Observer Efficiency in Boundary Discrimination Tasks Related to Assessment of Breast Lesions with Ultrasound

Craig K. Abbey[*,1,2], Roger J. Zemp[3], Jie Liu[4], Karen K. Lindfors[5], and Michael F. Insana[1,4]

[1]Dept. of Biomedical Engineering, University of California, Davis, CA 95616
[2]Department of Psychology, University of California, Santa Barbara, CA 93106
[3]Department of Biomedical Engineering, Texas A&M University, College Station. TX 77843
[4]Department of Bioengineering, University of Illinois at Urbana-Champaign. IL 93103
[5]Department of Radiology, UC Davis Medical Center, Sacramento, CA,93103.

ABSTRACT

The statistical efficiency of human observers in diagnostic tasks is an important measure of how effectively task relevant information in the image is being utilized. Most efficiency studies have investigated efficiency in terms of contrast or size effects. In many cases, malignant lesions will have similar contrast to normal or benign objects, but can be distinguished by properties of their boundary. We investigate this issue in the framework of malignant/benign discrimination tasks for the breast with ultrasound. In order to identify effects in terms of specific features and to control for other effects such as aberration or specular reflections, we simulate the formation of beam-formed radio-frequency (RF) data. We consider three tasks related to lesion boundaries including boundary eccentricity, boundary sharpness, and detection of boundary spiculations. We also consider standard detection and contrast discrimination tasks. We find that human observers exhibit surprisingly low efficiency with respect to the Ideal observer acting on RF data in boundary discrimination tasks (0.08%-3.3%), and that efficiency of human observers is substantially increased by Wiener-filtering RF frame data. We also find a limitation in efficiency is the computation of an envelope image from the RF data recorded by the transducer. Approximations to the Ideal observer acting on the envelope images indicate that humans may be substantially more efficient (10%-75%) with respect to the envelope Ideal observers. Our work suggests that significant diagnostic information may be lost in standard envelope processing in the formation of ultrasonic images.

Keywords: Ideal observer, Observer efficiency, Boundary discrimination tasks, breast sonography

1. INTRODUCTION

Ultrasonic imaging is considered a valuable adjunct in screening mammography where it is used to discriminate malignant disease from benign lesions (fibroadenoma, etc.) and cysts[1-3]. In its adjunctive role, ultrasound images are acquired from a generally high contrast lesion identified from some other source (X-ray imaging, palpation, etc.). In these cases, the difficulty is discriminating features of malignant disease from a clearly visible lesion. Many of the features used by clinicians to perform malignant/benign discrimination are defined at the boundary of a high-contrast lesion, and we refer to these as boundary discrimination tasks.

There is currently much interest in improving ultrasonic imaging in the breast for this purpose and for evaluating the efficacy of ultrasound as a stand-alone screening modality[4], particularly in women with mammographically dense breasts[5,6]. In order to know where improvement will have the most impact in a diagnostic imaging modality, it is important to identify where diagnostic information is being lost. We have been pursuing this question from a task-based perspective in which we analyze various elements of the image formation process[7]. This approach utilizes statistical models of scattering in the image formation process to define optimal discriminants (referred to as Ideal observers in the context of imaging) at various stages. Since the performance of an Ideal observer can be thought of as a measure of task-relevant information content, this approach allows us to identify sources of information loss in the imaging chain. We can also evaluate the human observer relative to an Ideal observer to determine how well the diagnostic information is accessed by image readers. A limitation of the Ideal observer approach is that defining the optimal discriminant

[*] Corresponding Author. Email: abbey@psych.ucsb.edu. Phone: 805 893 3853. FAX: 805 893 4303

Medical Imaging 2006: Image Perception, Observer Performance, and Technology Assessment, edited by Yulei Jiang, Miguel P. Eckstein, Proc. of SPIE Vol. 6146, 61460O, (2006) · 1605-7422/06/$15 · doi: 10.1117/12.655731

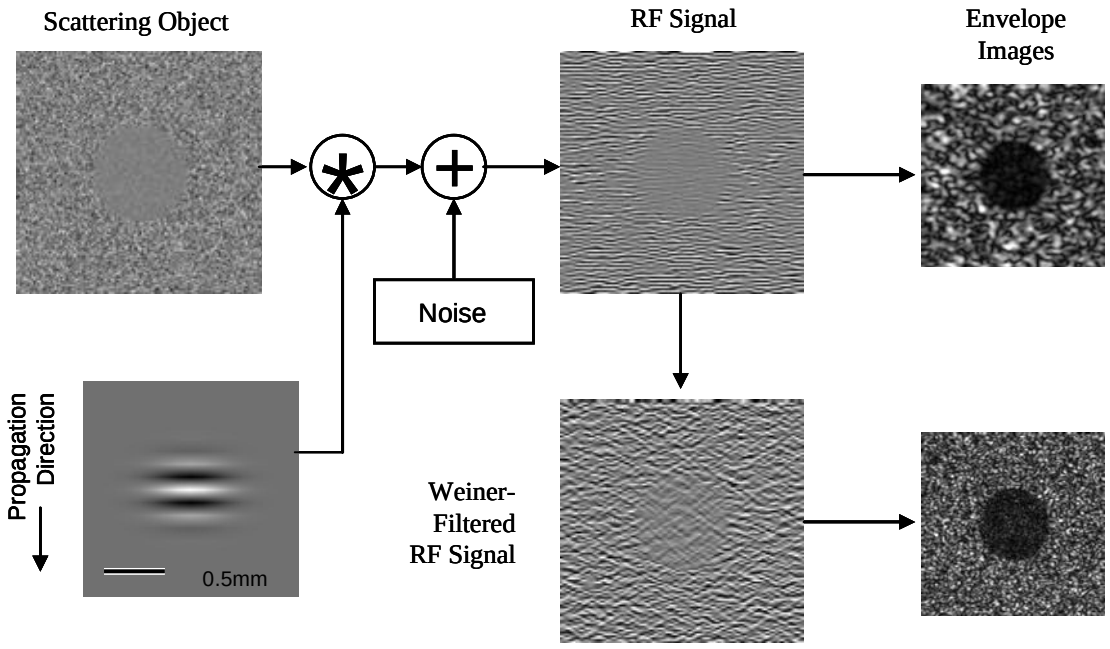

Figure 1. Diagram of the image formation model used in this work. The scattering object is convolved with the pulse profile and then acquisition noise is added to obtain a frame of radio frequency (RF) data. Under standard processing, the RF frame is demodulated, downsampled, and an envelope image is computed from it. We also investigate the effect of Wiener filtering the RF frame before computing the final envelope image.

function requires knowledge of the underlying probability distributions and hence simulated images are needed for implementation.

