

# An Adaptive Filter to Approximate the Bayesian Strategy for Sonographic Beamforming

Nghia Q. Nguyen\*, Craig K. Abbey, and Michael F. Insana

**Abstract**—A first-principles task-based approach to the design of medical ultrasonic imaging systems for breast lesion discrimination is described. This study explores a new approximation to the ideal Bayesian observer strategy that allows for object heterogeneity. The new method, called iterative Wiener filtering, is implemented using echo data simulations and a phantom study. We studied five lesion features closely associated with visual discrimination for clinical diagnosis. A series of human observer measurements for the same image data allowed us to quantitatively compare alternative beamforming strategies through measurements of visual discrimination efficiency. Employing the Smith–Wagner model observer, we were able to breakdown efficiency estimates and identify the processing stage at which performance losses occur. The methods were implemented using a commercial scanner and a cyst phantom to explore development of spatial filters for systems with shift-variant impulse response functions. Overall we found that significant improvements were realized over standard *B*-mode images using a delay-and-sum beamformer but at the cost of higher complexity and computational load.

**Index Terms**—Breast sonography, ideal observer, image quality, iterative Wiener filter, task-based design.

## I. INTRODUCTION

THE computational speed and configuration flexibility of current digital beamformers now make it realistic to consider implementing more complex alternatives to the current standard of delay-and-sum (DS) beamforming [1], [2]. Adaptive [3], [4], minimum error [5], and inverse operator [6] beamforming strategies are a few of the methods now being explored for their ability to improve spatial and contrast resolution while preserving the echo signal-to-noise ratio (SNR) [7], [8]. Beamforming refers to any modification of the amplitude and phase of transmitted and received array-element RF signals applied before envelope detection. This includes any pre- and post-summation filtering. Beamforming is primarily responsible for determining the spatial impulse response of the imaging system

that ultimately limits the diagnostic information presented to observers.

This research uses task-based methods to assess adaptive sonographic signal processing techniques applied to channel-summed radio-frequency (RF) echo signals. The tasks used for the study are chosen to represent features used to discriminate malignant from benign lesions in breast sonography. Our approach utilizes performance in these discrimination tasks to find optimal signal processing strategies.

Task based assessment of performance uses probabilistic models of tissue properties and the subsequent acquisition of signals as the basis for task performance [9], [10]. The ideal observer makes use of the statistics to optimally perform a detection, discrimination or more complex task [11], [12]. For this study, multivariate Gaussian models are used to represent RF data from malignant and benign classes [13], [14], and the optimal test statistic—the ideal Bayesian observer—is derived from the ratio between probability density functions (pdf) of the two classes. The ideal observer operating on echo signals provides an upper bound for task performance against the other observers, including trained humans. Thus it provides a measure of the efficiency for transferring diagnostic information through the image formation process. Our contribution has been to analyze the log-likelihood ratio computed for key diagnostic features of breast cancer diagnosis, and interpret the results as optimal strategies for transferring diagnostic information from echo signals into sonograms.

A practical challenge to computing ideal-observer test statistics is high dimensionality. The process requires inversion of two large covariance matrices. Brute-force inversion methods are rarely fruitful, so iterative methods are often necessary. In the initial presentation of this method, it was shown how these large covariance matrices could be partitioned into stationary and nonstationary components in order to expand the matrix in a power series [14]. A key theoretical finding of the analysis was that the first-order power series approximation is equivalent to a stationary Wiener deconvolution filter applied to the channel-summed RF signals. The resulting envelope images yielded a measurable improvement in performance when the task was discriminating low-contrast lesion features. However, performance was reduced for high-contrast lesions, even if the discrimination task is itself low contrast, specifically when observers were asked to discriminate anechoic and hypoechoic lesions. These previous findings suggested that, for imaging situations where there are large signal heterogeneities, filters must adaptively tune to the echo statistics wherever there is diagnostic information.

Manuscript received April 09, 2010; revised June 28, 2010; accepted July 06, 2010. Date of publication July 19, 2010; date of current version December 30, 2010. This work is supported by the National Institutes of Health under award CA118294. Asterisk indicates corresponding author.

\*N. Q. Nguyen is with the Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: nnguyen6@illinois.edu).

C. K. Abbey is with the Department of Psychology, University of California, Santa Barbara, CA 93106 USA (e-mail: abbey@psych.ucsb.edu).

M. F. Insana is with the Department of Bioengineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: mfi@illinois.edu).

Digital Object Identifier 10.1109/TMI.2010.2059035

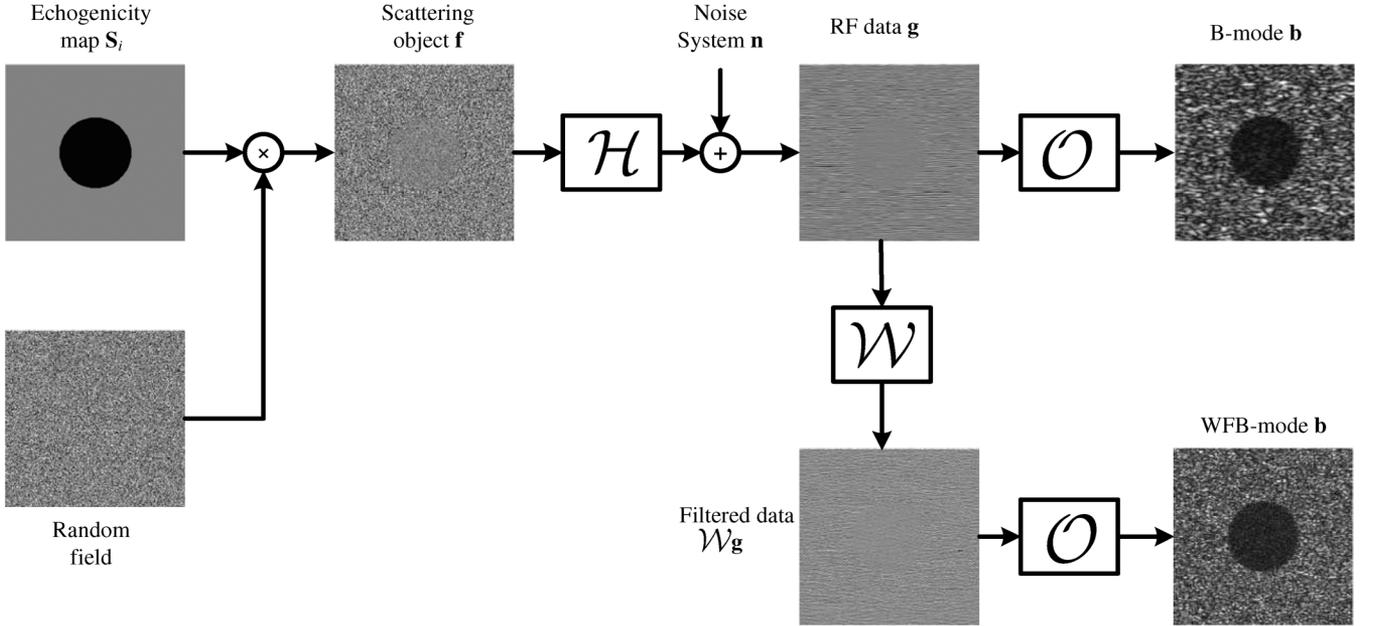


Fig. 1. Graphical representation of the process to generate sonographic images. *B*-mode images are formed using conventional delay-and-sum beamforming, whereas WFB-mode images include an additional Wiener filtering step applied before envelope detection.

