamforming transforms.

We assume that we have a well-defined system model, and for computational reasons we assume that it can be implemented by a discrete convolution of a system PSF with the sampled object. The assumption of a convolution-based system model is idealized, and at best confined to a small isoplanatic region. Furthermore, the wraparound arising from discrete convolutions is non-physical. We have tried to avoid any effects of wrap-around by making the image large enough that the edges of the image, where wrap-around occurs, are not relevant for task performance. However, even the assumption of a known system model is difficult in practice. The complexity of acoustic wave propagation in nonuniform media and nonlinear effects such as phase aberration make full specification of a system model difficult.

V. Conclusions

The purpose of this work has been to examine beamforming from the perspective of diagnostic information transfer as measured by the ideal observer. We have generalized this approach from previous work for the purpose of analyzing beamforming by developing a stochastic model of image formation based on individual receiver signals. This results in a different system model for each receiver channel in a moving aperture, as well as independent acquisition noise. The mathematical model for each of these channel systems can be assembled into a single large block-matrix system representing image formation data at the level of individual receiver channels. The model can be used to specify the statistics of the received signals which allows us to define an ideal observer decision variable for simple detection and discrimination tasks.

By analyzing the structure of the ideal observer's decision variable, we find that the signals from independent receiver channels can be compressed into a single signal by a 2-D matched filtering operation followed by summation across channels at each time point. We demonstrate

that this transformation does not lose any diagnostic information in the sense that the ideal observer acting on the compressed signal is equivalent to the ideal observer implemented before compression on all the acquired data. The fact that receive-channel signal compression can be implemented without loss of diagnostic information provides a rigorous theoretical justification for beamforming from the perspective of task performance.

We use five tasks related to breast sonography to show the degree to which the form of the beamforming transform can influence the transfer of information, as measured by statistical efficiency of task performance. At the level of the compressed RF signal, the 2DMF beamformer increases efficiency by a factor of 2.2 over standard delay-and-sum beamforming. The 2DMF beamformer modulates each channel signal both in amplitude and phase as well as laterally and axially. We have evaluated phase matched-filter, amplitude matched filter, and 1-D axial matched filter beamformers to assess the role of each of these components. The use of phase information achieves most of the gains of the optimal 2DMF, with a factor of 1.8 in increased efficiency relative to delay-and-sum beamforming. The amplitude plays a more modest role, showing only a factor of 1.2 improvement in efficiency relative to delay-and-sum. Somewhat surprisingly, the 1-D matched filter shows no improvement at all over delay-and-sum beamforming.

It is also clear that the form of the beamformer can have a strong effect on the diagnostic information of the final envelope image. The Smith-Wagner efficiency of the B-mode envelope image drops by more than two orders of magnitude in some cases, with a strong dependence of the beamforming system and the task being considered. Interestingly, application of a Weiner filter to the beamformed RF data essentially negates differences in the relative efficiency of different beamforming systems going from RF to an envelope. Nonetheless, our results show that post-compression processing (here we used a spatial Wiener filter) is necessary for the information in the 2DMF beamformer to transfer to a final envelope image.

Our results are all derived from simulations, where we know the underlying system model well. This is rarely the case with real scanners used in clinical applications. Nonetheless, our results suggest that efforts to develop more accurate system models in these more challenging environments—and the incorporation of these models into the processing of received data—can, in principle, have a substantial effect on task-relevant information in the final image.

References

[1] D. G. Wildes, R. Y. Chiao, C. M. W. Daft, K. W. Rigby, L. S. Smith, and K. E. Thomenius, "Elevation performance of 1.25D and 1.5D transducer arrays," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 44, pp. 1027–1038, Sep. 1997.

[2] A. Macovski, "Ultrasonic imaging using arrays," *Proc. IEEE*, vol. 67, no. 4, pp. 484–495, 1979.

[3] B. A. J. Angelsen, *Ultrasound Imaging: Waves, Signals, and Signal Processing*, vol. I and II: Trondheim, Norway: Emantec, 2000.

[4] A. Macovski, *Medical Imaging Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1983.

[5] C. K. Abbey, R. J. Zemp, J. Liu, K. K. Lindfors, and M. F. Insana, "Observer efficiency in discrimination tasks simulating malignant and benign breast lesions imaged with ultrasound," *IEEE Trans. Med. Imaging*, vol. 25, pp. 198–209, Feb. 2006.

[6] R. J. Zemp, M. D. Parry, C. K. Abbey, and M. F. Insana, "Detection performance theory for ultrasound imaging systems," *IEEE Trans. Med. Imaging*, vol. 24, pp. 300–310, Mar. 2005.

[7] A. D. Whalen, *Detection of Signals in Noise*. New York, NY: Academic Press, 1971.