Our previous work has investigated the loss of diagnostic information in ultrasonic imaging of the breast going from the beam-formed radio-frequency (RF) frame data to the final envelope image read by a human observer[7]. The evaluation of the Ideal observer on envelope images, a critical step in the assessment of diagnostic information transfer, is made technically challenging by the fact that there is not an analytic form for the resulting probability distributions. The Ideal observer was approximated using results from the early work of Smith and Wagner[8]. Their approximation was tailored to large area low-contrast lesions, which is not necessarily appropriate for high contrast boundary discrimination tasks.

In this work we extend our previous results by considering the formation of RF frame data from a scattering object as another step in the imaging chain, and by investigating the approximation of the Ideal observer acting on envelope images from a limited feature set. Analysis of the formation of RF frame data allows us to evaluate the effectiveness of the pulse-acquisition scheme in capturing diagnostic information from the object. This step is of interest for exploring alternative schemes such as coded pulses and acquisition of angular scattering data[9,10]. The additional analysis of the Ideal observer acting on envelope data is of interest for refining the estimates of diagnostic information lost in the formation of an envelope image as opposed to lost by the reader.

2. THEORY AND METHODS

Model of ultrasonic imaging. Ultrasonic backscatter (B-mode) images are generally formed by transmitting acoustic waves from a transducer array and receiving signals from the backscattered waves. The image can be thought of as estimating a scattering magnitude. Figure 1 illustrates the model of ultrasonic image formation used in this work.

The scattering object is presumed to be an independent Gaussian process with the important features of the task encoded in the variance at a given point. While the true scattering object should be defined in principle as a continuously defined function, we find it more convenient to work with a sampled function with a sampling rate well beyond the bandwidth of the pulse. We can therefore define an object frame, **f**, of sampled object scattering values. A frame of beam-formed radio-frequency data is generated from the object frame by a 2D convolution with the pulse

profile and addition of post-acquisition noise. If we define the frame of RF data as **g**, then we can write the linear transformation from object to RF as

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n}, \tag{2.1}$$

where **H** represents the linear operation of convolution by the pulse profile, and **n** is the post-acquisition noise.

Describing the interaction of the pulse with the scattering object as a convolution allows us to use Fourier techniques to speed up the necessary calculations – and requires the assumption of an isoplanetic focal region, but Equation 1 shows that the formulation is general enough to include arbitrary linear functions. The pulse profile we use (pictured in Figure 1) is determined from phantom studies, and models the linear array of a Sonoline Elegra Ultrasound System (Siemens Medical Solutions, Mountain View, CA) operating at 7.2 MHz with 51.7% FWHM bandwidth and a peak signal power to noise power SNR of 45.5 dB.

We will investigate the effect of filtering RF frame data using a Wiener filter that incorporates the pulse profile and the variance of the object scattering function. The filter profile, *w*, is defined in Fourier domain by

$$\hat{w}[k,l] = \frac{\sigma_{\text{obj}}^2 \overline{\hat{h}[k,l]}}{\sigma_{\text{obj}}^2 \left|\hat{h}[k,l]\right|^2 + \sigma_n^2}, \tag{2.2}$$

where $\hat{h}$ is the discrete 2D Fourier transform of the pulse profile, σ_{obj}^2 is the putative variance of the object scattering function, and σ_{n}^2 is the variance of post-acquisition noise in the RF frame. As we shall see, there is not a single object variance that will describe the object functions we use, and we believe that this assumption of a constant object variance limits the effectiveness of the Wiener filter in some cases.

The final stage in the ultrasonic image formation process is the computation of an envelope image, which we denote **b**. The envelope computation can be implemented as a sequence of operations. The RF frame data (or Wiener filtered frame data) are multiplied by a complex mixing matrix and lowpass filtered to get the in-phase and quadrature components of the complex envelope signal. These are downsampled to the sampling rate of the lowpass filter (a factor of 2 in this work) and the complex magnitude is computed to get the envelope signal. Finally, the signal is quantized so that it will fit into the 8-bit display range of a monitor. We can write this sequence of operations as

$$\mathbf{b} = Q\left(\left|\mathbf{DLMg}\right|\right), \tag{2.3}$$

where **M** is a diagonal mixing matrix whose diagonal elements are a complex exponential with frequency equivalent to the pulse frequency, **L** implements a lowpass filter in axial direction, **D** implements downsampling in the axial direction, and Q is a quantization operator that assigns the resulting value in each element one of 256 intensity bins.

In our previous work[7], we used a simplified model of envelope computation that neglected the downsampling and quantization steps. In this case, the envelope image is given by

$$\mathbf{b} = \left|\mathbf{LMg}\right|. \tag{2.4}$$

We present this formula because part of the motivation for the present study is to assess the impact of these additional steps in the envelope computation.

2-class discrimination tasks. The Ideal observer analysis begins by presuming that and image comes from one of two possible classes, malignant or benign. Images from each class will differ slightly in the ensemble statistical properties – the object variance in this work, and the goal of the imaging process is to most accurately estimate the class from which an image comes.

We assume that the difference between malignant and benign objects is contained in differences between the object variance maps. Under Gaussian assumptions, the distribution of the object scattering functions in each class are given by

$$\begin{aligned} \text{Benign: } \mathbf{f} &\sim \text{MVN}\left(\mathbf{0}, \sigma_{\text{obj}}^2 \left(\mathbf{I} + \mathbf{S}_0\right)\right) \\ \text{Malignant: } \mathbf{f} &\sim \text{MVN}\left(\mathbf{0}, \sigma_{\text{obj}}^2 \left(\mathbf{I} + \mathbf{S}_1\right)\right) \end{aligned} \tag{2.5}$$

where MVN indicates the multivariate normal distribution with mean given by the first argument (the zero vector in this case since specular reflections are not considered), and covariance matrix given by the second component. In this case