This paper describes an adaptive filter that better matches the optimal processing of the ideal observer. A binary segmentation is used to modify the Wiener filter to local statistics. Observer performance results show enhanced performance for all five of the diagnostic features examined when compared with the DS beamformer but with additional computational overhead. The concepts are demonstrated on a commercial system by imaging a tissue-mimicking phantom where results include a realistic, shift-variant model for the system impulse response.

## II. METHODS

### A. Numerical Modeling

The process of image formation is depicted graphically in Fig. 1. Incoherent tissue scattering is considered to be a zero-mean Gaussian process represented by the function  $f(\mathbf{x})$ , where  $\mathbf{x}$  represents spatial position. We approximate the continuous scattering function by a vector of sampled points  $\mathbf{f}$ . The imaging system maps object  $\mathbf{f}$  into an echo data vector  $\mathbf{g}$ . Image formation in pulse-echo ultrasound is well approximated by a linear transform under the first Born approximation [15]. Specifying the system by matrix  $\mathbf{H}$ , we approximate the continuous-space-to-discrete-time linear operator  $\mathcal{H}$  [16]. For the purposes of this work,  $\mathbf{g}$  are the channel-summed RF signals that comprise the A-lines before envelope detection. *B*-mode image data, denoted by the vector of pixel values  $\mathbf{b}$ , are obtained from  $\mathbf{g}$  through a nonlinear transform involving demodulation and scan conversion. We denote this process by the nonlinear operator  $\mathcal{O}$ .

In this work, the vectors  $\mathbf{g}$  and  $\mathbf{b}$  represent data in the scan plane of a 1-D linear array transducer. We will concentrate on the effect of filtering the RF echo signals before envelope detection. This can be thought of as applying a filtering operator  $\mathcal{W}$

before envelope detection. These processes are summarized by the equations

$$\mathbf{g} = \mathcal{H}f(\mathbf{x}) + \mathbf{n} \quad \text{and either} \quad \mathbf{b} = \mathcal{O}\mathbf{g} \quad \text{or} \quad \mathbf{b} = \mathcal{O}\mathcal{W}\mathbf{g},$$

where the additive component  $\mathbf{n}$  represents the system noise, modeled by an independent Gaussian process  $\mathcal{N}(\mathbf{0}, \sigma_n^2 \mathbf{I})$ . Since  $\mathcal{H}$  is approximated by the matrix  $\mathbf{H}$ , the first equation can be written in terms of a matrix multiplication, given by

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n}. \quad (1)$$

For the purpose of simulation, object vectors should be simulated numerically by generating  $\mathbf{f}$  with very high sampling frequencies for mimicking a convolution of analog signals. The output data  $\mathbf{g}$  is then downsampled to get the pixel size of a commercial system. In that case,  $\mathbf{f}$  and  $\mathbf{g}$  are not the same length and  $\mathbf{H}$  is a rectangular matrix of dimensions  $M \times N$ , where  $M \ll N$ . However, for a band-limited impulse response and fixed echo SNR, convolving before or after downsampling gives the same simulation results. Therefore,  $\mathbf{f}$  and  $\mathbf{g}$  can be the same length, which results in a square matrix for  $\mathbf{H}$ . The  $m$ th row of  $\mathbf{H}$  is the spatiotemporal impulse response that yields  $g[m]$ . Rows contain different impulse responses when modeling a shift-variant system. However, under the assumption of shift invariance,  $\mathbf{H}$  is a block-Toeplitz matrix, which provides advantages for computation.

We chose sampling intervals for 2-D signals that were 0.02 mm in range (corresponding to 40 Msample/s rate and sound speed  $c = 1.54$  mm/ $\mu$ s) and 0.2 mm in cross range (corresponding to the array pitch). The data were reordered to obtain column vectors for  $\mathbf{g}$ ,  $\mathbf{f}$ , and  $\mathbf{b}$  in (1). The final *B*-mode image  $\mathbf{b}$  is interpolated in cross range to render images for display with a square 0.02 mm pixel size. The interpolation

does not alter the information content of the envelope image, but mimics the final step of demodulation from RF data to  $B$ -mode images in ultrasound imaging systems.

We will consider a variety of discrimination tasks in which objects are associated with one of two classes, “Benign” (class 1) or “Malignant” (class 2). All task information is contained in the covariance of the object vector since we consider only incoherent scattering. The object covariance matrix is nonstationary,  $\sigma_{\text{obj}}^2(\mathbf{I} + \mathbf{S}_i)$ , where  $\sigma_{\text{obj}}^2$  is the variance of a Gaussian process and  $\mathbf{S}_i$  is a deterministic diagonal matrix for the  $i$ th class of data. The diagonal elements of  $\mathbf{S}_i$  specify the relevant object components that define each class. For example, the “echogenicity map” in Fig. 1 shows the diagonal elements of  $\mathbf{S}_i$  (the class,  $i$ , is arbitrary in this case) plotted as a 2-D image. In this case the covariance matrix defines a circular hypoechoic lesion.

### B. Pulse-Echo Impulse Response

Pulse-echo impulse responses were generated by the Field II program [17], [18]. System features were modeled after the VF10-5 linear array probe on the SONOLINE Antares System (Siemens Medical Solutions, Mountain View, CA). The array had 192 elements of dimensions  $0.18 \times 5.0 \text{ mm}^2$  separated by a 0.02 mm kerf, an element pitch of 0.2 mm, and a 96-element active aperture on transmit and receive. The  $f/2$  in-plane aperture had a 40-mm transmit/receive focal length. In elevation, the  $f/5$  aperture had a focal length of 25 mm. We applied a two-cycle excitation voltage and measured a 53% pulse-echo bandwidth about a 7.0 MHz center frequency. RF waveforms were sampled at 40 Msample/s. The system matrix  $\mathbf{H}$  was assembled from modeled impulse responses for this system. Except where noted, one matrix was used to both simulate and filter echo data.

### C. Discrimination Features

Consultation with a radiologist yielded a panel of five breast lesion features used for sonographic diagnosis [14], which is based on the BIRADS atlas [19]. Features became visual discrimination tasks by defining a malignant  $\mathbf{S}_2$  and benign  $\mathbf{S}_1$  matrix pair for each. Listed in order of malignant and benign, Task 1 involved detecting a low-contrast hypoechoic lesion versus a no-lesion background; Task 2 required discrimination of an elongated eccentric lesion from a circular lesion; Task 3 was discrimination of a soft, poorly defined boundary from a well-circumscribed boundary; Task 4 required discrimination of spiculated boundary irregularities from a smooth circular boundary; and Task 5 involved discriminating a very weakly scattering hypoechoic interior from an anechoic interior. Tasks 1 and 5 challenge the system to image large-area diagnostic features while 2-4 define lesion boundary features. Task 5 was distinct in that it was the only large-area, high-contrast lesion involved in a discrimination task. Variance maps of the five tasks were graphically illustrated ([14, Fig. 2]).

