INFORMATION THEORETIC DESIGN OF BREAST SONOGRAPHY

BY

NGHIA QUANG NGUYEN

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2009

Urbana, Illinois

Adviser:

Professor Michael F. Insana

ABSTRACT

First-principle approaches to the design of medical ultrasonic imaging systems for specific visual tasks are being explored. Our study focuses on breast cancer diagnosis and is based on the ideal observer concept for visual discrimination tasks, whereby tasks based on five clinical features are expressed mathematically as likelihood functions. Realistic approximations to the ideal strategy for each task are proposed as an additional beamforming procedure to maximize diagnostic image information content available to readers. Our previous study revealed that the Wiener filter, derived as a stationary approximation of the ideal observer and operating on RF echo data, generally improved discriminability except for one case involving high-contrast lesions. This study explores an adaptive, iterative Wiener filter that includes a lesion segmentation algorithm to improve discriminability for high-contrast lesions. Predicted performance is compared with that measured from trained human observers using psychophysical methods. The iterative Wiener filter was found to match the performance of the Wiener filter for low-contrast lesions and increase the performance for high-contrast tasks. This filter offers greater diagnostic performance for discriminating malignant and benign breast lesions, and it provides a rational basis for further task-specific imaging system design. The thesis also addresses some issues that may be encountered when applying the filter in clinical environments. We found that background variability improves the performance of spatial filters. We also studied the limitation of these filtering methods as key assumptions are violated.

To my family

ACKNOWLEDGMENTS

I would first like thank my adviser, Prof. Insana, whose valuable thoughts and vision have made this thesis possible. His guidance and encouragement are also a powerful inspiration to me. Second, I would like to thank Prof. Craig Abbey of the University of California, Santa-Barbara, for his creativity, contribution, and mentoring related to many aspects of the work. Thanks to people in the Ultrasonic Imaging Lab for their friendship and discussions. And finally, thanks to my wife Van and my parents for all their love, support, and encouragement. The thesis is dedicated to them.

This material is based upon work supported by the National Institutes of Health (NIH) under award No. CA118294.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

2AFC	two alternative force choice
AUC	area under ROC
BM	basement membrane
ECM	extracellular matrix
IWF	iterative Wiener filter
IWFB-mode	iterative Wiener filtered B-mode
MAP	maximum a posteriori
MRF	Markov random field
RF	radio frequency
ROC	receiver operating characteristic
SNR	signal-to-noise ratio
SW	Smith-Wagner
WF	Wiener filter
WFB-mode	Wiener filtered B-mode

CHAPTER 1

INTRODUCTION

A medical imaging system is an integration of operations that include acquiring, recording, processing, distributing, and displaying medical information. The performance of an imaging system is evaluated in the context of how well it delivers specific information. There are some features and complicated tradeoffs in these operations that impact the performance. For breast sonography, the information required is often benign/malignant classification of a lesion based on image features such as echogenicity and margin regularity. The diagnostic performance is best when the transducer beam energy is tightly focused uniformly throughout the field of view. Yet, in practice, beamforming demands compromises. For example, transducer apertures may be apodized to suppress sidelobe energy at the cost of widening the main lobe of the sound beam. Thus apodization trades off spatial resolution for improved contrast resolution, and depending on the clinical imaging objective this technique may or may not improve diagnostic performance. As processing speed increases, systems more readily are reconfigured to optimize performance for specific clinical examinations. Therefore, it is important to allow for flexibility in the signal processing strategy to adaptively optimize image formation algorithms for specific clinical tasks. In our research, we provide a method for introducing clinical information into a Bayesian discriminant function [1] – the ideal observer – to discover strategies for enhancing task performance through signal processing. The approach has some advantages in system optimization because the ideal observer captures all relevant information of the data, yielding optimal task performance [2].

The ideal observer approach is an application of statistical detection theory. It was first used in radar imaging during the World War II, with scientific background provided decades earlier by Hotelling, Thurstone, and Neyman & Pearson [3]. The approach was introduced into ultrasonic imaging quality assessment in 1983 by Smith, Wagner, and co-workers [4, 5]. They applied the ideal observer approach to sonographic diagnosis by showing the role of speckle size, lesion contrast and echo SNR in the detectability of large low-contrast lesions. In their research, the ideal observer was implemented on B-mode images, which are demodulated radio frequency (RF) echo signals. The ability of modern ultrasound systems to digitize and store RF

data allows us to focus on RF data, which contains all of the available diagnostic information obtained by the transducer detector. By translating clinical tasks into discriminant functions we find optimal task-based imaging strategies for processing the data before computing the final envelop image.

We previously described a spatial Wiener filtering (WF) technique that emerged from an approximation of an ideal observer analysis of breast sonography, in which the filter is applied to simulated, fixed-focus, RF echo data to produce B-mode images [6]. We first worked with a practicing mammographer to isolate discriminating image features and then translate those features into conditional probabilities – likelihood functions – that quantify the visual task. Breast lesion features, like margin smoothness, were introduced into tissue scattering functions by spatially modulating the variance of acoustic impedance. Thus the diagnostic task is specified entirely by the object covariance matrix. The sonographic system is represented as a linear transformation, modeled as a simple Gabor-pulse function, acting on the stochastic object to generate RF signals processed to form B-mode images. System properties and the subsequent signal and image processing performed to enhance human observer accessibility are combined to increase the efficiency of diagnostic information transferred from the object to the observer.

With the assumption of zero-mean Gaussian stochastic process for scattering properties of tissue [7], the RF data can be described by correlated multivariate Gaussian distribution. Therefore, the ideal observer decision variable is given explicitly by a quadratic function of the RF data involving the inverse of covariance matrices. However, the discriminant function of the ideal observer for visual tasks is difficult to compute exactly and interpret. We resolved the computational problem and gained insights by expanding the discriminant function into a power series and separating stationary components of covariance from the nonstationary components that carry information. We found that the first-order approximation to the ideal-observer strategy involved applying a spatial Wiener filter to fixed-focus beamformed RF signals before echo demodulation and scan conversion. Images from filtered echo signals, labeled as WFB-mode images, resulted in improved visual detectability for four of five features characteristic of malignant-benign lesion discrimination. For visual tasks where the statistical properties of the echo signal were accurately represented everywhere in the image field, WFB-mode images outperformed conventional B-mode images obtained using delay-and sum beamforming. However, in one visual task where the diagnostic information was contained in a region of lower echo SNR compared to surrounding regions, WFB-mode images reduced observer detection performance.

In this thesis, we extend the analysis to include a greater range of practical conditions by improving the linear approximations. The proposed method addresses deficiencies in the application of the power series expansion to calculate the inverse of the covariance matrices, which in return yields a better linear approximation. It leads to an iterative Wiener filtering

(IWF) approach that re-tunes the filter in regions of variable echo SNR to generate IWFB-mode images. Iteration enables the filter to adapt to variable tissue scattering conditions but requires a segmentation algorithm beforehand to identify regions that can benefit from iterative filtering. Most segmentations of medical ultrasound images are performed on B-mode images [8, 9]. However, our segmentation is performed on the WFB-mode images. This modification allows us to more accurately segment lesions and with less computation time. The results show that the new filter makes an improvement in visual performance for diagnostic information contained within the lesion interior, the task where we find a reduction in performance after Wiener filtering, while keeping the performance on par with that of the Wiener filter in other tasks. In this thesis, we also introduce a more realistic imaging ultrasound pulse that is generated by the Field II program [10, 11]. Field II is a public-domain simulator of ultrasonic pulses from single element and array transducers. We also examine the consequences of applying computational conveniences, such as a shift-invariant system and a circulant system matrix, to the analysis. Approximations are important for computation speed but must be carefully applied when they influence predictions of image discrimination performance. Finally, we expand the signal-known-exactly paradigm to allow for background variability when the background is known statistically. A focus of this thesis is to study how robust the Wiener-filter approach is to uncertainty in the filter parameters so that we can better understand how to apply these methods to clinical instruments.

CHAPTER 2

BACKGROUND AND MODELING

2.1 Background

In medical imaging classification, analysis of the system is possible only by modeling each component, from the object at the input and the system used to generate image data to the decision made by a human or computational observer. Our model for analysis is illustrated in Figure 2.1. The output of the system is observed and the observer then obtains some features regarding the possible objects being imaged, in the way the object is distorted by the image acquisition progress and random processes inside the system. In breast sonography, the features might be presence of a lesion, or some characteristics to classify a detected lesion into a particular class (e.g., benign or malignant). Based on the information obtained, the observer infers the class of the object at the input of the imaging system [1].

As mentioned in the Introduction, the ideal observer approach was first applied to B-mode image data because that is what humans view for diagnosis. RF data contain the phase of the echo signal, which can provide diagnostic information even if it is not readily accessible to the human eye-brain system. RF data is usually demodulated to discard phase information in the image presented to humans for diagnosis. In our research, we propose an investigation of the ideal observer in the RF domain, which means that the ideal observer acts on all available information. Exploring the ideal observer of the RF data, we can track the flow of information at each step of the image formation process to decide where information is lost or inaccessible to the visual system. Inaccessible information provides opportunities for applying image processing algorithms that enhance human observer performance.

Another advantage of moving to the RF domain is that we simplify the statistical model by applying a linear system approach when computing the ideal observer's test statistic. In the next section, we analyze the acquisition of RF data and describe how to model the formation of signals inside the ultrasonic system.

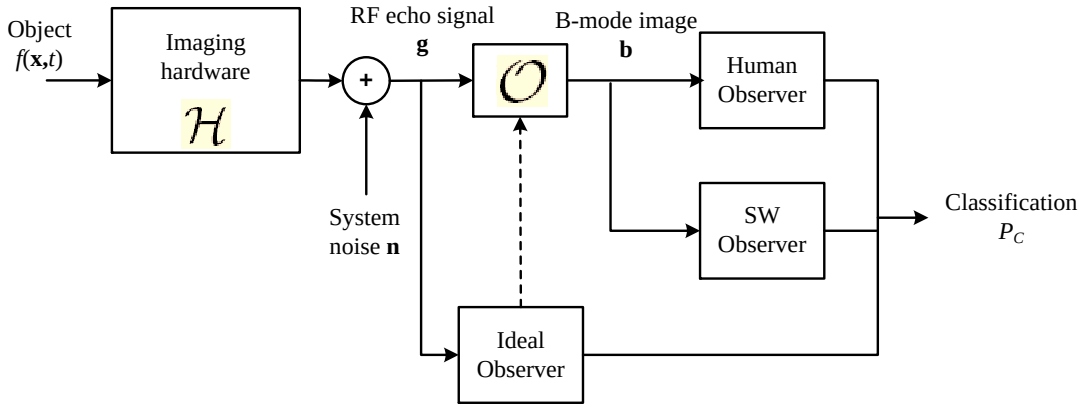

Figure 2.1: Model of data formation and observer processing as steps toward classification.

2.2 RF Data Acquisition

As discussed in Appendix A, the RF data recorded from a typical ultrasonic system can be modeled as the output of a linear process, given as

$$g(\mathbf{t}) = \int d\mathbf{x} h(\mathbf{x}, \mathbf{t}) f(\mathbf{x}) + n(\mathbf{t}), \tag{2.1}$$

where g is time-varying RF data. The function h is the pulse-echo spatio-temporal impulse response, f is the scattering object, and n is the system noise as measured at the output. Vector $\mathbf{t}$ represents the axial and lateral dimensions of the object. The scattering object f is the same as γ in Eq. (A.1). We change the denotation here to conveniently match the notation in our previous study [6].