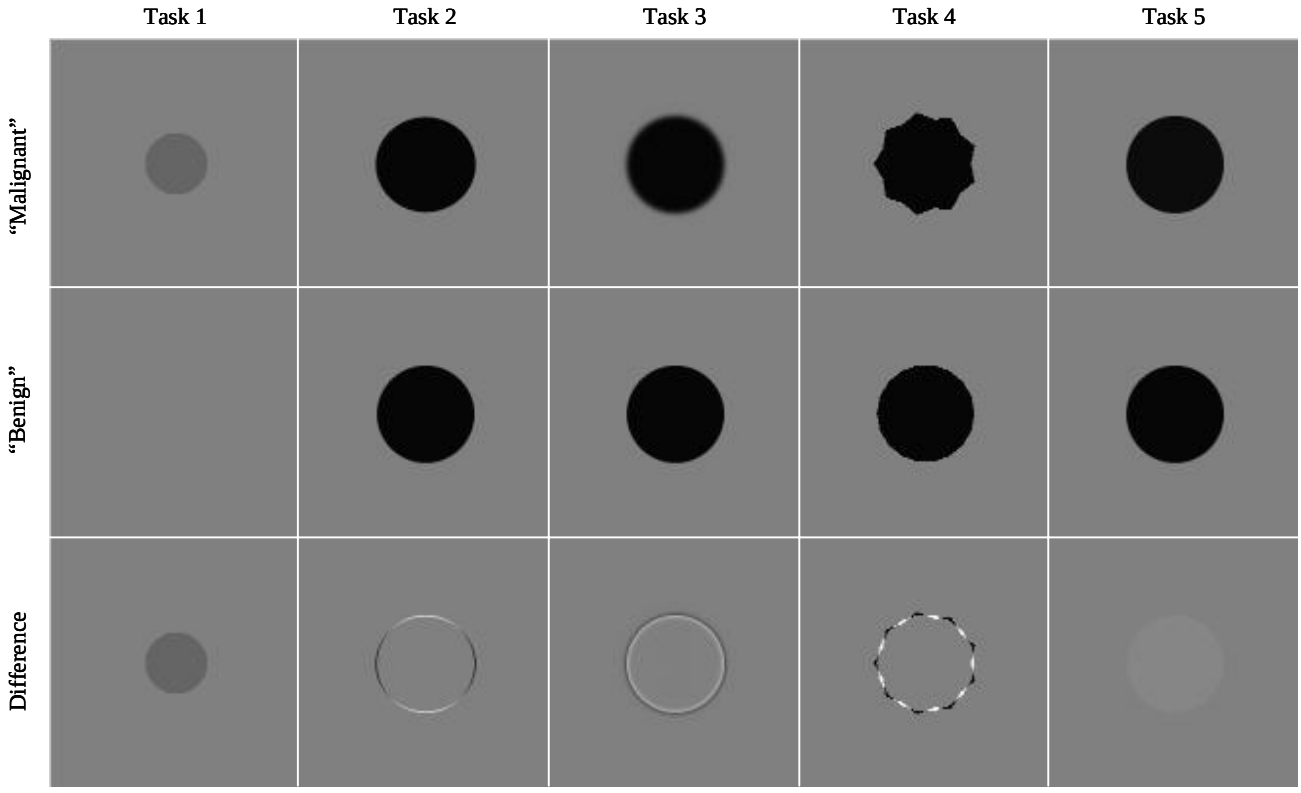

Figure 2. Variance profiles of "Malignant" and "Benign" scattering objects the 5 tasks. The third row is the difference in the variance profiles, which determines the object contrast for the task.

the covariance matrices in both classes are diagonal, and hence can be thought of as variance maps with differences from a uniform variance encoded in the $\mathbf{S}_0$ and $\mathbf{S}_1$ matrices (with diagonal elements that can range from -1 to ∞). Variance maps used in the five tasks investigated here can be found in Figure 2.

The $\mathbf{S}_0$ and $\mathbf{S}_1$ matrices also define the object contrast as we use it in this work. Because we will use a variety of different variance maps in the five tasks, it is helpful to have a common measure of the magnitude of differences between the malignant and benign maps. We define object contrast as the integrated difference in the variance maps,

$$C = \sum_{i=1}^{N} \left| \left[\mathbf{S}_1 - \mathbf{S}_0 \right]_{ii} \right| \Delta x \Delta y \,, \tag{2.6}$$

where Δx and Δy are the axial and lateral sampling lengths. If we consider Tasks 1 and 5 in Figure 2, we see that the difference in variance maps extends over the entire region of the lesion. We refer to these as large area tasks, and note that they include detection and contrast discrimination tasks. The variance maps of Tasks 2 – 4 are limited to an area immediately adjacent to the lesion boundary, and hence we refer to these as boundary discrimination tasks. These tasks discriminate malignant from benign lesions on the basis of lesion eccentricity (Task 2), the presence of an indistinct or "soft" boundary (Task 3), and spiculations on the boundary (Task 4). The selection of tasks was made on the basis of consultation with clinicians and from published atlases of malignant features of breast cancer in ultrasound images[1].

Equation 2.1 specifies the noisy linear operation that transforms the object to a frame of RF data. For the benign and malignant distributions defined in Equation 2.5 the corresponding RF frame distributions will be given by

$$\begin{aligned} \text{Benign: } \mathbf{g} &\sim \text{MVN}\left(\mathbf{0}, \boldsymbol{\Sigma}_0\right) \\ \text{Malignant: } \mathbf{g} &\sim \text{MVN}\left(\mathbf{0}, \boldsymbol{\Sigma}_1\right), \end{aligned} \tag{2.7}$$

where

$$\boldsymbol{\Sigma}_i = \sigma_{\text{obj}}^2 \mathbf{H}^t \left(\mathbf{I} + \mathbf{S}_i\right) \mathbf{H} + \sigma_{\text{n}}^2 \mathbf{I}, \quad i = 0,1. \tag{2.8}$$

Note that the task differences are still encoded in the covariance matrix, but this covaiance is no longer diagonal, and includes effects of the pulse profile and electronic noise.

Because of the nonlinear magnitude and quantization operations used in computing an envelope image, the class distributions are unknown. This limits our ability to analyze the Ideal observer acting on envelope images to approximate methods as we shall see below.

Ideal observers. We will implement Ideal observers (or approximations to them) to perform the various discrimination tasks at the level of the object, the RF frame and the envelope image. The Ideal observers are implemented by defining the decision variable of the observer and then evaluating the accuracy of the observer through Monte-Carlo studies.

Under the Gaussian assumptions made in this work, discriminating two classes on the basis of their covariance structure is well known to be optimally achieved by a quadratic form. This quadratic form is the log of the likelihood ratio for the two classes[11,12]. For the distributions specified for the scattering objects in Equation 2.5, this quadratic form is given by

$$\lambda_{\text{obj}}(\mathbf{f}) = -\frac{1}{2\sigma_{\text{obj}}^2}\mathbf{f}^t\left(\left(\mathbf{I}+\mathbf{S}_1\right)^{-1}-\left(\mathbf{I}+\mathbf{S}_0\right)^{-1}\right)\mathbf{f}\ . \tag{2.9}$$

Note that this decision variable is straight-forward to compute since the necessary matrix inversions are needed only on diagonal matrices. The Ideal observer in the RF domain is similarly defined,

$$\lambda_{\text{RF}}(\mathbf{g}) = -\frac{1}{2}\mathbf{g}^t\left(\boldsymbol{\Sigma}_1^{-1}-\boldsymbol{\Sigma}_0^{-1}\right)\mathbf{g}\ . \tag{2.10}$$

However, in this case the matrix inversions are nontrivial since they involve non-diagonal nonstationary covariance matrices. We have recently described a power-series method for computing such matrices that is efficient enough to be used in Monte-Carlo studies[7].