### D. Wiener Filtering [14]

Multivariate normal processes remain multivariate normal following linear transformations. Thus, passing the object through the noisy linear transformation in (1) results in another

Gaussian process for each class. The class covariance matrix for the RF data is given by

$$\Sigma_i = \sigma_{\text{obj}}^2 \mathbf{H}(\mathbf{I} + \mathbf{S}_i) \mathbf{H}^t + \sigma_n^2 \mathbf{I}. \quad (2)$$

The zero-mean MVN process characterizing  $\mathbf{g}$  under the two class hypotheses are

$$\begin{aligned} \mathbf{g}_1 &\sim \text{MVN}(\mathbf{0}, \Sigma_1) \\ \mathbf{g}_2 &\sim \text{MVN}(\mathbf{0}, \Sigma_2). \end{aligned} \quad (3)$$

The covariance matrices still capture all the relevant statistics of task performance, but now they are not diagonal because of the influence of the imaging system via  $\mathbf{H}$ .

The test statistic quantifying the ideal-observer response is given by the log-likelihood ratio [14], [20]

$$\lambda(\mathbf{g}) = \ln \frac{p(\mathbf{g}|2)}{p(\mathbf{g}|1)} \cong \mathbf{g}^t \left( \Sigma_1^{-1} - \Sigma_2^{-1} \right) \mathbf{g} \quad (4)$$

where  $p(\mathbf{g}|i)$  is the probability density function for the data given the  $i$ th condition. A larger value for this scalar variable indicates a greater likelihood for condition 2 than condition 1. An observer who adopts this strategy uses the data to maximize the area under the receiver operating characteristic (ROC) curve, and thus is referred to as the *ideal observer* for each of the discrimination tasks.

For shift-invariant  $\mathbf{H}$  approximated by a circulant matrix, we can separate each matrix into stationary and non-stationary components, respectively  $\Sigma_i = \Sigma_0 + \Delta \Sigma_i$ . Expanding the matrices in a power series and truncating after the first term we have [14]

$$\Sigma_1^{-1} - \Sigma_2^{-1} \simeq \Sigma_0^{-1} (\Delta \Sigma_1 - \Delta \Sigma_2) \Sigma_0^{-1}. \quad (5)$$

Consequently the linear approximation to (4) is

$$\lambda(\mathbf{g}) \approx \mathbf{g}^t \Sigma_0^{-1} \mathbf{H}(\mathbf{S}_2 - \mathbf{S}_1) \mathbf{H}^t \Sigma_0^{-1} \mathbf{g}. \quad (6)$$

The factor  $\mathbf{H}^t \Sigma_0^{-1} \mathbf{g}$  indicates that a Wiener filter is applied to the echo signal. However the filter depends on the stationary covariance  $\Sigma_0$ , which does not contain task information (see the Appendix). We named this the *stationary* Wiener filter to differentiate it from the adaptive filter described in the next section. An envelope generated by first Wiener filtering RF data is referred as a WFB-mode image.

### E. Iterative Wiener Filtering

When the echo SNR varies significantly in the field of view, as it does for the high-contrast lesions in Task 5, a stationary filter  $\mathbf{H}^t \Sigma_0^{-1}$  cannot be tuned to all regions simultaneously. This section describes a filter that adapts to spatially-varying echo SNR conditions.

Instead of separating covariance matrix components into stationary and nonstationary components, we form *average* and *difference* components, respectively

$$\Sigma_a = \sigma_{\text{obj}}^2 \mathbf{H}(\mathbf{I} + \mathbf{S}_a) \mathbf{H}^t + \sigma_n^2 \mathbf{I} \quad \text{and} \quad \Delta \Sigma = \sigma_{\text{obj}}^2 \mathbf{H} \Delta \mathbf{S} \mathbf{H}^t \quad (7)$$

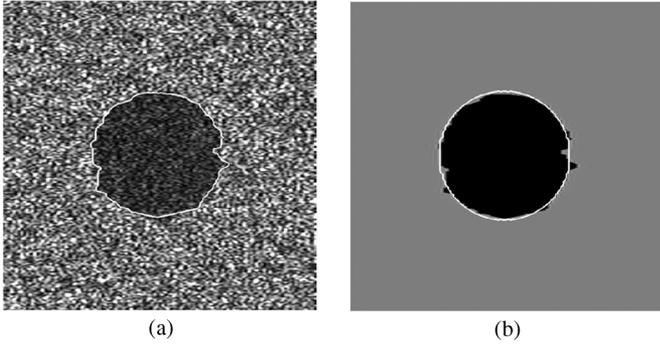


Fig. 2. (a) Segmentation of a WFB-mode image (linear scale). (b) Comparison of segmentation results (gray levels) to the echogenicity map  $\mathbf{S}$  (white ellipse).

where

$$\mathbf{S}_a = 0.5(\mathbf{S}_1 + \mathbf{S}_2) \quad \text{and} \quad \Delta\mathbf{S} = 0.5(\mathbf{S}_2 - \mathbf{S}_1).$$

As with the stationary filter, we expand the matrices in a power series and truncate after the first term to find

$$\Sigma_1^{-1} - \Sigma_2^{-1} \approx 2\Sigma_a^{-1}\Delta\Sigma\Sigma_a^{-1}. \quad (8)$$

The linear approximation to  $\lambda$  is now

$$\lambda(\mathbf{g}) \approx \mathbf{g}^t \Sigma_a^{-1} \mathbf{H} (\mathbf{S}_2 - \mathbf{S}_1) \mathbf{H}^t \Sigma_a^{-1} \mathbf{g}. \quad (9)$$

Equation (9) suggests that the filter  $\mathbf{H}^t \Sigma_a^{-1}$  be applied to the echo data, where the average covariance between two states is applied instead of the stationary background covariance. The advantage of this change is to allow the signal strength to vary significantly within any one image provided differences between compared images remain small. The task specificity is designed into this beamformer. A disadvantage of (9) is that  $\Sigma_a$  is the covariance matrix of a nonstationary process, so we cannot use Fourier techniques to quickly compute its inverse.

The power series approach may be applied to decompose  $\Sigma_a$  into stationary and non-stationary components,  $\Sigma_a = \Sigma_0 + \sigma_{\text{obj}}^2 \mathbf{H} \mathbf{S}_a \mathbf{H}^t$ , yielding an iterative formula for  $\sigma_{\text{obj}}^2 \mathbf{H}^t \Sigma_a^{-1} \mathbf{g}$  given by

$$\mathbf{q}_{i+1} = -\sigma_{\text{obj}}^2 \mathbf{H}^t \Sigma_0^{-1} \mathbf{H} \mathbf{S}_a \mathbf{q}_i$$

and

$$\mathbf{p}_{i+1} = \mathbf{p}_i + \mathbf{q}_{i+1}. \quad (10)$$

The iterative scheme is initialized by  $\mathbf{q}_0 = \mathbf{p}_0 = \sigma_{\text{obj}}^2 \mathbf{H}^t \Sigma_0^{-1} \mathbf{g}$ . Equation (10) begins with the stationary Wiener filter, and iteratively converges to  $\mathbf{p}_{i+1} = \sigma_{\text{obj}}^2 \mathbf{H}^t \Sigma_a^{-1} \mathbf{g}$  using the power series inverse approximation for  $\Sigma_a$ . We refer to the result as the *iterative Wiener filter* and to the images as *IWFB-mode images*. Twenty iterations were required to achieve an incremental change  $< 0.01\%$ . Thus IWFB-mode images can adapt to the task

as specified by  $\mathbf{S}_a$  provided that  $\mathbf{S}_a$  is known.  $\mathbf{S}_1$  and  $\mathbf{S}_2$  are estimated through image segmentation as we now describe.