The data structure for 2D B-mode acquisition is shown in Figure 2.2. A 1D linear array generates echo sequences oriented in columns which are A-lines. The RF data is stored digitally with elements $g(\mathbf{t}) = g(t_1, t_2) = g(lT, mLT)$. The L range samples are placed in columns for each A-line and assigned a discrete time stamp $t_1 = lT$, where $0 \leq l \leq L-1$ are integers and T is the temporal sampling interval. The A-lines are recorded sequentially from left to right, where $t_2 = mLT$, $0 \leq m \leq M-1$. The acquisition time is given by $t' = t_1 + t_2 = (l + mL)\,T$, and l is indexed completely from 0 to L - 1 before each increment of m. It is noted that commercial systems often employ nonsequential echo acquisition to speed recording. This method complicates echo-signal bookkeeping beyond our simple approach, but the basic linear model is the same.

For a 2D object, the conversion from temporal coordinates to spatial coordinates is as follows:

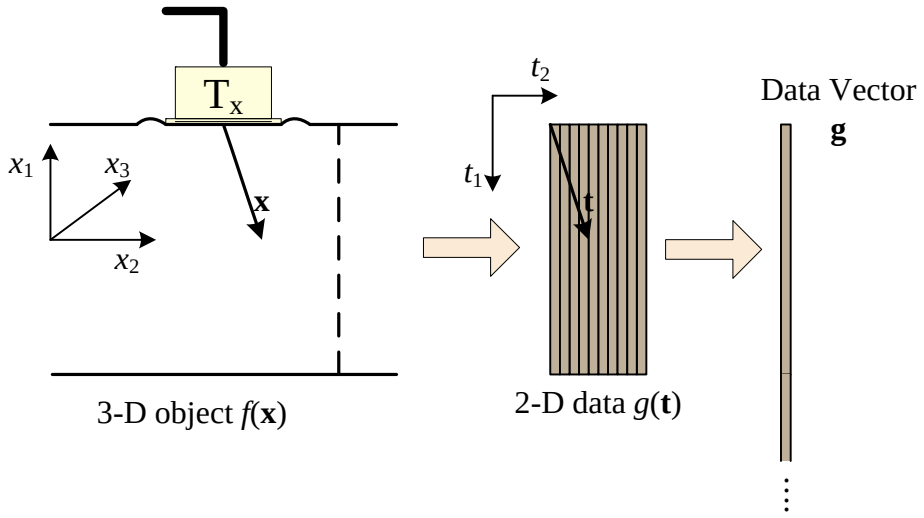

Figure 2.2: The geometry of mapping from an object in space to a data vector in the time domain [12].

$t_1 = lT \rightarrow lcT/2 = l\Delta X_1$, and $t_2 = mLT \rightarrow m\Delta X_2$, where c is the speed of sound, ΔX_1 and ΔX_2 are axial and lateral sampling intervals. Thus we can represent the conversion by a matrix multiplication, $\mathbf{t} = \mathbf{B}\mathbf{x}$, where $\mathbf{B}$ is a diagonal matrix. Specifically, the conversion is given as

$$\begin{bmatrix} t_1 \\ t_2 \end{bmatrix} = \begin{bmatrix} 2/c & 0 \\ 0 & LT/\Delta X_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}. \tag{2.2}$$

This conversion can be generalized for scanning in a 3-D object. Each point in the time domain (t_1, t_2) maps to points in the (x_1, x_2) plane where scattered energy is integrated over x_3.

By assuming that RF data is acquired in an *isoplanatic* region, where the impulse response is unchanged, h can be written in terms of a single independent variable [12],

$$h(\mathbf{x}, \mathbf{t}) = h(\mathbf{t} - \mathbf{B}\mathbf{x}). \tag{2.3}$$

Substituting Eq. (2.3) into Eq. (2.1), we obtain

$$g(\mathbf{t}) = \int d\mathbf{x} h(\mathbf{t} - \mathbf{B}\mathbf{x}) f(\mathbf{x}) + n(\mathbf{t}). \tag{2.4}$$

Thus, under the isoplanatic assumption, RF data can be considered as a noisy linear convolution of the scattering object and a system impulse response. Hereby, we use this assumption to generate the echo data in order to simplify computation. The formation of signals in an ultrasonic

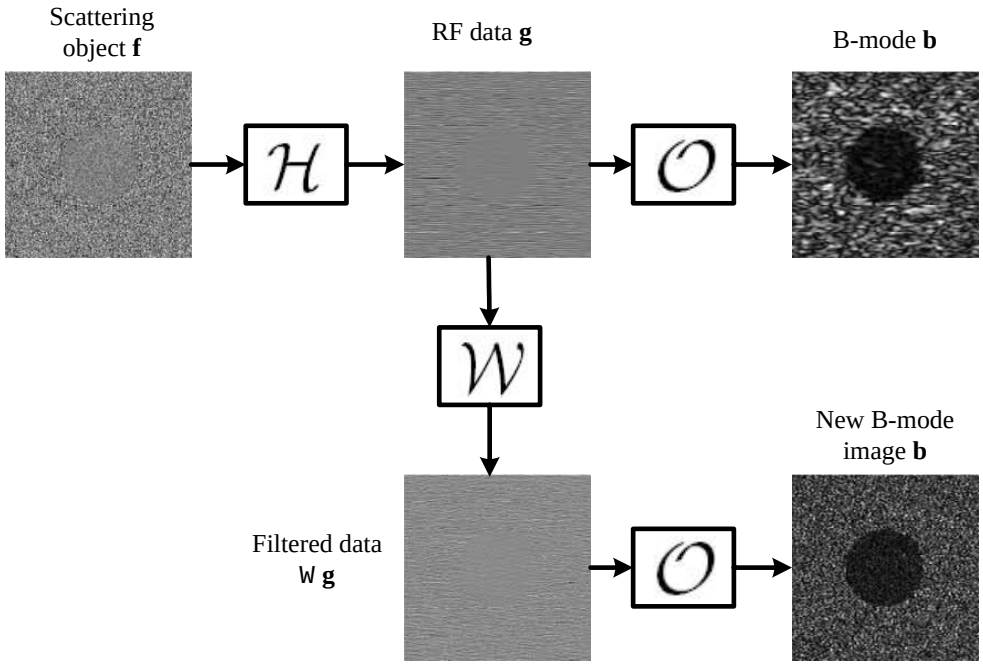

Figure 2.3: A graphical model of the sonographic process to generate B-mode images. The top row is conventional processing while the bottom row includes additional filtering, $\mathcal{W}$.

system is described in the next section.

2.3 Signal Modeling

The sonographic image formation process considered is graphically depicted in Figure 2.3. The tissue scattering field $f(\mathbf{x})$ at vector position $\mathbf{x}$ is represented by zero-mean Gaussian stochastic process [7, 13]. It is numerically modeled as a multiplication of a variance map with a 2D Gaussian random field, in which the variance map describes the lesion feature. The scattering object then maps into the RF echo signal vector $\mathbf{g}$ through the continuous-space-to-discrete-time linear operator $\mathcal{H}$ [2] to describe the *detection stage* of the imaging process [14]. Vector $\mathbf{g}$ may be the array of echo signals recorded from individual receiver channels, but in this work it is the beamformed RF echo signal. B-mode image data $\mathbf{b}$ are obtained from $\mathbf{g}$ through demodulation and scan conversion during the *display stage* of the image formation process. This processing is represented by a nonlinear, discrete-time-to-discrete-space operator $\mathcal{O}$ [12]. After beamforming and other linear operations applied before B-mode image rendering, the echo signals may be filtered via operator $\mathcal{W}$ resulting in new B-mode images. While objects $f(\mathbf{x})$ are

three dimensional, vectors $\mathbf{g}$ and $\mathbf{b}$ represent two-dimensional data acquired from an scan plane of a 1-D linear array transducer. These processes are summarized by the equations

$$\mathbf{g} = \mathcal{H}f(\mathbf{x}) + \mathbf{n} \text{ and either } \mathbf{b} = \mathcal{O}\mathbf{g} \text{ or } \mathbf{b} = \mathcal{O}\mathcal{W}\mathbf{g}, \tag{2.5}$$

where $\mathbf{n}$ is a vector of noise values from an independent zero-mean Gaussian process with variance σ_{n}^2.

In breast sonography, an ultrasound pulse is sent from a source transducer into the body and echoes are detected from back-scattered waves by the same transducer now serving as receiver. Scattering occurs because of spatial variability in acoustic impedance througout the propagation medium. In the context of the linear system described above, the whole process can be modeled as a linear convolution in the continuous space-time domain. Therefore we simulate the scattering function numerically by sampling continuous f at a rate higher than that represented in image $\mathbf{b}$. For example, a scattering function vector $\mathbf{f}$ may be generated by sampling $f(\mathbf{x})$ at a 0.002 mm interval. Images $\mathbf{g}$ and $\mathbf{n}$ are sampled at 40 Msample/s, which corresponds to a spatial stepping of 0.02 mm along the beam axis and 0.2 mm lateral to the beam axis. Image $\mathbf{b}$ may be rendered in 2D as an image with a square 0.02 mm pixel size for purposes of observation.

Operators $\mathcal{H}, \mathcal{W}, \mathcal{O}$ are applied via matrix multiplications, so that Eq. (2.5) may be expressed as $\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n}$ and $\mathbf{b} = \mathbf{O}\mathbf{W}\mathbf{g}$. For conventional B-mode display, $\mathbf{W} = \mathbf{I}$, the identity matrix. Note that each row of the $\mathbf{H}$ matrix is a sampled impulse response from the imaging system that is applied to one echo sample; for shift invariant systems, $\mathbf{H}$ has the form of a block-Toeplitz matrix [15]. The system response is modeled after the VF10-5 linear array probe on the SONOLINE Antares$^{\mathrm{TM}}$ System (Siemens Medical Solutions, Mountain View, CA) by using the Field II program [10, 11]. This makes the simulation more realistic than that of our previous study [6] which modeled the sonographic system as a Gabor pulse. Imaging system parameters were used in the echo simulations and the measurements are described below.

Beamformed RF echo waveforms are sampled at 40 Msample/s. This temporal sampling corresponds to a spatial sampling along the beam axis of 0.02 mm for a $c = 1540$ m/s tissue-like sound speed. The lateral sampling interval is set at 0.2 mm, equal to the element pitch. We set a 40-mm transmit/receive focal length and a 96-element ($\sim$20 mm) active aperture (f/2 in plane). The array has 192 total elements separated by a 0.02 mm element kerf. The elevational element length is 25 mm and it is focused at 40 mm. We applied a two-cycle excitation voltage and measured a 51% pulse-echo bandwidth about a 7.2 MHz center frequency. The beamformed, echo SNR was also measured at 32 dB near the focal length. Dynamic focusing and aperture growth were disabled. We assembled the system matrix $\mathbf{H}$ from the temporal responses (matrix

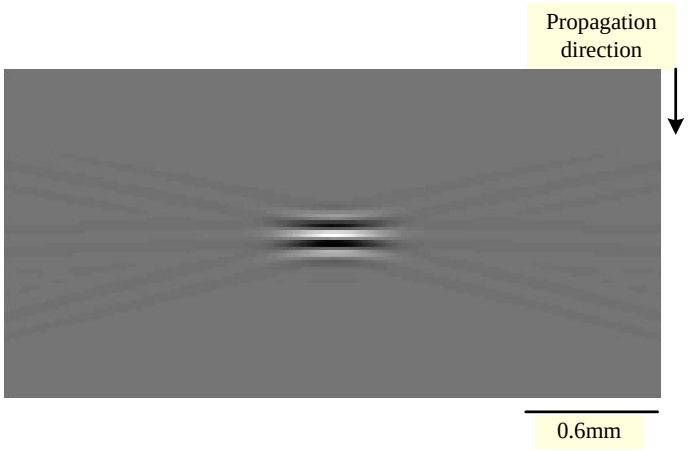

Figure 2.4: The simulated impulse response used to generate RF data and post-processing.

rows) at each spatial position in the imaging field. These **H** matrices were applied to both simulate the echo data and develop the Wiener filters for post-processing. The 2D simulated impulse response shown in Figure 2.4 is recorded to form a vector that becomes one row of the system matrix **H**.