For the Ideal observer applied to envelope images, we adopt an optimal feature approach similar in character to the Laguerre-Gauss Hotelling approach described by Barrett et al.[13] Here the decision variable is given by a weighted sum of feature values. The features are linearly related to the square of the envelope image, and can be written

$$\lambda_{\text{env}}(b) \cong \sum_{j=1}^{N_F} w_j \mathbf{t}_j^t \mathbf{b}^2\ , \tag{2.11}$$

where w_j are the feature weights, $\mathbf{t}_j$ are the linear templates used to generate each feature, and $\mathbf{b}^2$ is the element by element square of the envelope image. The square of the envelope image is used for continuity with the work of Smith and Wagner[8]. They found that in the context of a large area low-contrast lesion, the Ideal observer is well approximated by a the square of the envelope image integrated over the lesion area.

The optimal weights are determined from the first and second moments of the feature variables as

$$\mathbf{w} = \left(\boldsymbol{\Sigma}_{\mathbf{t},1}+\boldsymbol{\Sigma}_{\mathbf{t},0}\right)^{-1}\left(\boldsymbol{\mu}_1-\boldsymbol{\mu}_0\right), \tag{2.12}$$

where $\boldsymbol{\mu}_1$ and $\boldsymbol{\mu}_0$ are the mean responses of the feature values in the two classes, and $\boldsymbol{\Sigma}_{\mathbf{t},1}$ and $\boldsymbol{\Sigma}_{\mathbf{t},0}$ are the covariance matrices of the feature values in each class. The size of the feature covariance matrices is determined by the number of features (not the number of image pixels), and hence sample estimates of the first and second moments are often used.

Human observers. Human observer performance was evaluated through psychophysical studies using two-alternative forced choice (2AFC) methodology. In these studies, two images were shown to the human observers, one from the malignant and one from the benign class. Figure 3 shows the appearance of such images (in Task 3) for standard B-mode envelope images and for Wiener filtered envelope images. The observers are asked to identify the malignant image. Under the

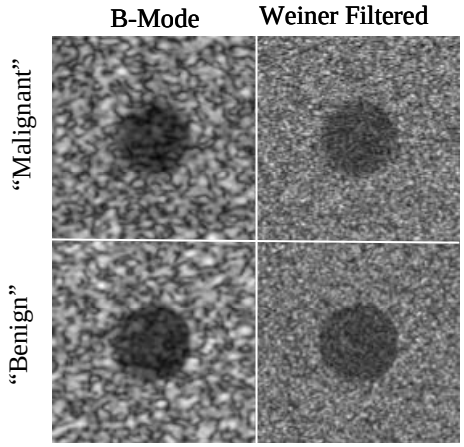

Figure 3. Examples of noisy B-mode and Wiener filtered images for Task 3 (at exaggerated contrast for display clarity).

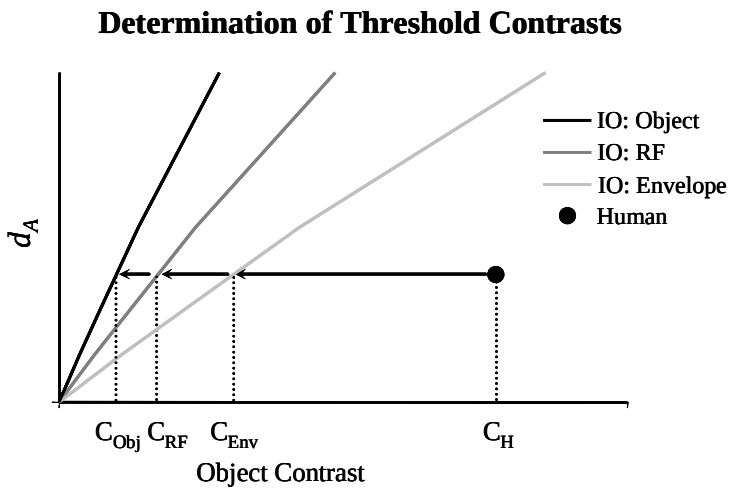

Figure 4. Computing contrast thresholds for observer efficiency. A human observer experiment with an object contrast C_H results in an observed level of detectability (●). The equivalent contrast needed for various ideal observers is found by the intersection of a line at the height of the human-observer detectability with the psychometric function of each observer.

assumption that human observers behave consistently, the proportion of correct responses in a 2AFC experiment is equivalent to the area under an ROC curve[14].

The object contrast of each task was manipulated in pilot studies to determine the object contrast needed to obtain approximately 80% correct performance on the standard envelope images in each task. A total of six observers participated completing all ten experiments (5 tasks each investigated for standard and Wiener filtered envelope images). Observers trained on 100 practice trials before each experiment. These trials used successively lower object contrast until the last 50 in which the observers operated at the level used in the data trials that followed. Human observer data was acquired from 400 trials in four sessions of 100 trials. Observers were encouraged to take short breaks between sessions to avoid fatigue. The order of the tasks was randomized across observers. The order of processing within each task was also randomized. The studies were performed in a darkened room on a calibrated monochrome monitor (Image Systems, Minnetonka, MN) using a perceptually linearized lookup table[15] over a luminance range of 0.5 to 150 Cd/m^2. The monitor pixel size was .3 mm. Viewing distance was not controlled, but observers generally maintained a comfortable viewing distance of approximately 40–50 cm.

Performance comparisons. Comparing human observers to Ideal observers typically involves equating one of two variables. Either the stimulus contrast can be equated and comparisons of performance are made, or performance is equated and comparisons of the contrast needed to achieve this performance are made. For this work, we use the equivalent performance approach. Figure 4 shows how a comparison between human observer and Ideal observers for the object, RF and envelope data are made. Performance data from the psychophysical studies (for humans) and Monte-Carlo studies (for Ideal observer) are converted to detectability scores by the formula

$$d_A = \sqrt{2}\Phi^{-1}\left(P_C\right), \tag{2.13}$$

where P_C is the proportion of correct responses and Φ is the cumulative normal distribution function. Equivalent contrasts for the Ideal observer are determined by finding the intersection of the psychometric function (plotting detectability as a function of object contrast) with a horizontal line that passes through the performance of the human observer.