#### F. Segmentation

We adopted a segmentation algorithm that makes use of a Markov random field (MRF) model. The core of the method is the adaptive clustering algorithm proposed by Papas [21] as applied to *B-mode* image data to be segmented into two regions: lesion and background. For  $b[m]$  as the pixel value at location  $m$ , determine the region  $r[m] = \ell$ , where  $\ell$  indicates lesion or background. We employ the maximum a posteriori (MAP) rule, which seeks to iteratively maximize the posterior probability density,  $p(r|b)$ . From Bayes' theorem,  $p(r|b)$  is proportional to the product of the prior density of the region process,  $p(r)$ , and the likelihood density for finding gray level  $b$  given region  $r$ ,  $p(b|r)$

$$p(r|b) = C p(b|r) p(r) \quad (11)$$

where  $C$  is a normalization constant. Applying the MRF model, we assume prior  $p(r)$  is given by a Gibbs density [22], [23], while likelihood  $p(b|r)$  has a Gaussian distribution with mean  $\mu_\ell$  and variance  $\sigma_\ell^2$ . The algorithm “adapts” to the image data by updating  $\mu_\ell$  and  $\sigma_\ell^2$  after each iteration.

Ashton *et al.* [24] successfully applied the algorithm to segment ultrasound *B-mode* images despite the Rayleigh pdf of *B-mode* speckles. They decomposed the image data into multiple resolution layers, and then invoked the central limit theorem to assume pixel values at the lowest level were described by a Gaussian pdf. Segmentation was applied at successively higher resolution layers through a computationally intensive process.

Our approach is to segment the logarithm of WFB-mode pixels instead of *B-mode* images. Because Wiener filtering reduces pixel correlation, we find log WFB-mode pixels are well represented by a Gaussian pdf, allowing  $\mathbf{b}$  to be quickly and effectively segmented. Fig. 2 shows an example where a two-region WFB-mode image was segmented. Comparing the segmented image to the original echogenicity map we can observe the segmentation errors that will reduce the effectiveness of IWFB-mode processing. Segmentation errors are primarily a problem for tasks involving discrimination of lesion boundary features.

The pixel whitening effects of the Wiener filter allow us to use WFB-mode pixels to coarsely estimate object background variance  $\hat{\sigma}_{\text{obj}}^2$ .  $\mathbf{S}_i$  is found by measuring the mean-square image value inside the lesion and dividing it by  $\hat{\sigma}_{\text{obj}}^2$ .

Fig. 3 shows example images processed three different ways. The effects of iterative Wiener filtering on discriminability are subtle, yet the changes in contrast resolution have a significant effect on human observer performance for Task 5 as shown below. First, we briefly summarize observer metrics.

#### G. Assessing Performance: Human and Ideal Observers

The visual discrimination performance of trained *human observers* was evaluated using the two-alternative forced-choice (2AFC) method under signal-known-exactly conditions [25],

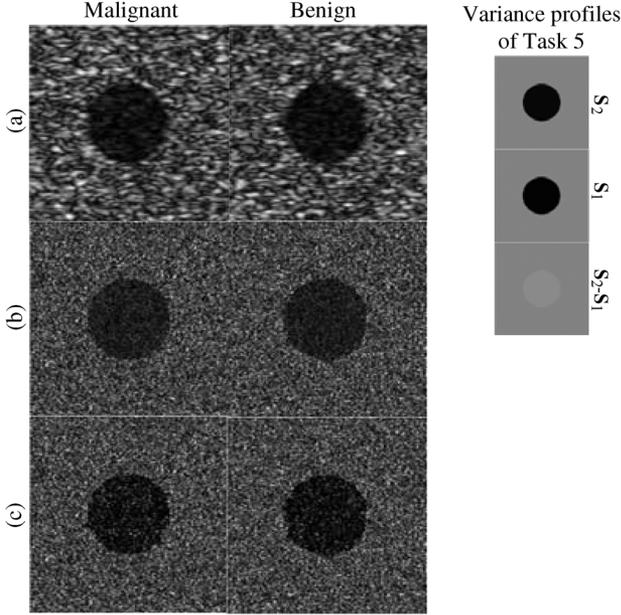


Fig. 3. Examples of (a) *B*-mode, (b) WFB-mode, and (c) IWFB-mode image pairs used for 2AFC studies of Task 5 (linear scale). The right side is illustrated for the variance profiles for Task 5 including  $S_2$  (malignant),  $S_1$  (benign), and the difference  $S_2 - S_1$ .

[26]. Observers viewed two simulated images during the  $j^{\text{th}}$  trial of a given task,  $\mathbf{b}_{i,j}$ . Images were placed side-by-side, one with a “benign” lesion feature,  $i = 1$ , and the other with the corresponding “malignant” feature,  $i = 2$ . Observers were informed of all feature parameters such as amplitude and location. After a training period, observers viewed 400 randomized image pairs per study and were asked to identify the simulated image with the malignant feature in each pair to measure the proportion of correct responses,  $P_C$ . The correctness of each response was immediately indicated to each observer. The echo SNR applied was the value computed experimentally using phantoms. The ratio of signal and noise power integrated over the signal bandwidth was 32 dB.

There were five observers for each task. Each participated in three sequential studies involving the same RF data to obtain  $P_C$  estimates from viewing *B*-mode, WFB-mode, and IWFB-mode images. Studies were performed for the five tasks involving different features of breast cancer sonography.

The visual discrimination performance of the ideal observer was evaluated using simulated RF data  $\mathbf{g}_{i,j}$  drawn from the same distributions and study conditions applied to humans. For each study (five tasks  $\times$  three processing schemes), 2000 pairs of RF data were processed to compute test statistics,  $\lambda(\mathbf{g}_{i,j})$ . If  $\mathbf{g}_{i,j}$  is an RF frame for condition  $i$ , then the score for trial  $j$  is

$$o_j = \text{step}(\lambda(\mathbf{g}_{2,j}) - \lambda(\mathbf{g}_{1,j})).$$

The step function equals 1 for positive arguments and 0 for negative arguments and the net scores yield  $P_C$  estimates.

Further details of human and machine observer experimental design, including observer training, viewer conditions, Monte

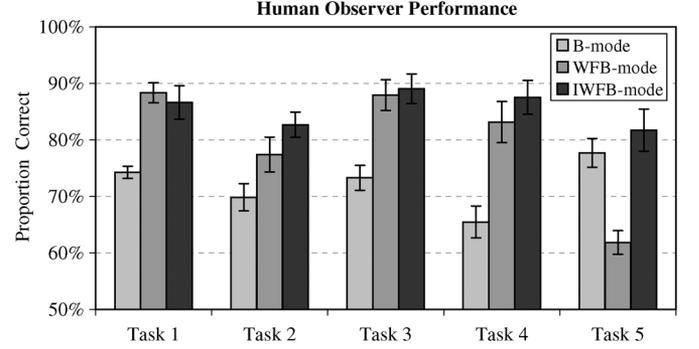


Fig. 4. Proportion correct for human observer studies involving three forms of signal processing as illustrated in Fig. 3 and the five diagnostic tasks described in Section II-C. Error bars indicate standard errors.

Carlo methods, and the object contrast factor, can be found in [14].