One motivation behind using the ideal observer approach is that it provides a fundamental measure of information content of image data specified for some considered task. In our research, the task comes from clinical application of ultrasonic imaging to discriminate malignant from benign lesions. And we have proposed five tasks in [6] based on features that trained experts are likely to look for in the process of detecting and discriminating malignant from benign lesions with ultrasound. In the next section, we investigate the development of cancer and variations of breast during malignant development to understand why the identified features play important roles in breast sonography diagnosis.

2.4 Object Modeling

2.4.1 Formation of malignant and benign lesions

We briefly describe the morphology of the breast first (see [16] for more details). The breast is a glandular organ with the glandular tissue embedded into a stroma, which includes fibroblast cells, smooth muscle cells, nerve cells and the extracellular matrix (ECM). The glandular tissue includes ductal trees that are composed of ducts and lobules. A typical tree consists of one

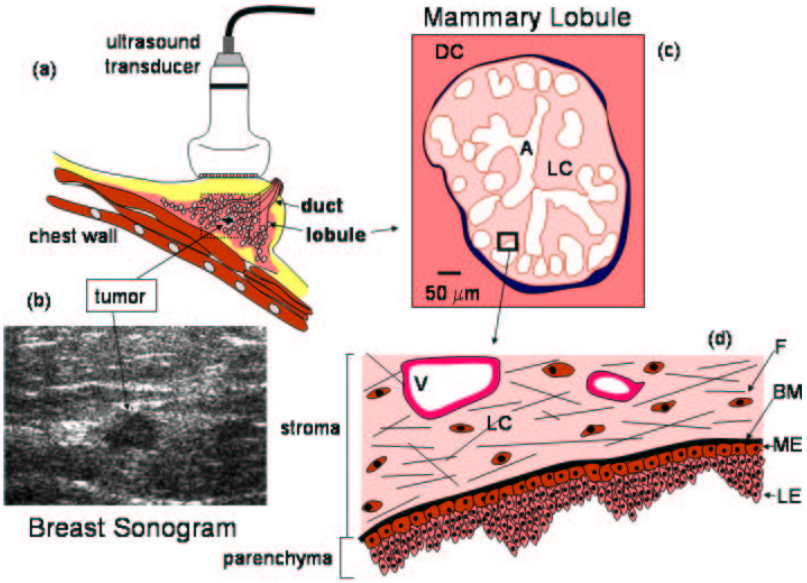

Figure 2.5: The gross and micro-anatomy of breast tissue: (a) Scanning the breast ultrasonically. (b) Sonogram. (c) The normal mammary lobule, in which A is an acini space, LC is the surrounding loose connective tissue stroma. (d) Basement membrane B separating LC from parenchymal tissues; LC contains fibroblast cells F and blood vessels V, parenchymal tissues include a layer of the myoepithelial cells ME and the luminal epithelium LE [17].

lactiferous duct starting from one opening on the nipple, branching into segmental, subsegmental and terminal ducts, each capped by lobules. The inner surface of ducts and lobules is covered by a single layer of luminal epithelial cells and a layer of myoepithelial cells. The main function of the epithelial cells is the secretion of milk while the myoepithelial cells are able to contract, increasing the pressure with the duct to push milk to the opening at the nipple.

The normal mammary lobule of the breast is shown in Figure 2.5(c),(d). The epithelial cells are separated from the stroma by the basement membrane (BM), which is a layer of extracellular collagen. The signaling molecules are attached on the BM for the interaction between epithelial cells and surrounding stroma. These molecules are essential for normal breast development and cyclic monthly maintenance. A normal organization is a balance between the tendency of the epithelium to stabilize morphometry by producing the BM and the stroma that induces structural changes by selectively eroding the BM ([17]).

The development of the breast cancer within a mammary lobule or duct is illustrated in Figure 2.6. Roughly 90% of all breast cancers begin in these tissues. Genetic alterations and regional environment changes to epithelial cells make them grow in an uncontrollable way. The BM, serving as a barrier between the epithelial cells and stroma, is degraded and finally broken

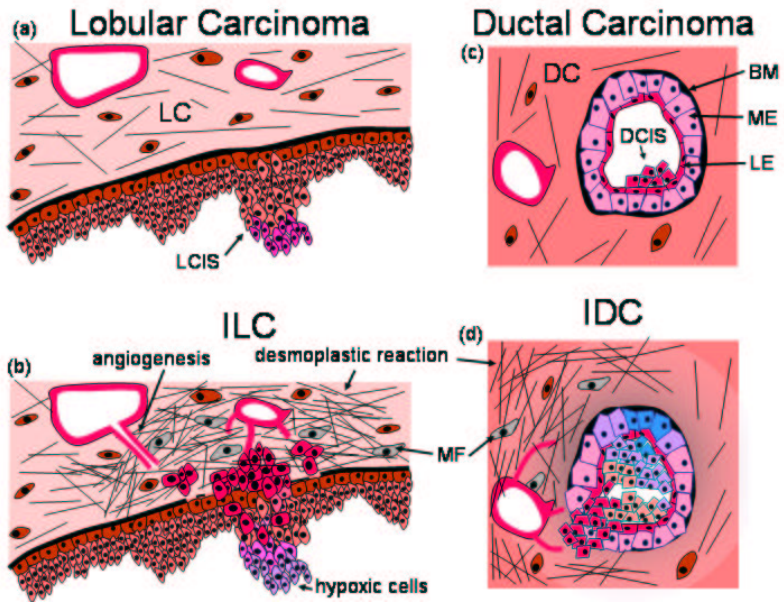

Figure 2.6: The progression of breast cancer in lobular (a,b) and ductal (c,d) [17].

down. The loss of the BM regulation initiates neoplastic transformation although the widespread breakdown of the BM is a histology indicator for malignant conversion. The physical contact between the epithelial cells with surrounding connective tissues promotes neovascularization, inflammation, and a structural remodeling for the ECM. The remodeling follows the formation of myofibroblasts from normal stroma fibroblasts. Ultrasonic methods allow us to observe the effects of this remodeling process even if we cannot resolve cells directly.

An ultrasonic scan of a malignant lesion is shown in Figure 2.7(a) with its boundary poorly defined and irregular. This is caused by the stroma surrounding the diseased duct or lobule trying to prevent the epithelial cells from expanding. The malignant lesion most often has a hypoechoic sonographic appearance. This fact can be explained by the edema of the tissues combined with changes in the collagen that scatters ultrasound.

The formation of a benign lesion is described in [16]. The processes of lobularization and arborization are sensitivity regulated, hormone-dependent. Changes in epithelial cells that disturb the hormonal influences may lead to delay or absence of hormone-related changes, which, in turn, may lead to aberrations or variations in morphology of the breast. This process is usually associated with the accumulation of fluid, milk, mucin, or cells within the lumen of the distended lobule. The process may result in structures that are detectable with ultrasound examinations. They are recognized as benign lesions although they represent the variations and aberrations of

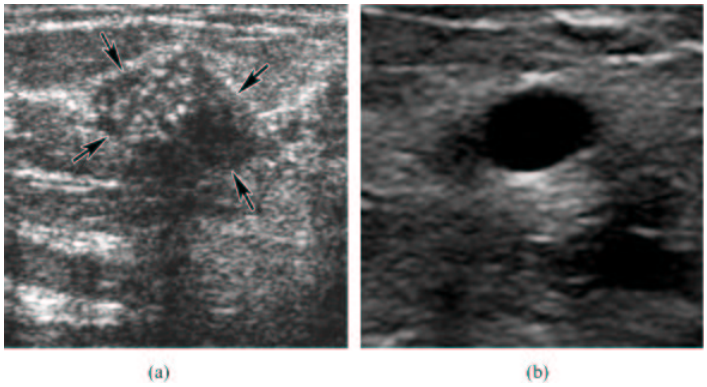

Figure 2.7: Ultrasonic scans of breast lesions that are (a) malignant, (b) cyst (benign).

the normal development and involution of the breast. The most common type of benign lesion is a cyst, which is illustrated in Figure 2.7(b). Since the cyst contains fluid, its sonogram has a sharply outlined round or oval mass, well-defined anterior, and anechoic interior, in contrast to the sonogram of the malignant lesion.

In the next section, we isolate the different sonographic features into five typical tasks to discriminate benign from malignant lesions.

2.4.2 Discrimination tasks

By consulting with a radiologist and from [18], we proposed a panel of five breast lesion features used for sonographic diagnosis in [6]. Those features became visual discrimination tasks by defining a malignant $\mathbf{S}_2$ and benign $\mathbf{S}_1$ matrix pair for each one of them. Listed in order of malignant and benign, Task 1 involves detecting a low-contrast hypoechoic lesion versus a no-lesion background; Task 2 requires discrimination of an elongated eccentric lesion from a circular lesion; Task 3 is discrimination of a soft, poorly defined boundary from a well-circumscribed lesion; Task 4 requires discrimination of spiculated boundary irregularities from a smooth circular boundary; and Task 5 involves discriminating a very weakly scattering hypoechoic interior from an anechoic interior. The variance maps of the five tasks are illustrated in Figure 2.8.

Another important discriminating sonographic feature between benign and malignant lesion is the prominent posterior shadow of malignant lesions as illustrated in Figure 2.7(a,b). We find this because the absorption of acoustic energy by the epithelial cells in malignant lesions is often greater than that in cystic fluid or other benign lesions. We do not consider the shadow feature in our research.

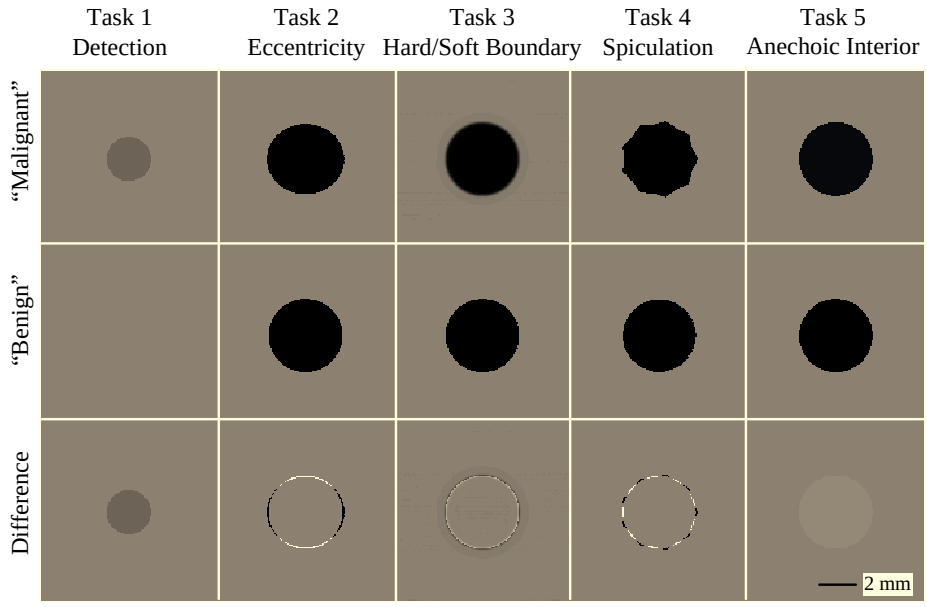

Figure 2.8: Variance profiles of "Malignant" and "Benign" scattering objects for five tasks. The lesion diameter is 3 mm in Task 1 and 5 mm in the other tasks [6].