3. RESULTS AND DISCUSSION

Observer performance. Human observer performance results for all five tasks with standard B-mode and Wiener filtered envelope images are shown in Figure 5. While there is some disparity in performance across the six observers, on average they are reasonably close to the targeted 80% correct level for the standard B-mode envelope images. The Wiener filtered envelope images show substantially higher performance in tasks 1 through 4. Average performance ranges from the high 87% in Task 2 to over 98% in Task 4. Notably, all six observers improve in these four tasks with Weiner filtering. These differences are all significant using a paired comparisons *t*-test ($p < 0.01$)[7].

However, the observers do worse in Task 5 with Weiner filtering where average performance falls from 81% in the standard B-mode envelope image to 73% for the Wiener-filtered envelope images. This decrease also occurs for all observers, and is of borderline significance ($p = 0.011$) in a paired comparisons *t*-test[7]. The relatively poorer performance with the Weiner filtered envelope images in Task 5 may be due to the fact that the Weiner filter is not well tuned to the lesion interior, where the magnitude of object scattering is considerably lower than in the area outside the lesion. Since the Weiner filter is tuned to the magnitude of scatter in the surrounding area, it may be inappropriately boosting some of the higher frequencies that would be predominantly object scatter at locations in the area surrounding

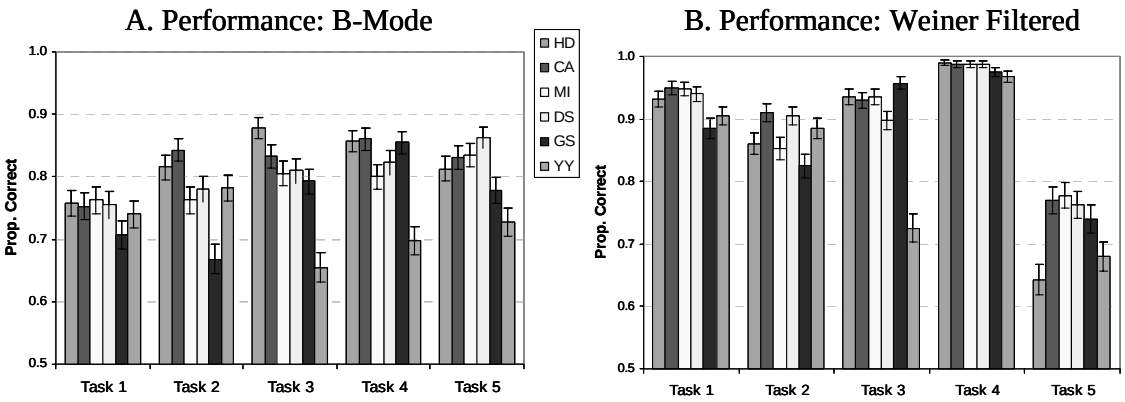

Figure 5. Human observer performance for standard B-mode (A) and Weiner filtered data (B). Proportion correct in 2AFC psychophysical experiments are given for each observer as a function of task. Error bars represent +/- 1 standard error.

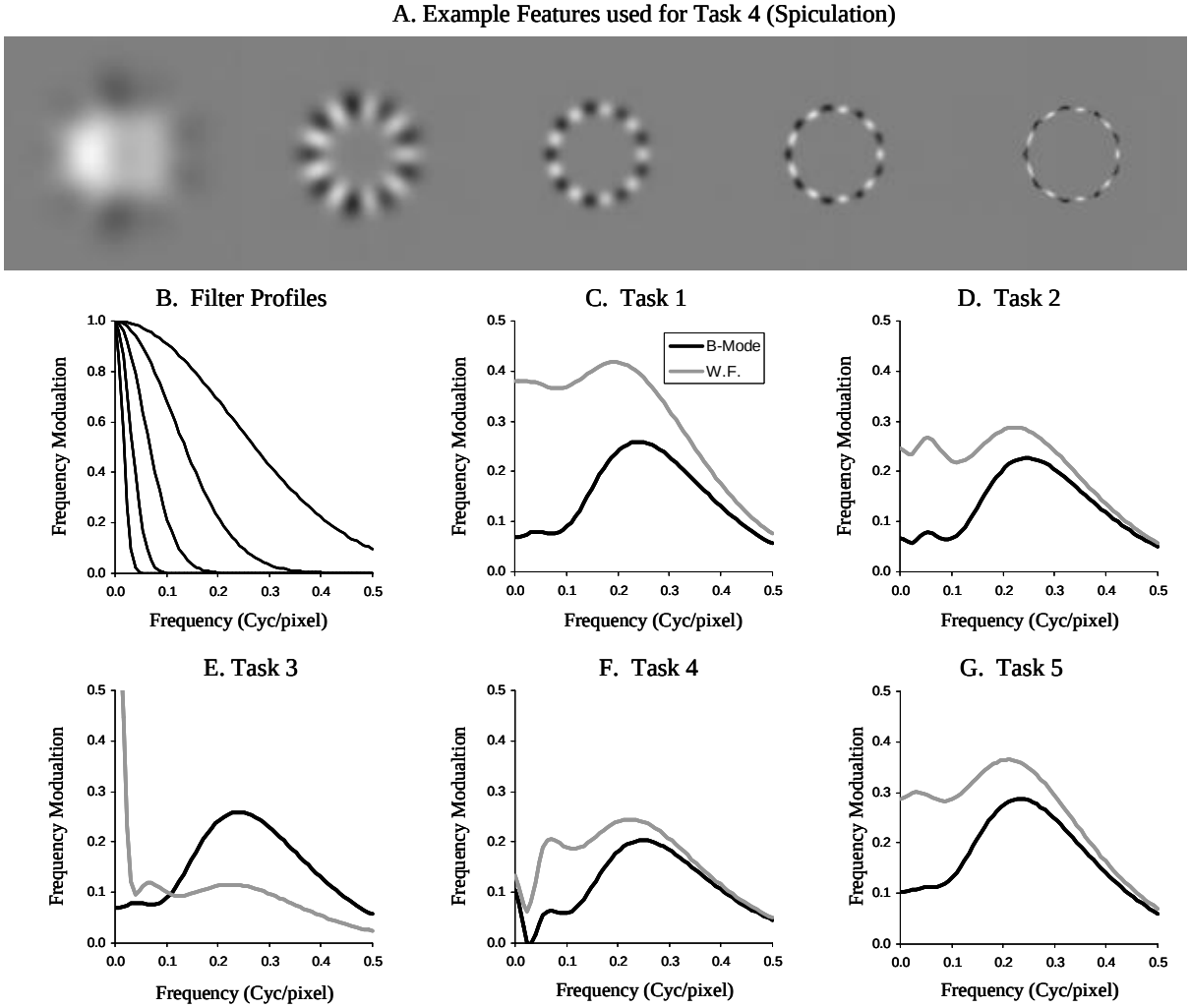

Figure 6. Channelized envelope ideal observer properties. The envelope ideal observer was computed from five features determined from lowpass filtering the Variance difference profile (A). The frequency modulation of the lowpass filters (B) was set so that the full-width at half max of the filter would increase by factors of 2 stating at 0.02 cyc/pixel. The frequency modulation of the resulting channelized ideal observers are plotted in C-G. The observers tailored to Wiener filtered envelope images consistently make greater use of lower spatial frequencies.