The efficiency of human observer discrimination with respect to the ideal observer is [27]

$$\eta = (C_I/C_H)^2 \quad (12)$$

where  $C$  is the contrast factor for a feature and subscripts  $H$  and  $I$  refer to values obtained from human and ideal observers. Pilot studies were used to find contrast settings between 0.7 and 0.8 in order to avoid ceiling and floor effects (since  $P_C$  is always between 0.5 and 1) and to remain in the region of high precision for 2AFC experiments [26].

We can gain additional insights about information flow through the image-formation/decision-making processes by also measuring the *Smith–Wagner observer* response. Smith *et al.* [13] described a test statistic acting on *B*-mode image data as

$$\lambda_{\text{SW}}(\mathbf{b}) = \mathbf{b}^t(\mathbf{S}_2 - \mathbf{S}_1)\mathbf{b} \quad (13)$$

and showed it was ideal for low-contrast lesion detection, our task 1. In the following application, we assume the Smith–Wagner (SW) observer is a reasonable approximation to the ideal observer acting on *B*-mode images for the other tasks. Measurements of  $\lambda_{\text{SW}}$  yield  $P_C$  values, so (12) can be expanded to

$$\eta = \left(\frac{C_I}{C_{\text{SW}}}\right)^2 \left(\frac{C_{\text{SW}}}{C_H}\right)^2 = \eta_{\text{SW}} \times \eta_{\text{H|SW}} \quad (14)$$

$\eta_{\text{SW}}$  may be interpreted as the efficiency by which information is transferred from the RF data into image data, while  $\eta_{\text{H|SW}}$  is the efficiency by which humans extract information from images. Separating the efficiency factors allows us to track sources of diagnostic information loss.

### III. OBSERVER PERFORMANCE MEASUREMENTS

Human observer studies were conducted at the Vision and Image Understanding Lab, UCSB, using methods described previously [14]. Five observers participated in 15 studies each after successfully completing a training regimen. Each study

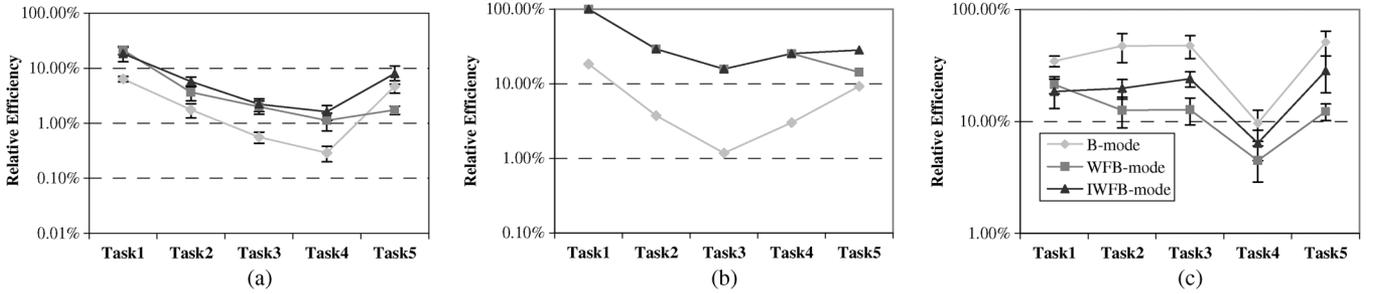


Fig. 5. (a) Human observer efficiencies for three beamformers,  $\eta$  from (14). (b) SW observer efficiency,  $\eta_{SW}$ . (c) Human efficiency relative to the SW observer,  $\eta_{H|SW}$ . Values in (a) equal the product of corresponding values from (b) and (c). Note that the ordinate scaling of relative efficiency axis is changed among the figures.

involved one observer viewing 400 image pairs from which the proportion correct,  $P_C$ , was found. Results averaged over five observers are charted in Fig. 4 for five tasks and three beamformers.

Examples of *B*-mode images, WFB-mode images, and IWFB-mode images used in human observer studies are shown in Fig. 3 where the images are generated under the shift-invariant assumption for the impulse response. The function is generated with fixed-focused for both transmitting and receiving, and without any apodization as we described in Section II-B. Although the image simulations of Fig. 4 involve a more realistic impulse response than that of our previous study [14], the rank order of  $P_C$  values for *B*-mode and WFB-mode images was preserved. Viewers of IWFB-mode images performed about the same as those viewing WFB-mode images for Tasks 1-4. However performance was significantly improved using IWFB-mode images for Task 5. Iterative filtering is able to adapt to the local statistics of the signal in a way that a stationary filter cannot. Segmentation facilitates the inclusion of task-specific information to enhance performance.

Data from Fig. 4 were applied to (14) to find the discrimination efficiencies plotted in Fig. 5. From Fig. 5(a), we find that humans are most efficient at low-contrast lesion detection (Task 1) and least efficient at discriminating smooth from spiculated lesion boundaries (Task 4). Fig. 5(b) shows that the greatest loss of information occurs during the process of converting RF signals into images, nearly an order of magnitude, and that spatial filtering is most effective at preserving information at this stage of the process. Fig. 5(c) suggests that humans are most effective at extracting task information from *B*-mode images and least effective using WFB-mode images. Human accessibility to information is between these two levels when IWFB-mode images are viewed. We have no explanation for the greater accessibility for *B*-mode images, however, we note that the S-W observer was designed to address Task 1. Its behavior for other tasks has not been studied.

It is interesting to see that the efficiency of transferring information from RF signals to image data is lowest for Task 3, Fig. 5(b), where viewers are challenged to discriminate soft and sharp lesion boundaries. Yet viewers have the least accessibility to information about lesion spiculation, i.e., Task 4 in Fig. 5(c). Recall that the point-wise product of curves in Fig. 5(b) and (c) give the net human visual efficiency summarized in Fig. 5(a).

Iterative Wiener filter performance is limited primarily by the accuracy of the segmentation algorithm that identifies image features (see Fig. 2). In situations where exact details of lesion features are known, i.e., we have  $\mathbf{S}_i$  *a priori*, the iterative Wiener filter provides near perfect ideal-observer performance for all five tasks, i.e.,  $P_C \rightarrow 1$ , and thus segmentation limits the transfer of diagnostic information from object  $\mathbf{f}$  into RF data  $\mathbf{g}$ . We were able to apply a two-level segmentation algorithm because for all tasks, except Task 3,  $\mathbf{S}_i$  involved just two levels of object scattering. Clinical studies where regional scattering heterogeneity is greater will require segmentation of image data with more than two intensity levels.

#### IV. EXPERIMENTAL IMPLEMENTATION

##### A. Line Spread Function

Filtering was applied to echo data acquired from a Siemens Sono-line Antares system with a VF10-5 transducer. System parameters were nominally the same as those applied to echo simulations from Section II-B. Data are recorded without applying time-gain-compensation. We still use fixed-focused for both transmitting and receiving at 40 mm and no apodization. The configuration helps show the dependence of the speckles upon the image depth and the improvement when Wiener filtering. The system was used to image a cyst phantom (Model #539, ATS Laboratories, Bridgeport, CT). The manufacturer-reported speed of sound in the phantom is 1450 m/s and the attenuation coefficient slope is  $0.5 \text{ dB cm}^{-1} \text{ MHz}^{-1}$ . The DS beamformed echo SNR in the phantom was measured to be 32 dB near the 40-mm focal length.