[8] D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics*. Los Altos, CA: Peninsula Publishing, 1988.

[9] H. H. Barrett, C. K. Abbey, and E. Clarkson, "Objective assessment of image quality. III. ROC metrics, ideal observers, and likelihood-generating functions," *J. Opt. Soc. Am. A*, vol. 15, pp. 1520–1535, Jun. 1998.

[10] J. A. Jensen, "Field: A program for simulating ultrasound systems," *Med. Biol. Eng. Comput.*, vol. 4, suppl. 1, pp. 351–353, 1996.

[11] J. A. Jensen and N. B. Svendsen, "Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 39, no. 2, pp. 262–267, 1992.

[12] C. K. Abbey, R. J. Zemp, J. Liu, K. K. Lindfors, and M. F. Insana, "Observer efficiency in boundary discrimination tasks related to assessment of breast lesions with ultrasound," in *SPIE Medical Imaging*, San Diego, CA, 2006, art. no. 61460O.

[13] W. W. Peterson, T. G. Birdsall, and W. C. Fox, "The theory of signal detectability," *Trans. IRE (PGIT)*, vol. 4, pp. 171–212, 1954.

[14] R. N. Bracewell, *The Fourier Transform and its Applications*, 2nd ed. New York, NY: McGraw-Hill, 1986.

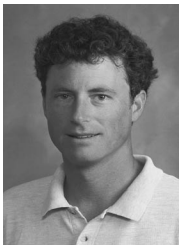
[15] S. W. Smith, R. F. Wagner, J. M. Sandrik, and H. Lopez, "Low contrast detectability and contrast/detail analysis in medical ultrasound," *IEEE Trans. Sonics Ultrason.*, vol. 30, pp. 164–173, May 1983.

[16] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed. Baltimore, MD: Johns Hopkins University Press, 1989.

[17] D. A. Guenther and W. F. Walker, "Generalized cystic resolution: A metric for assessing the fundamental limits on beamformer performance," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 56, pp. 77–90, Jan. 2009.

[18] D. Vilkomerson, J. Greenleaf, and V. Dutt, "Towards a resolution metric for medical ultrasound imaging," in *Proc. 1995 Int. IEEE Ultrasonics Symp.*, 1995, pp. 1405–1410.

[19] K. Ranganathan and W. F. Walker, "Cystic resolution: a performance metric for ultrasound imaging systems," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 54, pp. 782–792, Apr. 2007.



Craig K. Abbey received his Ph.D. degree in applied mathematics from the University of Arizona in 1998. He was a postdoctoral fellow in medical physics at Cedars-Sinai Medical Center and UCLA from 1998 to 2001. From 2001 to 2004, he was a member of the faculty in biomedical engineering at UC Davis, where he retains an adjunct position. His primary affiliation is in the Dept. of Psychology at UC Santa Barbara. His research focuses on the transfer of diagnostic information in medical imaging systems utilizing theoretical analysis of image statistics as well as visual psychophysics for evaluating human observer performance.

Nghia Q. Nguyen received his B.S. degree in electronics and telecommunications from Hanoi University of Technology in 2002, and his M. S. degree in electrical and computer engineering from the University of Illinois at Urbana-Champaign in 2009, where he is working toward his Ph.D. degree. He was a researcher in Optical Communications and Networking at Research Institute of Posts and Telecoms (RIPT) of Vietnam from 2002 to 2005, and is currently a Research Assistant at

the University of Illinois at Urbana-Champaign. His research focuses on ultrasound medical imaging, signal processing, inverse problems, and objective assessment of image quality.

Michael F. Insana received his Ph.D. degree in medical physics from the University of Wisconsin–Madison in 1983. His dissertation topic was acoustic scattering and absorption in biological tissues. He was a research physicist at the FDA from 1984 to 1987, where he developed methods for describing tissue microstructure from the statistical properties of echo signals. From 1987 to 1999 he was with the Department of Radiology at the University of Kansas Medical Center. There, he directed an imaging research lab that applied ultrasonic imaging to the evaluation of progressive renal failure and breast cancer. Between 1999 and 2004, Mike was Professor of Biomedical Engineering at the University of California, Davis where he directed the graduate program. He is currently Professor and Head of the Department of Bioengineering at the University of Illinois at Urbana-Champaign. His research includes the development of ultrasonic methods for imaging soft tissue microstructure, viscoelasticity, and blood flow. The goal is to understand basic mechanisms of tumor formation and responses to therapy. He is also interested in the principles of imaging system design and performance evaluation, signal processing, detection and estimation. He is a Fellow of the Acoustical Society of America, the Institute of Physics, and the American Institute of Medical and Biological Engineering. He is also a Senior Member of the IEEE and Associate Editor for *IEEE Transactions on Medical Imaging*.