2.5 Ideal Observer

Our principal interest is in the response of the ideal observer acting on the RF data, $\mathbf{g}$, as recorded immediately following beamformation where the diagnostic information content is greatest. Any post-detection processing may improve accessibility of information to human or model observers but it cannot increase the overall information content of the data. Task information is contained in the covariance matrices of the scattering object $\mathbf{f}$ given by $\sigma^2_{\text{obj}}(\mathbf{I} + \mathbf{S}_i)$, where σ^2_{obj} is the variance of the Gaussian scattering process and $\mathbf{S}_i$ is a diagonal matrix for the ith class of data included in the binary decision process. We set the diagonal elements of $\mathbf{S}_i$ to specify feature(s) under hypotheses H_i that define the task as in the first two rows of Figure 2.8. The difference between two variance maps is displayed in the bottom row, and it illustrates the contrast between the two hypotheses.

Multivariate normal processes remain multivariate normal following linear transformations. So the zero-mean MVN process defining $\mathbf{f}$ becomes for $\mathbf{g}$ under the two hypotheses, H_1 and H_2,

$$\begin{aligned} \text{H}_1 : \ \mathbf{g} &\sim \text{MVN}\ (\mathbf{0}, \boldsymbol{\Sigma}_1) \\ \text{H}_2 : \ \mathbf{g} &\sim \text{MVN}\ (\mathbf{0}, \boldsymbol{\Sigma}_2)\ . \end{aligned} \tag{2.6}$$

Specifically, the pdf of $\mathbf{g}$ under each hypothesis is given as

$$p(\mathbf{g}|\mathrm{H}_i) = \frac{1}{(2\pi)^{N/2}|\mathbf{\Sigma}_i|^{1/2}} \exp\left(-\frac{1}{2}\mathbf{g}^t\mathbf{\Sigma}_i^{-1}\mathbf{g}\right), \tag{2.7}$$

where N is the total number of samples in the acquisition frame and the size of column vector $\mathbf{g}$, $N = LM$, and the variances is given as

$$\mathbf{\Sigma}_i = \sigma_{\mathrm{obj}}^2\mathbf{H}(\mathbf{I} + \mathbf{S}_i)\mathbf{H}^t + \sigma_n^2\mathbf{I}\,, \text{ for } i = 1, 2. \tag{2.8}$$

The ideal observer is based on the minimum-error decision rule, which tells us to choose H_2 (the malignant case) when

$$\Lambda(\mathbf{g}) = \frac{p(\mathbf{g}|\mathrm{H}_2)}{p(\mathbf{g}|\mathrm{H}_1)} > \frac{\Pr(\mathrm{H}_1)}{\Pr(\mathrm{H}_2)}\,, \tag{2.9}$$

where $\Pr(\mathrm{H}_1)$ and $\Pr(\mathrm{H}_2)$ are prior probabilities for H_1 and H_2 without any information from $\mathbf{g}$. Typically, $\Pr(\mathrm{H}_1) = \Pr(\mathrm{H}_2) = 0.5$.

The likelihood ratio $\Lambda(\mathbf{g})$ is written as

$$\Lambda(\mathbf{g}) = \frac{|\mathbf{\Sigma}_1|^{1/2}}{|\mathbf{\Sigma}_2|^{1/2}} \frac{\exp\left(-\frac{1}{2}\mathbf{g}^t\mathbf{\Sigma}_2^{-1}\mathbf{g}\right)}{\exp\left(-\frac{1}{2}\mathbf{g}^t\mathbf{\Sigma}_1^{-1}\mathbf{g}\right)}. \tag{2.10}$$

The log-likelihood ration is given by

$$\lambda(\mathbf{g}) = \frac{1}{2}\mathbf{g}^t(\mathbf{\Sigma}_1^{-1} - \mathbf{\Sigma}_2^{-1})\mathbf{g} + \frac{1}{2}(\log(|\mathbf{\Sigma}_1|) - \log(|\mathbf{\Sigma}_2|)). \tag{2.11}$$

Combining terms and factors not related to the data into the threshold, we obtain the test statistic quantifying the ideal-observer response, which is given as [19]

$$\lambda(\mathbf{g}) \cong \mathbf{g}^t(\mathbf{\Sigma}_1^{-1} - \mathbf{\Sigma}_2^{-1})\mathbf{g} \leftrightarrow t\,. \tag{2.12}$$

The test statistic λ is a scalar valued quantity that is compared with threshold t to decide between the two hypotheses. Values larger than the threshold indicate H_2 is more likely, while lower values indicate H_1 is more likely. An observer who adopts this strategy uses data in an optimal way to maximize the area under the ROC curve, and thus is referred to as the *ideal observer*.

There are several ways to quantify ideal performance. One convenient way is to define the

lesion SNR from the means $\overline{\lambda}_i$ and variances σ_i^2 of the log-likelihood ratio under two hypotheses:

$$\mathrm{SNR}_I = \frac{\overline{\lambda}_2 - \overline{\lambda}_1}{\sqrt{\frac{1}{2}\left(\sigma_2^2 + \sigma_1^2\right)}}. \tag{2.13}$$

The SNR_I is closely related with the area under ROC for normal distributions of log-likelihood test statistics λ and can be calculated in closed-form in some simple situations [20, 21].

2.6 Human Observer

In medical imaging, the decision maker is the human observer, whose performance depends on diagnostic information available from the observed images, any additional post-processing, display algorithms, and intrinsic limitations including training and internal noises. Therefore, maximizing the diagnostic information in image data does not always mean maximizing human observer performance. Differences between human and ideal observers suggest that there may be a post-processing algorithm that enhances accessibility of the information to humans and reduces the performance gap.

The human observer is evaluated through psychophysical studies. We choose the two-alternative-force-choice method (2AFC) in this research. Performance measured in terms of proportion of correct can be connected directly to the area under the ROC curve [22, 23].

There are two key parameters related to the human observer performance [6]. The first parameter is the *object contrast factor* that is independently set by the study designer to challenge viewers appropriately. The object contrast factor quantitatively defines the difference between the variance maps in each task, which is the integrated absolute value of the variance deviation profiles,

$$C = A \sum_i \left| [\mathbf{S}_1 - \mathbf{S}_2]_{ii} \right| , \tag{2.14}$$

where $A = \Delta x \Delta y$ is the sampling interval area for RF data. C controls the difficulty of the task presented to the viewer (increasing this parameter makes the task easier to perform) so that we can challenge viewers equivalently for all tasks.

The second parameter used in the research is *observer efficiency*. For a system where the ideal observer can be computed, the efficiency of the human observer relative to the ideal observer is defined as [22]

$$\eta = \frac{\mathrm{SNR}_{human}^2}{\mathrm{SNR}_I^2}, \tag{2.15}$$

where SNR_I is defined as in Eq. (2.13). SNR_{human} is also known as the detectability index

d_{A}, which is found from the proportion of correct responses P_C in a 2AFC experiment [3]. For normally distributed responses, the relationship involves an error function and is given by

$$d_A = 2\mathrm{erf}^{-1}(2P_C - 1). \tag{2.16}$$

When human and ideal performance are comparable, the efficiency approaches one and in that case, the human observer is able to use all available diagnostic information in the data to perform the visual task. If the efficiency is much less than one, the human observer is inefficient at extracting the relevant information in the data to perform the task, and post-processing designed to enhance the human observer performance is recommended.

Human observer studies are expensive and time consuming. We adopt a computational observer, the Smith-Wagner (SW) observer [6], to represent human responses in order to evaluate efficiency. We use the SW observer to narrow down the list of post-processing approaches, and eventually perform human observer experiments as the ultimate measure of efficiency.

2.7 Smith-Wagner Observer

In [4, 5], Smith, Wagner and their co-workers have derived the ideal observer acting on B-mode images. The observer was implemented in the context of a 2AFC psychophysics study where signals are known exactly, and under assumptions of low-contrast large lesions, fully developed speckles and no additive noise at the output. In 2AFC experiments for each task, an observer must decide between two hypotheses:

H_1 : the lesion in image 1 is malignant and the lesion in image 2 is benign or

H_2 : the lesion in image 1 is benign and the lesion in image 2 is malignant.

Denoting the data of image 1 as $\mathbf{b}_1$ and the data of image 2 as $\mathbf{b}_2$, a decision function for the ideal SW observer is the ratio of a likelihood function, given as

$$\Lambda_{\mathrm{SW}}(\mathbf{b}_1, \mathbf{b}_2) = \frac{p(\mathbf{b}_1, \mathbf{b}_2|\mathrm{H}_2)}{p(\mathbf{b}_1, \mathbf{b}_2|\mathrm{H}_1)}, \tag{2.17}$$

where $p(\mathbf{b}_1, \mathbf{b}_1|\mathrm{H}_i)$ is the joint conditional pdf of the data of images $\mathbf{b}_1$ and $\mathbf{b}_2$. With the assumption of complex, multivariate, Gaussian random process for RF data, the B-mode image

data have a Rayleigh distribution [4], given by

$$p(\mathbf{b}_1, \mathbf{b}_2|\mathrm{H}_1) = \prod_{i=1}^{S} \frac{b_{1i}}{\psi_m} \exp\left(\frac{-b_{1i}^2}{2\psi_m}\right) \prod_{j=1}^{S} \frac{b_{2j}}{\psi_b} \exp\left(\frac{-b_{2j}^2}{2\psi_b}\right)$$
$$p(\mathbf{b}_1, \mathbf{b}_2|\mathrm{H}_2) = \prod_{i=1}^{S} \frac{b_{1i}}{\psi_b} \exp\left(\frac{-b_{1i}^2}{2\psi_b}\right) \prod_{j=1}^{S} \frac{b_{2j}}{\psi_m} \exp\left(\frac{-b_{2j}^2}{2\psi_m}\right), \tag{2.18}$$

where the parameters $2\psi_m$ and $2\psi_b$ are the mean backscattering intensities respectively for the malignant lesion image and the benign lesion image. S is the number of independent data speckles in the target area available for the decision maker. This number is different from the size N of the column vector data $\mathbf{b}$ [4].

Substituting Eq. (2.18) into Eq. (2.17), we find

$$\Lambda_{\mathrm{SW}}(\mathbf{b}_1, \mathbf{b}_2) = \prod_{i=1}^{S} \exp\left[\frac{b_{1i}^2}{2}\left(\frac{1}{\psi_m} - \frac{1}{\psi_b}\right)\right] \prod_{j=1}^{S} \exp\left[\frac{-b_{2j}^2}{2}\left(\frac{1}{\psi_m} - \frac{1}{\psi_b}\right)\right]. \tag{2.19}$$

Taking the log of the likelihood ratio we obtain,

$$\begin{aligned} \lambda_{\mathrm{SW}}(\mathbf{b}_1, \mathbf{b}_2) &= \left(\frac{\psi_b - \psi_m}{2\psi_m\psi_b}\right)\left(\sum_{i=1}^{S} b_{1i}^2 - \sum_{j=1}^{S} b_{2j}^2\right) \\ &= \mathbf{a}^t\mathbf{b}_1^2 - \mathbf{a}^t\mathbf{b}_2^2, \end{aligned} \tag{2.20}$$

where the sums in the first equation in Eq. (2.20) are taken over all independent speckle spots inside each lesion. The second equation represents the log-likelihood in terms of energies of whole images, in which $\mathbf{a}$ is a weighted vector of a_i's, $a_i = 0$ for pixel i outside the lesion, and $a_i = (\psi_b - \psi_m)/(2\psi_m\psi_b)$ for pixel i inside the lesion.