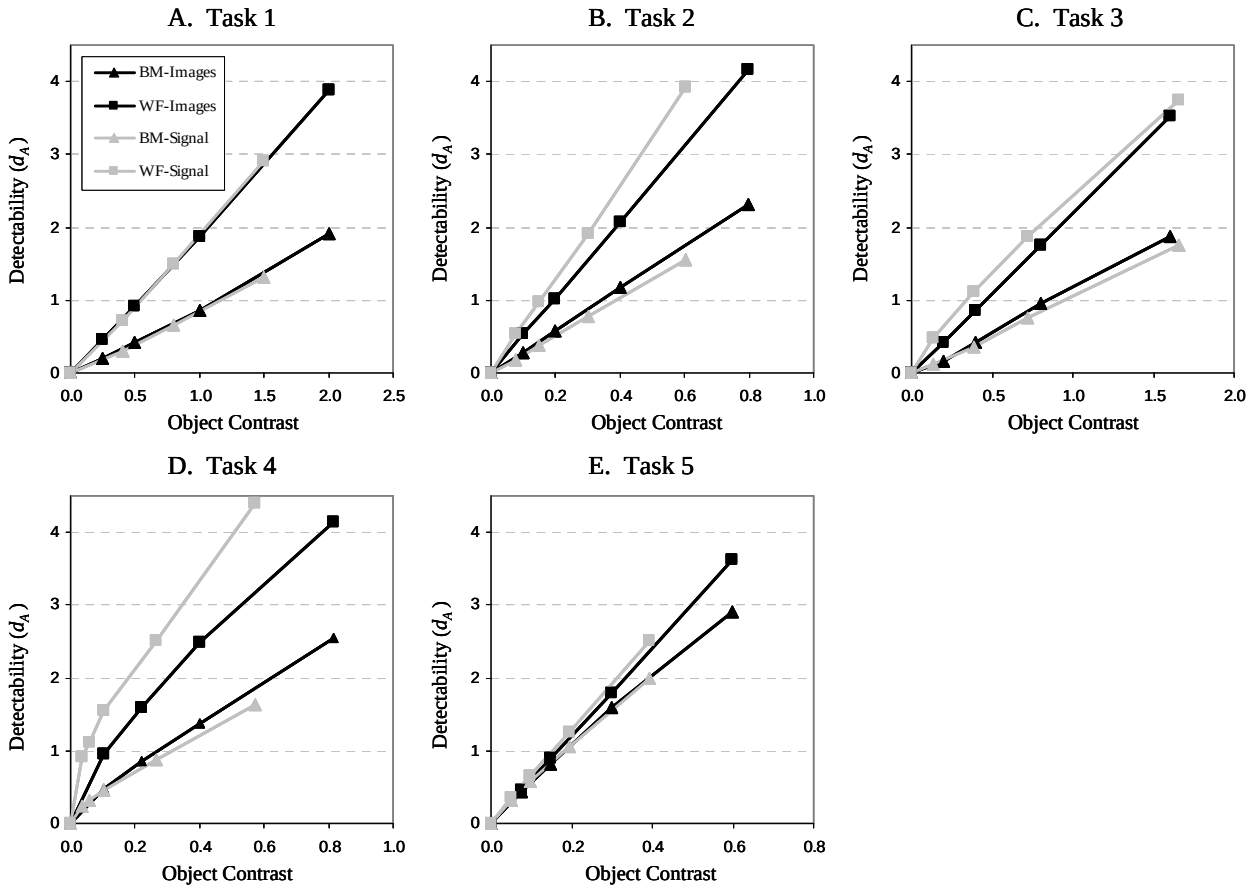

Figure 7. Envelope Ideal Observer performance. The plots show detectability as a function of object contrast for a "channelized" envelope ideal observer computed from the actual images (dark plots) and with the Smith-Wagner approximation computed from the demodualted RF signal (light plots). Both standard B-mode envelope processing (▲) and Wiener filtering enveloped (■) are shown. For tasks 1 and 5, the two methods of computation are roughly equivalent. However, in the boundary discrimination tasks (2, 3, and 4), there is often a substantial difference. The "channelized" ideal observer approximation acting on the final images results in improved performance for standard B-Mode processing and degraded performance for Wiener filtered envelope images.

the lesion, but are predominantly electronic noise in the low-scattering lesion interior. A better under standing of Wiener filter in this context is the subject of future work.

Envelope Ideal observers. As described above, we used an optimal feature combination approach to approximate the Ideal observer acting on envelope images, and we will compare this to previous work where we used the Smith-Wagner formulation of the Ideal observer. All results for this work were generated from Monte-Carlo studies with 6400 samples for each task. Each Monte-Carlo sample generated signal absent and signal-present objects at multiple contrasts for use in defining a psychometric function. Feature responses from these samples were used to determine the optimal weights (training) as well as evaluate the performance of the observer (testing). The large number of Monte-Carlo samples relative to the number of features used should remove any bias in performance from using common training and testing sets.

The feature combination approach used here has the ability to use multiple features in a way that may allow for improvement over the Smith-Wagner observer. However, we also apply these filters to downsampled and quantized images that may reduce performance. Figure 6A shows the set of 5 features used to approximate the Ideal observer in Task 4. These features are generated by taking the difference in variance profiles, $\mathbf{S}_1 - \mathbf{S}_0$, downsampling by a factor of 2 in the axial direction and then lowpass filtering by one of 5 Gaussian lowpass filters. The frequency response of these filters is shown in Figure 6B, and has half-max values at 0.02, 0.04, 0.08, 0.16, and 0.32 cycles/pixel respectively.