The Wiener filter was developed using the measured line-spread function (*lsf*), which is the pulse-echo impulse response integrated over the elevational axis. The line scatterers are of 0.12-mm-diameter nylon monofilament. Since phantom objects are 2-D (line and cylinder inclusions oriented to give point and circular targets in the scan plane), the *lsf* is most appropriate to use in the system matrix  $\mathbf{H}$  of the Wiener filter.

The *lsf* varied with depth as shown in the *B*-mode image of Fig. 6(a), which is used to investigate the shift-variant impulse responses for Wiener filtering. For the shift-invariant impulse response,  $\mathbf{H}$  is assumed as a circulant matrix for experiment data and composed from a single *lsf* recorded from a line scatterer positioned at the 40-mm focal length. For the shift-variant impulse

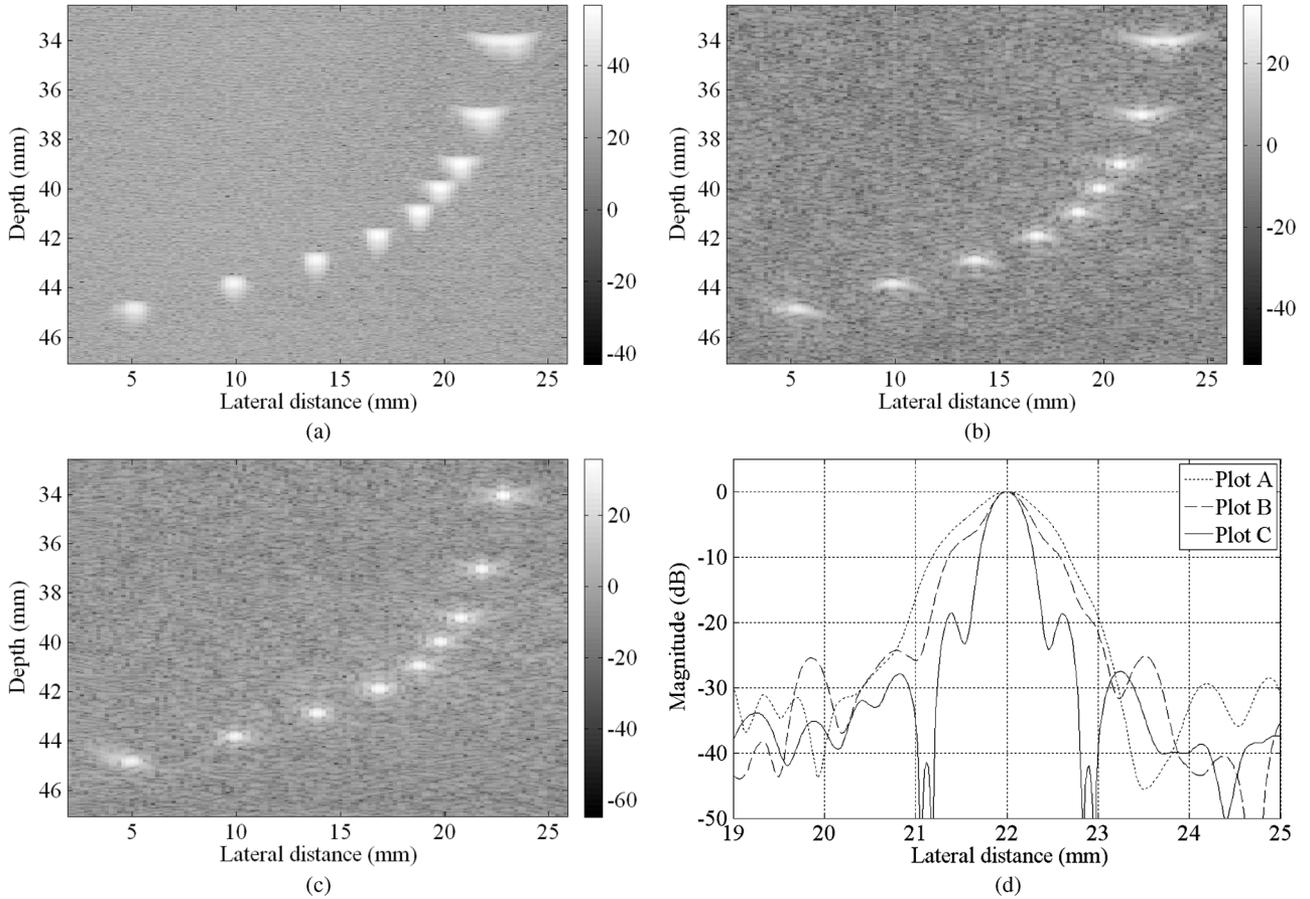


Fig. 6. Log-compressed images for a line target phantom measured using the Siemens Sono-line Antares system at 7 MHz. (a)  $B$ -mode image. (b) WFB-mode image with shift-invariant impulse response assumed. (c) WFB-mode image with shift-variant impulse response assumed and the multi- $lsf$  filter applied. (d) The plots of three lateral lines (normalized) across the target at the depth of 37 mm, in which plot A is for the  $B$ -mode image in (a), plot B is for the WFB-mode image in (b) with shift-invariant impulse response, and plot C is for the WFB-mode image in (c) with shift-variant impulse response.

response,  $\mathbf{H}$  is composed from five  $lsf$ 's recorded at regular intervals between 34 and 45 mm depths (the depth of *isoplanatic* patches are 2 mm, see the Appendix). Applying the method described in the Appendix below, a Wiener filter for a shift-variant system was formed. To reduce  $lsf$  noise to negligible levels for filter development, we recorded and time-averaged 1000 frames from stationary line scatterers.

Fig. 6(b) and (c) shows WFB-mode images for filters made assuming shift-invariant and shift-variant systems, respectively. While the impulse response of the  $B$ -mode image in Fig. 6(a) is strongly depth-dependent, the WFB-mode image in Fig. 6(b) shows improved spatial resolution near the 40 mm focal length. However, the shift-variant filter used in the image of Fig. 6(c) demonstrates a more uniform and improved spatial resolution. Nevertheless, deconvolution remains incomplete because of low-level side-lobe energy.

### B. Cyst Phantom Experiment

The same two Wiener filters were applied to an 8-mm-diameter, anechoic, circular target in the phantom that was positioned at a depth of 40 mm. The  $B$ -mode image is shown in the upper-left corner of Fig. 7, where speckle is large and nonuniform over the plane. Average speckle size is reduced in the WFB-mode image (upper right) filtered by only one  $lsf$

recorded at 40 mm depth. However, speckle is not spatially uniform, and the cyst boundary is distorted because the RF data is filtered with an unmatched  $lsf$ . The WFB-mode image generated with multiple shift-variant  $lsf$ 's (bottom left) has a more circular boundary and uniform small speckle, but at the cost of increased computation (see Appendix). Finally, the corresponding IWFB-mode image (bottom right) has enhanced contrast and is able to most clearly represent a cyst-like target. Segmentation errors tend to erode the margin and suggest a more complicated boundary than the simple circle we know is present.