The decision then is made based on the comparison

$$\lambda_{\mathrm{SW}}(\mathbf{b}_1, \mathbf{b}_2)) = \mathbf{a}^t\mathbf{b}_1^2 - \mathbf{a}^t\mathbf{b}_2^2 \gtrless 0. \tag{2.21}$$

The decision favors hypothesis H_2 for positive values and H_1 otherwise. The comparison is equivalent to

$$\mathbf{a}^t\mathbf{b}_1^2 \gtrless \mathbf{a}^t\mathbf{b}_2^2. \tag{2.22}$$

So instead of calculating the test statistic decision variable as in Eq. (2.21), we calculate each sum separately for each image and compare the two. The SW observer is the exact ideal observer for task 1 on B-mode images. To apply it to all five tasks, we approximate the SW observer by

[6],

$$\lambda_{\mathrm{SW}}(\mathbf{b}) = \mathbf{b}^t(\mathbf{S}_2 - \mathbf{S}_1)\mathbf{b}. \tag{2.23}$$

The model is conjectured from some observers modeling for the human response [2], although it must be said that there has been no theoretical basis for doing so. With the use of the Smith-Wagner observer as the ideal observer acting on the envelop image, we can decompose visual detection into two stages,

$$\eta_{\mathrm{H}} = \eta_{\mathrm{SW}} \times \eta_{\mathrm{H|SW}} \ , \tag{2.24}$$

where the term η_{SW} is the efficiency of the SW observer acting on envelope images with respect to the ideal observer acting on RF data, and $\eta_{\mathrm{H|SW}}$ is the efficiency of human observer with respect to the Smith-Wagner observer acting on B-mode images. Separating efficiency in this manner enables us to identify the sources of information loss. η_{SW} is considered as a measurement of information lost by the demodulation process, and $\eta_{\mathrm{H|SW}}$ is a measurement of information loss by the human observer.

2.8 Summary

In this chapter, we have described the ultrasonic system as a linear model, and numerically modeled the signals of the system. By analyzing various features of sonograms that distinguish benign and malignant lesions, we have yielded a panel of five typical discrimination tasks, and classified the input signal (i.e., the scattering object) into two hypotheses for each task. The hypotheses of the scattering object lead to the hypotheses for the RF echo-signal, used to derive the ideal object acting on RF domain. The ideal observer is explored in the next chapter to find an optimal post-processing strategy for each task. We have also developed tools, such as the Smith-Wagner observer, to measure the visual detection efficiency of the human observer at performing tasks representative of those involved in clinical diagnosis. In the research, we are not measuring human detection efficiency for medical diagnostic tasks under clinical situations. We control the tasks absolutely and provide viewers with significant prior information not available clinically, so as to specifically measure the advantages of various signal processing approaches that affect performance.

CHAPTER 3

METHODS

In the previous chapter, we have described the ultrasound system as a linear system and the input as a nonstationary, Gaussian stochastic process. We then derived the ideal observer for binary discrimination from data obtained in the RF domain. The ideal observer's strategy is to square the RF data recorded by the instrument and multiply by the difference between the two inverses of the covariance matrices. This processes yields a scalar test statistic that is compared to a threshold value when making a decision. The test statistic is often too difficult to compute analytically or numerically, so approximations are sought. In this chapter, we explore approximations to the ideal observer strategy for discrimination to find signal processing techniques that are easily computed and yet approach ideal performance.

3.1 Power Series Approach

Although the exact test statistic λ is well defined by Eq. (2.12), it is prohibitively expensive to compute because of the need to invert two very large covariance matrices. We showed previously [6] that λ can be expanded in a power series for typical imaging conditions. The power series expansion used to approximate a matrix inverse is [24]

$$(\mathbf{I} - \mathbf{A})^{-1} = \sum_{k=0}^{\infty} \mathbf{A}^k, \tag{3.1}$$

which holds if the eigenvalues of $\mathbf{A}$ are between -1 and 1.

To apply the power series, we must assume a linear shift invariant system and that $\mathbf{H}$ is a circulant matrix, which allows us to diagonalize $\mathbf{H}$ by a Fourier transformation,

$$\mathbf{H} = \mathbf{F}^{-1}\mathbf{T}\mathbf{F}. \tag{3.2}$$

$\mathbf{F}$ is the 2-D forward discrete Fourier transform matrix, and $\mathbf{T}$ is a diagonal matrix whose elements are the eigenvalues of $\mathbf{H}$.

Under these assumptions, we were able to separate stationary from nonstationary components of $\mathbf{\Sigma}_i$ from Eq. (2.8) using

$$\begin{aligned}\mathbf{\Sigma}_i &= \sigma_{\text{obj}}^2\mathbf{H}(\mathbf{I}+\mathbf{S}_i)\mathbf{H}^t+\sigma_n^2\mathbf{I}\\ &= \sigma_{\text{obj}}^2\mathbf{H}\mathbf{H}^t+\sigma_n^2\mathbf{I}+\sigma_{\text{obj}}^2\mathbf{H}\mathbf{S}_i\mathbf{H}^t\\ &= \mathbf{\Sigma}_0+\Delta\mathbf{\Sigma}_i,\end{aligned} \tag{3.3}$$

where $\mathbf{\Sigma}_0=\sigma_{\text{obj}}^2\mathbf{H}\mathbf{H}^t+\sigma_n^2\mathbf{I}$ is the stationary component and $\Delta\mathbf{\Sigma}_i=\sigma_{\text{obj}}^2\mathbf{H}\mathbf{S}_i\mathbf{H}^t$ is the nonstationary component.

Because $\mathbf{\Sigma}_0$ is stationary, it can be decomposed by using Fourier techniques as follows:

$$\mathbf{\Sigma}_0=\mathbf{F}^{-1}\mathbf{N}_0\mathbf{F}, \tag{3.4}$$

where $\mathbf{N}_0$ is a diagonal matrix whose elements are

$$[\mathbf{N}_0]_{ii}=\sigma_{\text{obj}}^2\left|[\mathbf{T}]_{ii}\right|^2+\sigma_n^2. \tag{3.5}$$

$\mathbf{T}$ is the diagonal matrix from the decomposition of $\mathbf{H}$ defined in Eq (3.2). Therefore $\mathbf{\Sigma}_0$ can be inverted easily by using the Fourier technique as

$$\mathbf{\Sigma}_0^{-1}=\mathbf{F}^{-1}\mathbf{N}_0^{-1}\mathbf{F}. \tag{3.6}$$

When this is combined with the nonstationary part $\Delta\mathbf{\Sigma}_i$, we can apply a power series expansion to calculate the matrix inverse. First, we write the covariance matrix in the form

$$\mathbf{\Sigma}_i=\mathbf{\Sigma}_0^{1/2}\left(\mathbf{I}+\mathbf{\Sigma}_0^{-1/2}\Delta\mathbf{\Sigma}_i\mathbf{\Sigma}_0^{-1/2}\right)\mathbf{\Sigma}_0^{1/2}. \tag{3.7}$$

The power series expansion gives the inverses as

$$\begin{aligned}\mathbf{\Sigma}_i^{-1} &= \mathbf{\Sigma}_0^{-1/2}\left(\mathbf{I}+\mathbf{\Sigma}_0^{-1/2}\Delta\mathbf{\Sigma}_i\mathbf{\Sigma}_0^{-1/2}\right)^{-1}\mathbf{\Sigma}_0^{-1/2}\\ &= \mathbf{\Sigma}_0^{-1/2}\left(\sum_{k=0}^{\infty}\left(-\mathbf{\Sigma}_0^{-1/2}\Delta\mathbf{\Sigma}_i\mathbf{\Sigma}_0^{-1/2}\right)^k\right)\mathbf{\Sigma}_0^{-1/2}.\end{aligned} \tag{3.8}$$

The power series approach is feasible only if eigenvalues of matrix $\mathbf{\Sigma}_0^{-1/2}\Delta\mathbf{\Sigma}_i\mathbf{\Sigma}_0^{-1/2}$ have values between -1 and 1. To study this condition further, we use the definition of the 2-norm of a matrix $\mathbf{A}$ in [24],

$$\|\mathbf{A}\|_2=\sqrt{\text{the largest eigenvalue of } \mathbf{A}^t\mathbf{A}}. \tag{3.9}$$

Thus if σ is the eigenvalue of $\mathbf{A}$ which has the largest absolute value, then $\|\mathbf{A}\|_2 = |\sigma|$.

Using the equation for the nonstationary part in (3.3), we have

$$\begin{aligned} \left\|\boldsymbol{\Sigma}_0^{-1/2}\Delta\boldsymbol{\Sigma}_i\boldsymbol{\Sigma}_0^{-1/2}\right\|_2 &= \left\|\boldsymbol{\Sigma}_0^{-1/2}\sigma_{\text{obj}}^2\mathbf{H}\mathbf{S}_i\mathbf{H}^t\boldsymbol{\Sigma}_0^{-1/2}\right\|_2 \\ &\leq \sigma_{\text{obj}}^2\left\|\boldsymbol{\Sigma}_0^{-1/2}\mathbf{H}\right\|_2\|\mathbf{S}_i\|_2\left\|\mathbf{H}^t\boldsymbol{\Sigma}_0^{-1/2}\right\|_2, \end{aligned} \tag{3.10}$$

with the later inequality coming from a property of the norm [24]. Combining the decompositions of $\boldsymbol{\Sigma}_0$ and $\mathbf{H}$ in Eqs. (3.2) and (3.4), and noting that $\boldsymbol{\Sigma}_0^{-1/2} = \mathbf{F}^{-1}\mathbf{N}_0^{-1/2}\mathbf{F}$, we obtain

$$\boldsymbol{\Sigma}_0^{-1/2}\mathbf{H} = \mathbf{F}^{-1}\mathbf{K}\mathbf{F}, \tag{3.11}$$

where $\mathbf{K}$ is a diagonal matrix with its elements given by

$$[\mathbf{K}]_{ii} = \frac{[\mathbf{T}]_{ii}}{\sqrt{\sigma_{\text{obj}}^2\left|[\mathbf{T}]_{ii}\right|^2 + \sigma_n^2}}. \tag{3.12}$$

By denoting $\sigma_h = \|\mathbf{H}\|_2 = \max_i\left|[\mathbf{T}]_{ii}\right|$, combining with the fact that the function $f(x) = x/\sqrt{x^2+1}$ is strictly increasing, we find

$$\left\|\boldsymbol{\Sigma}_0^{-1/2}\mathbf{H}\right\|_2 = \frac{\sigma_h}{\sqrt{\sigma_{\text{obj}}^2\sigma_h^2 + \sigma_n^2}}. \tag{3.13}$$

Similarly, we have

$$\left\|\mathbf{H}^t\boldsymbol{\Sigma}_0^{-1/2}\right\|_2 = \frac{\sigma_h}{\sqrt{\sigma_{\text{obj}}^2\sigma_h^2 + \sigma_n^2}}. \tag{3.14}$$

As described in Section 2.5, $\mathbf{S}_i$ is a normalized diagonal matrix that encodes nonstationary information defining the task as spatial variations in the scattering object. Therefore $\|\mathbf{S}_i\|_2 = 1$. Using it for the inequality in Eq. (3.10), and combining with Eqs. (3.13), (3.14), we find

$$\begin{aligned} \left\|\boldsymbol{\Sigma}_0^{-1/2}\Delta\boldsymbol{\Sigma}_i\boldsymbol{\Sigma}_0^{-1/2}\right\|_2 &\leq \sigma_{\text{obj}}^2\left\|\boldsymbol{\Sigma}_0^{-1/2}\mathbf{H}\right\|_2\|\mathbf{S}_i\|_2\left\|\mathbf{H}^t\boldsymbol{\Sigma}_0^{-1/2}\right\|_2 \\ &= \frac{\sigma_{\text{obj}}^2\sigma_h^2}{\sigma_{\text{obj}}^2\sigma_h^2 + \sigma_n^2} < 1. \end{aligned} \tag{3.15}$$

Thus $\left\|\boldsymbol{\Sigma}_0^{-1/2}\Delta\boldsymbol{\Sigma}_i\boldsymbol{\Sigma}_0^{-1/2}\right\|_2 < 1$ provided that $\sigma_n^2 > 0$, so the power series in Eq. (3.8) converges. Thus the power series expansion can be used to calculate matrix inverses.