The Ideal feature is constructed by a weighted sum of feature values as described in Equation 2.11. However, it is also of interest to know how these Ideal features modulate frequencies of the Smith-Wagner observer. This can be done

by plotting a composite frequency response determined from combining the frequency modulation functions in Figure 6B with the optimal weights determined from Equation 2.12. Plots of this composite frequency response are found for each task in Figures 6C through 6G. While the plots are somewhat variable (indicating that there may be better feature sets to consider), the consistent difference between the approximate Ideal observer using standard B-mode envelope images and Wiener filtered envelope images is that low spatial frequencies are more informative in the Wiener filtered envelope images.

Performance comparisons between an approximate Ideal observer computed as described above and the Smith-Wagner approximation to the Ideal observer used previously are given for each task in Figure 7. The two approaches are very closely matched in Tasks 1 and 5, indicating fairly stable performance in the large area tasks. However in Tasks 2 - 4, the optimal feature combination approximation consistently outperforms the Smith-Wagner observer for standard B-Mode envelope images. For the Wiener filtered envelope images, this effect is reversed with the Smith-Wagner observer consistently outperforming the approximation based on optimal feature combinations. We believe that these results indicate that the low-frequency suppression observed for B-mode envelope images in Figure 6 translates into improved performance overall. However, the deleterious effects of downsampling and quantization are strong enough to reduce performance for the Wiener filtered images.

Efficiency results. The main result of this work is a characterization of the ultrasonic imaging model in terms of the relative efficiency of each step, as shown in Figure 8. The overall efficiency of the imaging process is the product of relative efficiencies of each step. For a task in which human observer performance is evaluated at an object contrast of C_{hum}, and equivalent performance of the Ideal observer acting on the object, the RF frame data, and the envelope images is given by C_{obj}, C_{RF}, and C_{env} respectively, the overall efficiency of the human observer can be decomposed into the product of three relative efficiencies given by

$$\begin{aligned} \text{EFF} &= 100\% \times \left(\frac{C_{\text{obj}}}{\text{C}_{\text{hum}}} \right)^2 \\ &= 100\% \times \left(\frac{C_{\text{obj}}}{\text{C}_{\text{RF}}} \right)^2 \left(\frac{C_{\text{RF}}}{\text{C}_{\text{env}}} \right)^2 \left(\frac{C_{\text{env}}}{\text{C}_{\text{hum}}} \right)^2 . \end{aligned}$$

The three plots of Figure 8 correspond to each of the relative efficiencies above.

The efficiency of the RF data is given in Figure 8A. In this case we find that the large area tasks lead to efficiencies of 10-40%. The boundary discrimination tasks show approximately an order of magnitude decrease in efficiency with values around 1%. We also see that the standard B-mode and envelope images yield roughly the same efficiency. This finding is not surprising since the effects are computed pre-envelope, and hence they indicate that the efficiency is relatively constant across different performance levels. The one task for which substantial differences are observed is Task 4, the spiculated boundary task. However this may be an aliasing effect since the size of spiculations at this level is at the limit of the object sampling size. The low efficiency of the RF frame data for boundary discrimination tasks indicate that the pulse-acquisition system is better suited to large area tasks.

The relative efficiency of the Envelope images in Figure 8B shows a substantial difference between the standard B-mode envelope images and Wiener filtered envelope images. This indicates that the (approximate) Ideal observer acting on the envelope images is considerably less efficient on standard B-mode envelope images in Tasks 1 -4. In Task 5, the efficiencies are approximately equivalent. Figure 8B implies a considerable loss of diagnostic information in the computation of a standard B-mode envelope image.

The relative efficiency of human observers to the envelope Ideal observer is given in Figure 8C. Tasks 1-3 show high relative efficiencies with values from 30% to 60%. Task 4 shows somewhat less efficiency with values of 8% and 17% for standard B-mode and Wiener filtered envelope images respectively. Task 5 shows yields the lowest efficiency around 5%. Compared to Figure 8B, the efficiencies for B-mode and Wiener filtered envelope images are fairly well matched for Tasks 1 – 4. In these tasks, the computation of an envelope image is the factor limiting diagnostic performance, not the human observer's ability to read the resulting image. In Task 5, the situation reverses with most of the lower efficiency coming from the human observer's ability to read the image. This is not overly surprising in light of the mismatched Wiener filter as described above. The filter will tend to inappropriately boost higher frequencies in the lesion interior, and human observers are known to be less efficient in highpass noise[16,17].

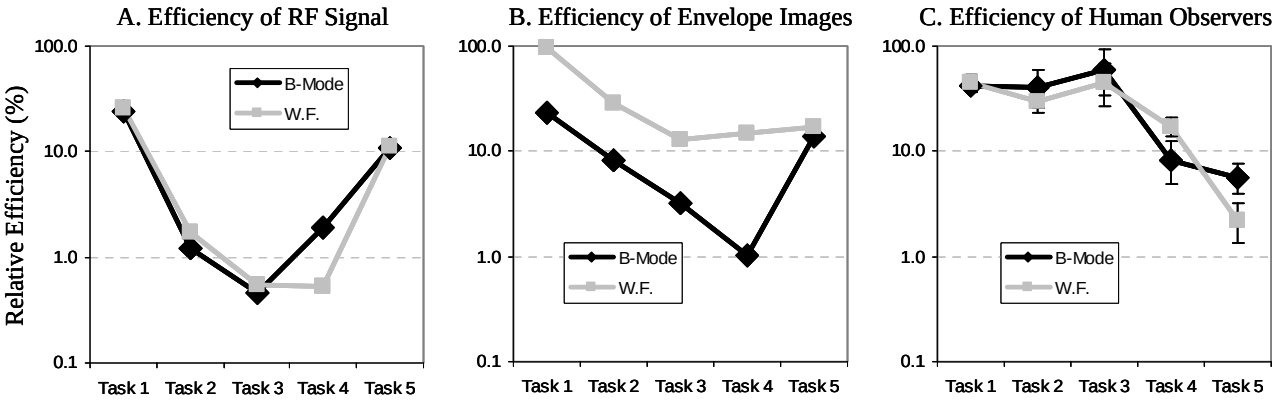

Figure 8. Efficiency at various stages of the imaging chain. The plots show relative efficiency going from object to RF signal (A), From the RF signal to the envelope image (B), and from the envelope image to the human observer. The most consistent source of lower efficiency in the standard B-Mode processing compared to Wiener filtered RF data comes from the computation of an envelope image.