## V. DISCUSSION AND CONCLUSION

The ideal observer approximation is extended to improve visual discrimination for high-contrast features by introducing an iterative Wiener filter. IWFB-mode images decorrelate speckle, as do WFB-mode images, but are able to better preserve contrast resolution for contrast-limited tasks, e.g., Task 5. The human-observer performance studies show that IWF provides the same high discrimination level as the stationary WF for a low-contrast large-area detection task (Task 1), and three boundary discrimination tasks (Tasks 2-4). However, IWF significantly improves efficiency for a high-contrast large-area task (Task 5), where echo nonstationarity from object heterogeneity degrades the stationary WF. The improvement in performance comes at

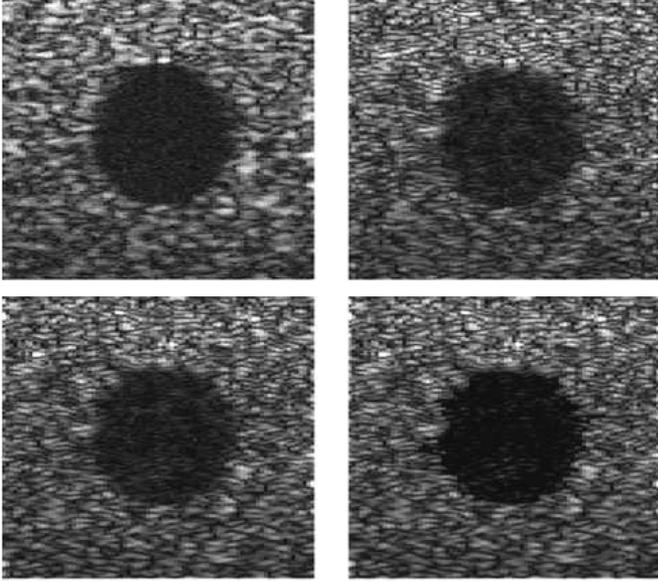


Fig. 7. Images of a cyst-like, 8-mm circular target centered at a 40 mm depth for a commercial system (linear scale). Standard *B*-mode image (upper-left), WFB-mode image with shift invariant  $\mathbf{H}$  (upper-right), WFB-mode image with shift variant  $\mathbf{H}$  (lower-left), and IWFB-mode image (lower-right).

the cost of approximately five-fold increase in computational load (IWFB-modes require 20 s on a dual-core PC, 2.13 GHz Processor, 2G RAM).

The deconvolution methods improve feature discriminations because the density of independent image samples available to observers increases as the average speckle area is reduced. The additional information is helpful for contrast-limited large-area Tasks 1 and 5 as well as those involving resolution-limited edge discrimination, Tasks 2–4. The Wiener filter produces significant improvements when the impulse response is known exactly and the image data are from a wide-sense stationary distribution. It fails when the assumptions are violated, as occurred with the Wiener filter in Task 5, and yet is robust enough for use with commercial systems. Performance improves across all five tasks for IWFB-mode images because we added task-specific information in filter development when it was necessary. We strive to develop beamformers that are fast and robust across the wide spectrum of clinical features, and yet can adapt when necessary to special conditions. The ideal observer approach provides a framework for that development.

The use of realistic imaging pulses in image simulations, as provided by Field-II, resulted in a generally lower observer efficiency when compared to results from sonograms simulated with Gaussian pulses (compare Fig. 5(a) above with Fig. 7(a) from [14]). The reduction results from lower feature contrast generated using the less-compact Field II pulses.

Human visual discrimination efficiency for the five lesion features considered is less than 10% for *B*-mode imaging. Spatial filtering was found to improve the transfer of object information into the image data, but it reduces somewhat the ability of humans to access the information. Thus there is a potential role for image processing of the final envelope image to increase accessibility. Note that the Smith–Wagner observer [13] was devel-

oped as the ideal observer for Task 1 but is unlikely to perform optimally for other tasks.

Task performance of spatial-filter beamformers is improved significantly by including any shift variance in the pulse-echo impulse response in the filter as seen in the phantom experiments. The greatest challenge when applying this method in the clinical environment is to estimate accurately the pulse-echo impulse response for the Wiener filter. However, it is very difficult to measure this impulse response function for commercial systems, because it is affected by phase aberrations, imperfection of transducers, and undesirable artifacts inside the systems. In this research, we found that a line-spread function can be used to develop Wiener filters in place of the impulse response when imaging cylindrical objects. Accurate estimation of point-spread functions throughout the field will be needed to improve clinical imaging. Loss of visual discrimination from an inaccurate point-spread function is also an interesting topic and motivates further study.

For any pulse-echo experiment, the most effective processing—from perspectives of both an optimal Bayesian observer (information transfer) and from psychophysically measured human observer performance—requires detailed knowledge of the system impulse response to decorrelate RF signals and thereby reduce the effects of speckle in the resulting image.

## APPENDIX A

### SHIFT-VARIANT IMPULSE RESPONSE

This appendix provides an expansion of the shift-invariant methodology to include depth-varying impulse responses in isoplanatic regions. We divide the field into small regions (patches) along the beam axis within which the impulse response is shift invariant. The expression for RF data becomes

$$\mathbf{g} = \sum_{j=1}^k \mathbf{H}_j \mathbf{f}_j + \mathbf{n} = \mathbf{H} \mathbf{f} + \mathbf{n} \quad (\text{A-1})$$

where  $k$  is the number of patches along the depth of field,  $\mathbf{H}_j$  is a block Toeplitz system matrix constructed from the  $j$ th impulse response  $h_j$ , and  $\mathbf{f}_j = \mathbf{E}_j \mathbf{f}$  is the scattering vector for the object in patch  $j$ .  $\mathbf{E}_j$  is a diagonal matrix of 1s corresponding to patch  $j$  such that

$$\sum_{j=1}^k \mathbf{E}_j = \mathbf{I} \quad (\text{A-2})$$

for identity matrix  $\mathbf{I}$ . Also  $\mathbf{H} = \sum_{j=1}^k \mathbf{H}_j \mathbf{E}_j$ . Although  $\mathbf{H}_j$  are block Toeplitz matrices,  $\mathbf{H}$  is not.

The covariance matrices for  $\mathbf{g}$  in (2) becomes

$$\boldsymbol{\Sigma}_i = \sigma_{\text{obj}}^2 \sum_{j=1}^k \mathbf{H}_j \mathbf{E}_j (\mathbf{I} + \mathbf{S}_i) \mathbf{E}_j^t \mathbf{H}_j^t + \sigma_{\text{n}}^2 \mathbf{I} \quad (\text{A-3})$$

with the decision variable given in (4). Separating the covariance matrices as into background and task-specified components, we

obtain, respectively

$$\begin{aligned}\Sigma_0 &= \sigma_{\text{obj}}^2 \sum_{j=1}^k \mathbf{H}_j \mathbf{E}_j \mathbf{E}_j^t \mathbf{H}_j^t + \sigma_n^2 \mathbf{I} \\ \Delta \Sigma_i &= \sigma_{\text{obj}}^2 \sum_{j=1}^k \mathbf{H}_j \mathbf{E}_j \mathbf{S}_i \mathbf{E}_j^t \mathbf{H}_j^t.\end{aligned}\quad (\text{A-4})$$

Truncating the power series expansion after the first term, we obtain the linear approximation

$$\lambda(\mathbf{g}) \cong \frac{1}{2} \mathbf{g}^t \Sigma_0^{-1} (\Delta \Sigma_2 - \Delta \Sigma_1) \Sigma_0^{-1} \mathbf{g} \quad (\text{A-5})$$

where

$$\Delta \Sigma_2 - \Delta \Sigma_1 = \sigma_{\text{obj}}^2 \sum_{j=1}^k \mathbf{H}_j \mathbf{E}_j (\mathbf{S}_2 - \mathbf{S}_1) \mathbf{E}_j^t \mathbf{H}_j^t. \quad (\text{A-6})$$