With the use of the power series, we write each component in the decision variable as

$$\begin{aligned}\mathbf{g}^t\boldsymbol{\Sigma}_i^{-1}\mathbf{g} &= \mathbf{g}^t\boldsymbol{\Sigma}_0^{-1/2}\left(\sum_{k=0}^{\infty}\left(-\boldsymbol{\Sigma}_0^{-1/2}\Delta\boldsymbol{\Sigma}_i\boldsymbol{\Sigma}_0^{-1/2}\right)^k\right)\boldsymbol{\Sigma}_0^{-1/2}\mathbf{g} \\ &= \mathbf{g}^t\left(\sum_{k=0}^{\infty}\left(-\boldsymbol{\Sigma}_0^{-1}\Delta\boldsymbol{\Sigma}_i\right)^k\right)\boldsymbol{\Sigma}_0^{-1}\mathbf{g}, \end{aligned} \tag{3.16}$$

which yields an iteration for calculation, given as

$$\begin{aligned}\mathbf{q}_{i+1} &= -\boldsymbol{\Sigma}_0^{-1}\Delta\boldsymbol{\Sigma}_i\mathbf{q}_i \text{ and} \\ \mathbf{p}_{i+1} &= \mathbf{p}_i + \mathbf{q}_{i+1}. \end{aligned} \tag{3.17}$$

The iteration is started with $\mathbf{p}_0 = \mathbf{q}_0 = \boldsymbol{\Sigma}_0^{-1}\mathbf{g}$.

In this research, each component is determined after 50 iterations, where the decision variable of the ideal observer is changed by less than 0.5% per iteration.

3.2 Wiener Filtering as Approximation to Ideal Strategy

By truncating the power series at $k = 1$, the inverse covariance term in Eq. (3.8) is given as

$$\boldsymbol{\Sigma}_i^{-1} \simeq \boldsymbol{\Sigma}_0^{-1} - \boldsymbol{\Sigma}_0^{-1}\Delta\boldsymbol{\Sigma}_i\boldsymbol{\Sigma}_0^{-1}\,. \tag{3.18}$$

Therefore $\boldsymbol{\Sigma}_1^{-1} - \boldsymbol{\Sigma}_2^{-1}$ is approximated by

$$\begin{aligned}\boldsymbol{\Sigma}_1^{-1} - \boldsymbol{\Sigma}_2^{-1} &\simeq \boldsymbol{\Sigma}_0^{-1}\left(\Delta\boldsymbol{\Sigma}_1 - \Delta\boldsymbol{\Sigma}_2\right)\boldsymbol{\Sigma}_0^{-1} \\ &= \sigma_{\text{obj}}^2\boldsymbol{\Sigma}_0^{-1}\mathbf{H}\left(\mathbf{S}_1 - \mathbf{S}_2\right)\mathbf{H}^t\boldsymbol{\Sigma}_0^{-1}\,. \end{aligned} \tag{3.19}$$

Substituting Eq (3.19) into Eq (2.12), we find

$$\lambda(\mathbf{g}) \approx \sigma_{\text{obj}}^2\mathbf{g}^t\boldsymbol{\Sigma}_0^{-1}\mathbf{H}(\mathbf{S}_2 - \mathbf{S}_1)\mathbf{H}^t\boldsymbol{\Sigma}_0^{-1}\mathbf{g}\,. \tag{3.20}$$

The matrix multiplication $\sigma_{\text{obj}}^2\mathbf{H}^t\boldsymbol{\Sigma}_0^{-1}\mathbf{g}$ is equivalent to a filtering of the RF data. Combining with the decompositions of $\mathbf{H}$ and $\boldsymbol{\Sigma}_0$ in Eqs. (3.2) and (3.4), we have $\mathbf{H}\boldsymbol{\Sigma}_0^{-1} = \mathbf{F}^{-1}\boldsymbol{\Phi}\mathbf{F}$, where $\boldsymbol{\Phi}$ is a diagonal matrix with its elements given by

$$[\boldsymbol{\Phi}]_{ii} = \frac{\sigma_{\text{obj}}^2\left[\mathbf{T}^t\right]_{ii}}{\sigma_{\text{obj}}^2\left|[\mathbf{T}]_{ii}\right|^2 + \sigma_n^2}. \tag{3.21}$$

$[\mathbf{T}]_{ii}$ are eigenvalues of $\mathbf{H}$ and also the Fourier transform of the system response. Thus $\sigma^2_{\text{obj}}\mathbf{H}^t\boldsymbol{\Sigma}_0^{-1}$ is recognized as the Wiener filter for the scattering object of a low-contrast lesion when its covariance matrix is approximated by $\sigma^2_{\text{obj}}\mathbf{I}$.

When the linear approximation holds, the ideal strategy is to Wiener filter the echo signals and multiply the squared result by the task, $\mathbf{S}_2 - \mathbf{S}_1$. It suggests that the RF data be Wiener filtered before computing the B-mode images. This strategy is investigated in [6] and was found to also improve human observer performance provided the assumptions for linear approximation are not violated (see the results in Section 4.1.1). The main goal of the current study is to observe consequences of high contrast tasks where the linear approximation may not hold. Based on the analysis, we propose an alternative procedure for the task when the Wiener filter fails to make an improvement in visual performance.

3.3 Iterative Wiener Filter

In this section, a method for adapting the Wiener filter to spatially varying SNR conditions is proposed. Rather than separating covariance matrix components into stationary and nonstationary components, we form *average* and *difference* components. First, the average variance profile $\mathbf{S}_a = 0.5(\mathbf{S}_1 + \mathbf{S}_2)$ and difference variance profile $\Delta\mathbf{S} = 0.5(\mathbf{S}_2 - \mathbf{S}_1)$ are found to define covariance components

$$\boldsymbol{\Sigma}_a = \sigma^2_{\text{obj}}\mathbf{H}(\mathbf{I} + \mathbf{S}_a)\mathbf{H}^t + \sigma^2_{\text{n}}\mathbf{I} \text{ and } \Delta\boldsymbol{\Sigma} = \sigma^2_{\text{obj}}\mathbf{H}\Delta\mathbf{S}\mathbf{H}^t \,. \tag{3.22}$$

Hence the covariances for the two class states are

$$\boldsymbol{\Sigma}_1 = \boldsymbol{\Sigma}_a - \Delta\boldsymbol{\Sigma} \text{ and } \boldsymbol{\Sigma}_2 = \boldsymbol{\Sigma}_a + \Delta\boldsymbol{\Sigma} \,. \tag{3.23}$$

Written as

$$\begin{aligned} \boldsymbol{\Sigma}_1 &= \boldsymbol{\Sigma}_a^{1/2}\left(\mathbf{I} - \boldsymbol{\Sigma}_a^{-1/2}\Delta\boldsymbol{\Sigma}\boldsymbol{\Sigma}_a^{-1/2}\right)\boldsymbol{\Sigma}_a^{1/2} \\ \boldsymbol{\Sigma}_2 &= \boldsymbol{\Sigma}_a^{1/2}\left(\mathbf{I} + \boldsymbol{\Sigma}_a^{-1/2}\Delta\boldsymbol{\Sigma}\boldsymbol{\Sigma}_a^{-1/2}\right)\boldsymbol{\Sigma}_a^{1/2} \,, \end{aligned} \tag{3.24}$$

the power series expansion gives the inverse covariances

$$\begin{aligned}\boldsymbol{\Sigma}_1^{-1} &= \boldsymbol{\Sigma}_a^{-1/2}\left(\sum_{k=0}^{\infty}\left(\boldsymbol{\Sigma}_a^{-1/2}\Delta\boldsymbol{\Sigma}\boldsymbol{\Sigma}_a^{-1/2}\right)^k\right)\boldsymbol{\Sigma}_a^{-1/2}\\ \boldsymbol{\Sigma}_2^{-1} &= \boldsymbol{\Sigma}_a^{-1/2}\left(\sum_{k=0}^{\infty}\left(-\boldsymbol{\Sigma}_a^{-1/2}\Delta\boldsymbol{\Sigma}\boldsymbol{\Sigma}_a^{-1/2}\right)^k\right)\boldsymbol{\Sigma}_a^{-1/2}.\end{aligned} \tag{3.25}$$

Truncating the power series at $k = 1$, we obtain

$$\boldsymbol{\Sigma}_1^{-1} - \boldsymbol{\Sigma}_2^{-1} \approx 2\boldsymbol{\Sigma}_a^{-1}\Delta\boldsymbol{\Sigma}\boldsymbol{\Sigma}_a^{-1}\ . \tag{3.26}$$

Combining (2.12), (3.23), and (3.26), the linear approximation to the ideal observer's test statistic becomes

$$\lambda(\mathbf{g}) \approx \mathbf{g}^t\boldsymbol{\Sigma}_a^{-1}\mathbf{H}(\mathbf{S}_2 - \mathbf{S}_1)\mathbf{H}^t\boldsymbol{\Sigma}_a^{-1}\mathbf{g}\ . \tag{3.27}$$

Equation (3.27) also suggests that the echo signal be filtered. However the filter $\mathbf{H}^t\boldsymbol{\Sigma}_a^{-1}$ uses the average covariance between the two states rather than the stationary background covariance. The advantage of this change is to allow the signal strength to vary significantly within any one image provided that differences between images to be compared remain small. A disadvantage of the new filter is that $\boldsymbol{\Sigma}_a$ is the covariance matrix of a nonstationary process, so we cannot use Fourier technique to compute its inverse.

The power series approach may be applied to decompose $\boldsymbol{\Sigma}_a$ into stationary and nonstationary components, $\boldsymbol{\Sigma}_a = \boldsymbol{\Sigma}_0 + \sigma_{\text{obj}}^2\mathbf{H}\mathbf{S}_a\mathbf{H}^t$, yielding an iterative formula for $\sigma_{\text{obj}}^2\mathbf{H}^t\boldsymbol{\Sigma}_a^{-1}\mathbf{g}$ given by

$$\begin{aligned}\mathbf{q}_{i+1} &= -\sigma_{\text{obj}}^2\mathbf{H}^t\boldsymbol{\Sigma}_0^{-1}\mathbf{H}\mathbf{S}_a\mathbf{q}_i \ \text{ and}\\ \mathbf{p}_{i+1} &= \mathbf{p}_i + \mathbf{q}_{i+1}\ .\end{aligned} \tag{3.28}$$

The iterative scheme is initialized by $\mathbf{q}_0 = \sigma_{\text{obj}}^2\mathbf{H}^t\boldsymbol{\Sigma}_0^{-1}\mathbf{g} = \mathbf{p}_0$.