4. SUMMARY AND CONCLUSION

This work has added two new results to what we have published previously analyzing ultrasonic imaging in terms of performance on tasks related to discriminating malignant and benign lesions in the breast. We have investigated the efficiency of image formation process by comparing the efficiency of the Ideal observer acting on the detected radio-frequency signal, with a hypothetical Ideal observer given access to the actual stochastic scattering functions of the tissue itself. We have also revisited the Ideal observer acting on envelope images. These results show how early and late components of the imaging process affect task performance, and taken in combination with our previous work, they present a sequential analysis of the ultrasonic imaging chain that is reminiscent of the cascaded systems analysis of Yao and Cunningham[18]. However, our results also rest on some strong assumptions that require further investigation. We summarize each result and the associated assumptions below.

The relative efficiency going from stochastic tissue scattering functions to the RF signal shows that, compared to low contrast detection and discrimination tasks, boundary discrimination tasks are not nearly as efficiently captured by the US system. We find as much as an order of magnitude drop in efficiency for the boundary discrimination tasks relative to the large area (low contrast) detection and discrimination tasks. This bottleneck in the transfer of diagnostic information suggests, from a theoretical viewpoint, that there may be considerable benefit from improvement in pulse formation and pulse-acquisition strategies.

This result was developed with the assumption of a rich scattering medium that leads to fully developed speckle[19]. We have also assumed that we can sample the tissue scattering function – which is continuous – without changing the relative differences in efficiency. A finer sampling can only improve Ideal observer performance, however we assume it will improve performance in all tasks roughly equivalently. Finally our result relies on the model of RF acquisition we use and convergence of the power-series inversion of covariance matrices associated with it.

Our second result came from reconsidering the Ideal observer acting on an envelope image. In this work we included downsampling and quantization in the formation of an envelope image. We also used a multi-filtered "channelized" approach to finding an optimal linear weighting[13]. We found that this approach lead to slightly improved performance in standard B-mode images (from the ability to suppress lower spatial frequencies), but somewhat lower performance for the envelope images computed after Weiner filtering the RF frame data (perhaps from decimation or intensity quantization). The new envelope image Ideal observer shows that human observers are about equally efficient at performing tasks on envelope images with or without Wiener filtering.

The new efficiency results confirm that a second bottleneck in the transfer of diagnostic information for boundary discrimination tasks comes in the formation of an envelope image. Wiener filtering the RF frame data improved Ideal observer efficiency on the envelope data by factors ranging from 4 to over 10. Since human observers are roughly equivalent in efficiency from envelope images with or without Wiener filtering, improvements in the method of computing an envelope image (such as Wiener filtering) appear to transfer fairly directly to improved human observer performance. The scale of improvement was true also for the detection task, but notably absent from the discrimination task. We believe the lack of improvement for low-contrast discrimination reflects the fact that the Wiener filter used was not tuned to the lower level of tissue scatter present in the lesion interior.

The Ideal observer acting on the envelope images assumes that a linear observer acting on the square of the envelope image is approximately optimal. This was the assumption in the Smith Wagner formulation, but that work also assumed a large-low-contrast lesion, which is not the case with boundary discrimination tasks. It is also not clear that the limited set of five features used to generate an "optimal" template is in fact sufficient. Nonetheless, there appear to be considerable gains to be made by processing the RF frame data before computing the final envelope image.

ACKNOWLEDGMENTS

This work was supported in part by the NIH (R01-CA082497 and R21-CA102733). The authors are grateful to the observers for their diligence in the psychophysical studies.

REFERENCES

1. *Breast Imaging Reporting and Data System Atlas*. Reston, VA: American College of Radiology, 2003.
2. L. W. Bassett and C. Kimme-Smith, "Breast sonography," *Am. J. Roentgenol.*, vol. 156, no. 3, pp. 449–455, 1991.
3. H. M. Zonderland, E. G. Coerkamp, J. Hermans, M. J. van de Vijver, and A. E. van Voorthuisen, "Diagnosis of breast cancer: Contribution of US as an adjunct to mammography," *Radiology*, vol. 213, no. 2, pp. 413–422, 1999.
4. T. M. Kolb, J. Lichy, and J. H. Newhouse, "Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: An analysis of 27,825 patient evaluations," *Radiology*, vol. 225, no. 1, pp. 165–175, 2002.
5. T. M. Kolb, J. Lichy, and J. H. Newhouse, "Occult cancer in women with dense breasts: Detection with screening US-diagnostic yield and tumor characteristics," *Radiology*, vol. 207, no. 1, pp. 191–199, 1998.
6. P. Crystal, S. D. Strano, S. Shcharynski, and M. J. Koretz, "Using sonography to screen women with mammographically dense breasts," *Am. J. Roentgenol.*, vol. 181, no. 1, pp. 177–182, 2003.
7. 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. Imag. 25*(2): 198-209, 2006.
8. 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. SU-30, pp. 156–163, 1983.
9. M. O'Donnell, Coded excitation systems for improving the penetration of real-time phased array imaging systems," IEEE Trans UFFC, vol. 39(3), pp. 341-351, 1992.
10. W.F. Walker, "C- and D-weighted ultrasonic imaging using the translating apertures algorithm," *IEEE Trans Ultrason Ferroelectr Freq Control*, 48(2):452-61, 2001.
11. K. Fukunaga, "Introduction to statistical pattern recognition," Academic Press Inc., Boston, USA, 1990.
12. R. F.Wagner and H. H. Barrett, "Quadratic tasks and the ideal observer," in *Proc. SPIE*, vol. 767, pp. 306–309. 1987.
13. H.H. Barrett, C.K. Abbey, and B.D. Gallas, "Stabilized estimates of Hotelling-observer detection performance in patient structured noise," in *Proc. SPIE*, 3340, pp 27-43, 1998.
14. D.M. Green and J.A. Swets, *Signal Detection Theory and Psychophysics*, NY: Wiley, 1966.
15. E. A. Krupinski and H. Roehrig, "The influence of a perceptually linearized display on observer performance and visual search," *Acad. Radiol.*, vol. 7, no. 1, pp. 8–13, 2000.
16. K.J. Myers, H.H. Barrett, M.C. Borgstrom, D.D. Patton, and G.W. Seeley, "Effect of noise correlation on detectability of disk signals in medical imaging," *J Opt Soc Am A*. 2(10):1752-9, 1985.
17. C.K. Abbey and M.P. Eckstein, "Estimates of human-observer templates for simple detection tasks in correlated noise," *Proc. SPIE* (E.A. Krupinski, Ed.) 3981:70-77, 2000.
18. J. Yao and I.A. Cunningham, "Parallel cascades: new ways to describe noise transfer in medical imaging systems." *Med Phys.* 28(10):2020-38, 2001
19. J.W. Goodman, "Statistical properties of laser speckle patterns," in *Laser Speckle and Related Phenomena*, (J.C. Dainty ed.) Springer-Verlag, Berlin, 1975.