Noting that for  $j \neq l$ ,  $\mathbf{E}_j (\mathbf{S}_2 - \mathbf{S}_1) \mathbf{E}_j^t = 0$ , and (A-6) becomes

$$\begin{aligned}\Delta \Sigma_2 - \Delta \Sigma_1 &= \sigma_{\text{obj}}^2 \left( \sum_{j=1}^k \mathbf{H}_j \mathbf{E}_j \right) (\mathbf{S}_2 - \mathbf{S}_1) \left( \sum_{j=1}^k \mathbf{E}_j^t \mathbf{H}_j^t \right) \\ &= \sigma_{\text{obj}}^2 \mathbf{H} (\mathbf{S}_2 - \mathbf{S}_1) \mathbf{H}^t.\end{aligned}\quad (\text{A-7})$$

Substituting (A-7) into (A-5) we obtain

$$\lambda(\mathbf{g}) \approx \frac{\sigma_{\text{obj}}^2}{2} \mathbf{g}^t \Sigma_0^{-1} \mathbf{H} (\mathbf{S}_2 - \mathbf{S}_1) \mathbf{H}^t \Sigma_0^{-1} \mathbf{g}. \quad (\text{A-8})$$

Thus for the Wiener filter  $\mathbf{H}^t \Sigma_0^{-1} \mathbf{g}$

$$\mathbf{H} = \sum_{j=1}^k \mathbf{H}_j \mathbf{E}_j \quad (\text{A-9})$$

and

$$\Sigma_0 = \sigma_{\text{obj}}^2 \sum_{j=1}^k \mathbf{H}_j \mathbf{E}_j \mathbf{E}_j^t \mathbf{H}_j^t + \sigma_n^2 \mathbf{I}. \quad (\text{A-10})$$

Since  $\mathbf{E}_j \mathbf{E}_l^t = 0$  for  $j \neq l$ ,

$$\Sigma_0 = \sigma_{\text{obj}}^2 \mathbf{H} \mathbf{H}^t + \sigma_n^2 \mathbf{I}. \quad (\text{A-11})$$

The Wiener filter for a shift-variant system has the form of a shift-invariant system but is composed of a partial sum of block-Toeplitz matrices, and  $\Sigma_0$  is the covariance of a nonstationary process. Consequently Fourier techniques cannot be applied to calculate inverse  $\Sigma_0$ .

To implement the filter, we note that multiplication of the covariance inverse by vector  $\mathbf{y}$  can be found by solving the linear equation  $\mathbf{y} = \Sigma_0 \mathbf{x}$ . Since  $\Sigma_0$  is a symmetric matrix, the equation can be solved by using the *gradient conjugate* approach [28]. Thus the Wiener filter can be derived under more realistic conditions at the cost of computation time.

## REFERENCES

- [1] K. E. Thomenius, "Evolution of ultrasound beamformers," in *Proc. IEEE Ultrason. Symp.*, 1996, pp. 1615–1622.
- [2] B. A. J. Angelsen, *Ultrasound Imaging: Waves, Signals, and Signal Processing*. Trondheim, Norway: Emantec, 2000, vol. I, II.
- [3] Z. Wang, J. Li, and R. Wu, "Time-delay and time-reversal-based robust Capon beamformers for ultrasound imaging," *IEEE Trans. Med. Imag.*, vol. 24, no. 10, pp. 1308–1322, Oct. 2005.
- [4] J.-F. Synnevåg, A. Austeng, and S. Holm, "Adaptive beamforming applied to medical ultrasound imaging," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 54, no. 8, pp. 1606–1613, Aug. 2007.
- [5] F. Lingvall and T. Olofsson, "On time-domain model-based ultrasonic array imaging," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 54, no. 8, pp. 1623–1633, Aug. 2007.
- [6] R. Lavarello, F. Kamalabadi, and W. D. O'Brien, Jr., "A regularized inverse approach to ultrasonic pulse-echo imaging," *IEEE Trans. Med. Imag.*, vol. 25, no. 6, pp. 712–722, Jun. 2006.
- [7] J. Li and P. Stoica, *Robust Adaptive Beamforming*. Hoboken, NJ: Wiley Interscience, 2005.
- [8] D. A. Guenther and W. F. Walker, "Robust finite impulse response beamforming applied to medical ultrasound," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 56, no. 6, pp. 1168–1188, Jun. 2009.
- [9] R. F. Wagner and D. G. Brown, "Unified SNR analysis of medical imaging systems," *Phys. Med. Biol.*, vol. 30, no. 6, pp. 489–518, 1985.
- [10] H. H. Barrett, "Objective assessment of image quality: Effects of quantum noise and object variability," *J. Opt. Soc. Am. A*, vol. 7, no. 7, pp. 1266–1278, 1990.
- [11] 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, no. 6, pp. 1520–1535, 1998.
- [12] P. Khurd, B. Liu, and G. Gindi, "Ideal AFROC and FROC observers," *IEEE Trans. Med. Imag.*, vol. 29, no. 2, pp. 375–386, Feb. 2010.
- [13] 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, no. 3, pp. 164–173, May 1983.
- [14] 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 with ultrasound," *IEEE Trans. Med. Imag.*, vol. 25, no. 2, pp. 198–209, Feb. 2006.
- [15] R. J. Zemp, C. K. Abbey, and M. F. Insana, "Linear system model for ultrasonic imaging: application to signal statistics," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 50, no. 6, pp. 642–654, Jun. 2003.
- [16] H. H. Barrett and K. J. Myers, *Foundations of Image Science*. Hoboken, NJ: Wiley, 2004.
- [17] 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, Mar. 1992.
- [18] J. A. Jensen, "Field: A program for simulating ultrasound systems," *Medical and Biological Engineering and Computing*, vol. 34, no. Suppl. 1, pt. 1, pp. 351–353, 1996.
- [19] American College of Radiology, *Breast Imaging Report and Data System Atlas*. Reston, VA, 2003.
- [20] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*. New York, NY: Wiley, 1968, pt. I.

- [21] T. N. Pappas, "An adaptive clustering algorithm for image segmentation," *IEEE Trans. Signal Process.*, vol. 40, no. 4, pp. 901–914, Apr. 1992.
- [22] B. J. Besag, "On the statistical analysis of dirty pictures," *J. R. Stat. Soc.: Series B*, vol. 48, no. 3, pp. 259–302, 1986.
- [23] J. S. Geman and D. Geman, "Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6, no. 6, pp. 721–741, 1984.
- [24] E. A. Ashton and K. J. Parker, "Multiple resolution Bayesian segmentation of ultrasound images," *Ultrason. Imag.*, vol. 17, no. 4, pp. 291–304, 1995.
- [25] D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics*. New York: Wiley, 1966.
- [26] A. E. Burgess, "Comparison of receiver operating characteristic and forced choice observer performance measurement methods," *Med. Phys.*, vol. 22, no. 5, pp. 643–655, 1995.
- [27] W. P. Tanner and T. G. Birdsall, "Definition of  $d'$  and  $\eta$  as psychophysical measures," in *Signal Detection and Recognition by Human Observer: Contemporary Readings*, J. A. Swets, Ed. New York: Wiley, 1964, pp. 147–163.
- [28] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: Johns Hopkins Univ. Press, 1996.