Equation (3.28) begins with the Wiener filter, and through i iterations, $\mathbf{p}_{i+1}$ converges to $\sigma_{\text{obj}}^2\mathbf{H}^t\boldsymbol{\Sigma}_a^{-1}\mathbf{g}$. We refer to the procedure as forming an *iterative Wiener filter* for generating IWFB-mode ultrasonic images. The iterative Wiener filter adapts to the statistical properties of the RF data throughout the field as specified by $\mathbf{S}_a$. Therefore the process requires that we first estimate $\mathbf{S}_a$, which we can do by estimating $\mathbf{S}_1$ and $\mathbf{S}_2$ using image segmentation.

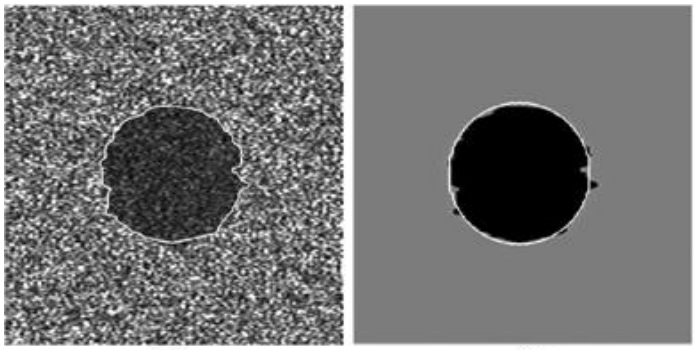

Figure 3.1: Segmentation of the variance profile from the Wiener filtered envelope (a) and its error to compare to the exact signal (b).

3.4 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, which was applied for images consisting of very few levels [25]. The algorithm begins from development of a model for the posterior probability density function $p(x|y)$, where y is the observed image segmented into regions x. At the site s of the image, $x_s = i$ means that the site belongs to region i. By Bayes' theorem, we have

$$p(x|y) = p(y|x)p(x), \tag{3.29}$$

where $p(x)$ is a prior density of the region process, which imposes spatial continuity, and $p(y|x)$ is the conditional density of the observed image given the distribution of the regions. It constrains the region magnitude to be close to that of the data. By using MRF, the density of x is given by a Gibbs density [26, 27], while the conditional density is modeled as a white Gaussian process with mean μ_i and variance σ_i^2 characterizing region i. The decision is made by MAP rule based on the combined probability density function $p(x|y)$. This is done pixel-by-pixel in the whole image and in many iterations until we converge on x. *"Adaptive"* means parameters μ_i and σ_i^2 are updated after each iteration.

The algorithm was first used to segment the ultrasound B-mode images by Ashton and Parker [28]. However, the magnitude of conventional B-mode images has a Rayleigh distribution, which does not allow use of this method directly. Therefore, the author first decomposed B-mode images into multiple layers (wavelet decomposition), and based on the central limit theorem, they assumed the intensity of pixels at the lowest level has a Gaussian distribution. Segmentation is then applied for each layer from the lowest to the highest resolution. This process is computa-

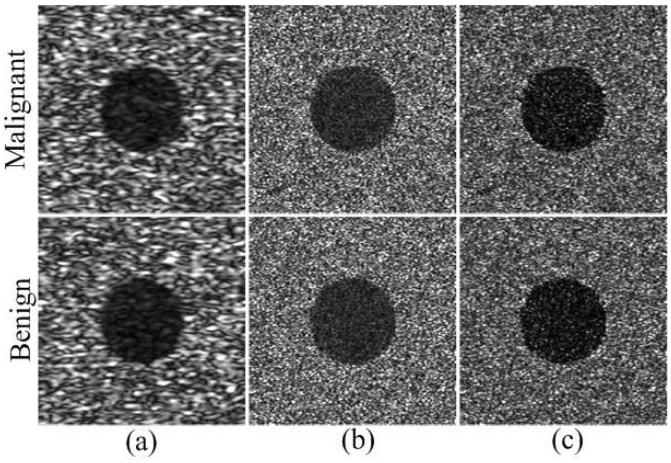

Figure 3.2: Examples of (a) B-mode, (b) WFB-mode and (c) IWFB-mode images for malignant and benign features of Task 5.

tionally intensive. In our application, we modified the process by segmenting the WFB-mode image instead. There are some advantages to using the Wiener filtered image despite any errors due to variable echo SNR. First, the Wiener filtered data approximates a deconvolution of RF data followed by B-mode rendering. The WFB-mode image has reduced pixel correlation and provides greater region delineation compared with B-mode images. Also, the logarithm of image values is approximately Gaussian distributed as required for segmentation purposes with an acceptable error. Thus, performing segmentation on WFB-mode images simplifies the complexity and shortens computation time. Figure 3.1(a) shows an example of the estimated shape of the variance profile obtained from segmentation and Fig 3.1(b) shows its errors incurred when comparing the segmented image with the exact signal. Segmentation errors reduce the effectiveness of IWFB-mode processing and therefore will adversely affect human observer performance.

The pixel whitening effects of the Wiener filter enable us to consider WFB-mode images as a coarse estimate of the scattering object, $f(\mathbf{x})$, so that we can roughly estimate σ^2_{obj} from the mean square pixel value of the background region. $\mathbf{S}_i$ may be estimated by measuring the mean-square image value inside the lesion and dividing it by the estimated $\hat{\sigma}^2_{\text{obj}}$. However, we modify the process slightly by just dividing the squared magnitude of each pixel inside the lesion by $\hat{\sigma}^2_{\text{obj}}$. This modification is to adapt to the statistical properties of speckled image data inside the lesion region.

Example of B-mode, WFB-mode, and IWFB-mode images for Task 5 are illustrated in Figure 3.2. In this example, the benign-image lesion area contains noise and a cyst, while the malignant-image lesion area contains weak scattering echo and noise that makes it nonuniformly

brighter than the benign one.

CHAPTER 4

RESULTS AND DISCUSSION

In this chapter, we evaluate the effects of the Wiener and iterative Wiener filters over the panel of the five tasks, based on the ideal and human observers described in Chapter 2. In [6], where the system response is modeled by a Gabor pulse, we have found that the Wiener filter makes an improvement in four of the five tasks, and yet it reduces performance for task 5. In this chapter, we evaluate the Wiener filter with a more realistic pulse, generated by the Field II program [10, 11], and the new iterative Wiener filter. Throughout the results, we try to understand the failure of the Wiener filter in Task 5 and then explain the adaptivity of the new filter for enhancing performance.

4.1 Human Observer

4.1.1 Previous results [6]

The application of the Wiener filter in the RF domain has been investigated in our previous study with the system impulse response modeled by a Gabor pulse function. Six observers participated in each of the five visual detection tasks by viewing both B-mode and WFB-mode images. To control for case variability in each task, B-mode and WFB-mode were generated from the same RF data. The studies were performed in a darkened room on a calibrated monochrome monitor (Image Systems, Minnetonka, MN) using a perceptually linearized lookup table over a luminance range of 0.5 to 150 cd/m^2. The monitor pixel size was 0.3 mm. Viewing distance was not controlled, but generally ranged from 40 to 50 cm.

Human observer performance is charted in Figure 4.1. There is a substantial increase in human observer performance with the WFB-mode images for Tasks 1-4. Performance for Tasks 2 and 3 increases by approximately 10 percent, while for Tasks 1 and 4 the increase was as much as 16 percent. However, performance decreases in Task 5 for WFB-mode images relative to the B-mode images, from around 81% to 73%. This reduction is discussed in the next sections with the more realistic simulated impulse response. Using the Gabor pulse for simulation ignores common

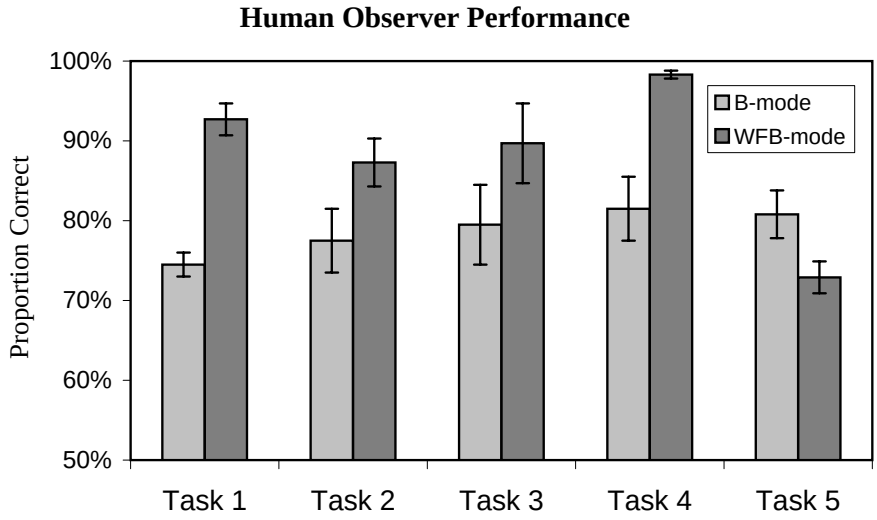

Figure 4.1: Human observer performance with system impulse response modeled by a Gabor pulse [6].

conditions that generate strong sidelobes and/or grating lobes for system impulse response.

4.1.2 Human observer performance

In current research, the human observer experiments are conducted at the Psychophysics Lab, University of California at Santa Barbara, under the direction of our collaborator, Professor Craig Abbey. Five observers participated in a total of 15 studies, which investigate each of the 5 visual tasks with B-mode images, WFB-mode images, and IWFB-mode images. In each task, performance is found by averaging 2AFC responses from observers viewing 400 pairs of images. A pilot study is taken for adjusting the contrast so that the average performance on B-mode images in each task is around 70% - 80% [3]. Since the range of AFC response is from 50% to 100%, setting the average performance at moderate levels allows us to fully understand the filters' effect. For the IWFB-mode images, the iterative process is calculated until 20 iterations. After 20 iterations, the increments in each iteration vary negligibly (less than 0.1%). The rendered images for viewing are square images of 300 pixels, with pixel size 0.04 mm. The lesion inside the images has a diameter of 3 mm for Task 1, and 5 mm for Tasks 2-5. The impulse response is generated by the Field II program illustrated in Figure 2.4.

Human observer results are summarized in Figure 4.2. In the first four tasks, we find that humans viewing B-mode images yield the lowest percent correct for the 2AFC experiments. Both filters generate images that help increase the performance. However, the WF fails to improve

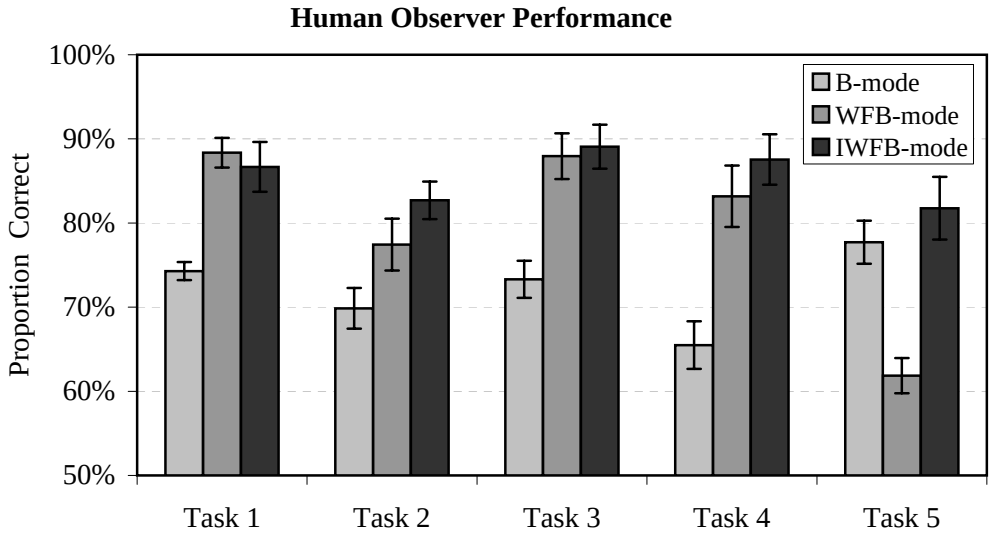

Figure 4.2: Human observer performance for B-mode and filtered B-mode images with a more realistic impulse response.

Task 5 performance. It decreases from 79% to 63%. Combined with the results with Gabor pulse in the previous section, we have replicated the previous results but for a more realistic pulse profile with system parameters. The IWF makes improvements in all five tasks. In Task 5, the IWF makes a big improvement compared with the WF, increasing from 63% to 82%. On the other tasks, the performance of IWFB-mode images is on par with that of the WFB-mode images.

4.1.3 Discussion

The effect of the WF at influencing performance for the five tasks demonstrates a common trade-off in ultrasonic image quality between spatial resolution and contrast resolution. Increased spatial resolution often occurs at the expense of contrast resolution and vice versa [29]. With weak additive system noise, the Wiener filter acts as a high-pass filter, and thus improves the spatial resolution. Using the definition of speckle size in [4], we find the WF reduces speckle size approximately 3.5 times (from 0.0648 mm^2 for standard B-mode images to 0.018 mm^2 for the WFB-mode images). Smaller speckles sharpen the edge of the lesion (as shown in Figure 3.2), facilitating lesion feature detection. High frequency enhancement coupled with high spatial frequency tasks explains why the visual performance in the first four tasks is improved by using the Wiener filter. In Task 5, the Wiener filter still enhances spatial resolution but does not aid observers at detection. The diagnostic information is the average contrast lesion,

which is a low spatial frequency tasks. The process enhances high frequency at the expense of low frequency, which contains detection information, and therefore the Wiener filter reduces detection performance for Task 5.

The IWF is a WF combined with an iterative process to tune the contrast inside the lesion. The iterative process of the filter increases the contrast resolution of the two lesions while it still improves the spatial resolution. That is why the IWF can work not only for the first four tasks but also for Task 5 where the WF failed. A factor that can reduce the human observer performance on the IWFB-mode images is the segmentation errors, especially in Task 3, where the soft, poorly defined boundary of the lesion conflicts with the two-level assumption of the segmentation algorithm. However, the results show that human observer performance on the IWFB-mode images is as good as that on WFB-mode images even in Task 3. This can be explained by the segmented version of the soft boundary malignant lesion becoming smaller and more spiculated, making it easy to discriminate from the benign lesion that is larger and rounder. A disadvantage of the IWF is the long computation time for both the segmentation and iterations (it takes approximately 20 sec to calculate one IWF version on a PC with 2.13 GHz processor, 2 Gb RAM).

It is noted that the failure of the WFB-mode images for Task 5 is due to the poor approximation of the first-order expansion of the covariance matrices. The success of IWB-mode images is due to an improvement in the approximation for the ideal observer response.

4.2 Ideal Observer

4.2.1 Results

In the context of the 2AFC method, the calculation of the ideal observer is slightly different from what we presented in Section 2.5. Instead of calculating the ideal observer response for one image and then comparing with a threshold, we calculate the responses for one pair of images, containing both benign and malignant lesions. The responses are then compared with each other and the result is scored by a step function. In our simulation, 2000 pairs of RF echo frames were generated for each task. Each pair was an independent realization of the random process. If we let $\mathbf{g}_{1i}$ be the RF frame from H_1 in the ith trial and $\mathbf{g}_{2i}$ be the RF frame for H_2, then the trial score, o_i, is given by

$$o_i = \mathrm{step}\left(\lambda\left(\mathbf{g}_{2i}\right) - \lambda\left(\mathbf{g}_{1i}\right)\right), \tag{4.1}$$

where the step function has a value of 1 for positive argument and 0 for negative. Thus, assuming that the malignant lesion is generated under hypothesis H_2, the score is 1 if $\lambda\left(\mathbf{g}_{2i}\right) > \lambda\left(\mathbf{g}_{1i}\right)$,

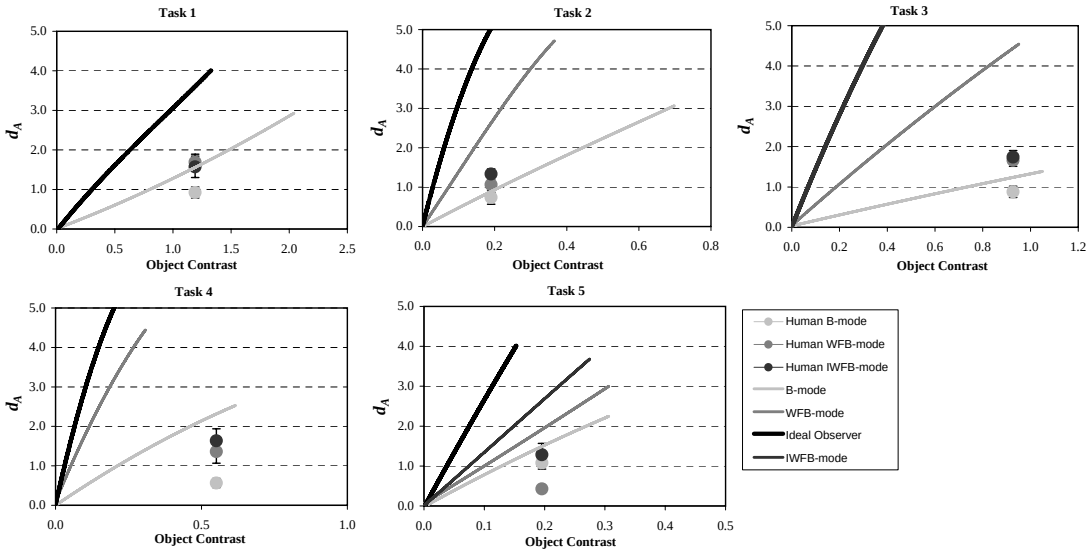

Figure 4.3: The ideal observer performance represented in terms of the detectability index d_a is plotted as a function of object contrast. The Smith-Wagner observer performance on envelope images is also plotted, as well as the human observer performance. The legend next to the plot of Task 5 is applied for all five tasks.

and 0 for otherwise. The proportion of correct, P_C, is defined as the average of the score trials. Through P_C, we calculate the detectability index as in Eq. (2.16).

The detectability index d_A is calculated and plotted as a function of the object contrast for all five tasks, as shown in Figure 4.3. In the figure, d_A values for the Smith-Wagner observer are plotted for comparison. In Task 1, d_A's of the ideal observer, the SW observer on WFB-mode images and IWFB-mode images are identical, so only one for the ideal observer is plotted. From Task 2 to Task 4, the SW observers on WFB-mode images and IWFB-mode images are the same, so only the one for the WFB-mode images is plotted. In Task 5, the detectability indices are totally different and plotted as four separate curves in the figure. The legend next to the plot of Task 5 is used for all five tasks.

4.2.2 Discussion

As we described in Chapter 2, the ideal observer uses all of the task-relevant information content of the data to make the decision; therefore, its performance is optimal. The Wiener filters, derived from linear approximations of the ideal observer, yield optimal performance if the linear approximations are accurate. In this section, we discuss how accuracies of the linear approximations are related to the filters' performance. In the plots of Figure 4.3, the greater slope

suggests the corresponding data has more diagnostic information. The RF data has the greatest slope in all five plots because it contains the most diagnostic information, including the phase information which is inaccessible to the human observer. However, in Task 1, the three d_A's for ideal observer, SW observers of WFB-mode and IWFB-mode images are identical, suggesting that the step function in Eq. (4.1) reaches stable values just after one iteration. The reason is the elements in $\mathbf{S}_i$, encoding for low-contrast/no lesion of Task 1, have values very small or even zero, making the linear approximation good enough to calculate the matrix inverses. Thus in Task 1, the linear approximations are good enough to preserve the diagnostic information.

In the other tasks that include high contrast lesions (Tasks 2-5), the d_A of the ideal observer is separated from those of SW observers, indicating that we lose information by doing the linear approximations. In general, the linear approximation of the IWF filter is better than that of the WF filter, because the norm of the matrix in power series (3.25) is compressed via $\mathbf{S}_2 - \mathbf{S}_1$. However, in tasks 2-4, the two d_A's for SW observers on filtered B-mode images are the same, i.e., both linear approximations are the same. There is only an improvement for the IWF filter in Task 5 when the d_A for the SW observer on IWFB-mode images has greater slope than that on WFB-mode images. This can be explained by investigating the higher order terms of the power series in Eq. (3.8). The higher order terms can be written as

$$\begin{aligned}\left(\boldsymbol{\Sigma}_0^{-1/2}\Delta\boldsymbol{\Sigma}_i\boldsymbol{\Sigma}_0^{-1/2}\right)^k &= \left(\boldsymbol{\Sigma}_0^{-1/2}\sigma_{\text{obj}}^2\mathbf{H}\mathbf{S}_i\mathbf{H}^t\boldsymbol{\Sigma}_0^{-1/2}\right)^k \\ &= \sigma_{\text{obj}}^2\boldsymbol{\Sigma}_0^{-1/2}\mathbf{H}\mathbf{S}_i\left(\sigma_{\text{obj}}^2\mathbf{H}^t\boldsymbol{\Sigma}_0^{-1}\mathbf{H}\mathbf{S}_i\right)^{k-1}\mathbf{H}^t\boldsymbol{\Sigma}_0^{-1/2}, \qquad (4.2)\end{aligned}$$

where $k \geq 2$. In Eq. (4.2), $\sigma_{\text{obj}}^2\mathbf{H}^t\boldsymbol{\Sigma}_0^{-1}\mathbf{H}$ is recognized as a filter. $\mathbf{S}_i$ is a diagonal matrix which is considered as a column vector when multiplied with a vector data. Thus $\sigma_{\text{obj}}^2\mathbf{H}^t\boldsymbol{\Sigma}_0^{-1}\mathbf{H}\mathbf{S}_i$ is recognized as a filtered version of variance maps $\mathbf{S}_i$. Using Eqs. (3.2), (3.4), we get $\sigma_{\text{obj}}^2\mathbf{H}^t\boldsymbol{\Sigma}_0^{-1}\mathbf{H} = \mathbf{F}^{-1}\boldsymbol{\Omega}\mathbf{F}$ where $\boldsymbol{\Omega}$ is the diagonal matrix with its elements given by

$$[\boldsymbol{\Omega}]_{ii} = \frac{\sigma_{\text{obj}}^2\left|[\mathbf{T}]_{ii}\right|^2}{\sigma_{\text{obj}}^2\left|[\mathbf{T}]_{ii}\right|^2 + \sigma_n^2}, \qquad (4.3)$$

where $[\mathbf{T}]_{ii}$ is the frequency response of the impulse response function in Eq. (3.2). Thus, $[\boldsymbol{\Omega}]_{ii}$ is the frequency response of the 2D filter whose matrix representation is $\sigma_{\text{obj}}^2\mathbf{H}^t\boldsymbol{\Sigma}_0^{-1}\mathbf{H}\mathbf{S}_i$. With the impulse response function generated by the Field II program shown in Figure 2.4, the frequency response $[\boldsymbol{\Omega}]_{ii}$ is illustrated in Figure 4.4.

The frequency spectra of $\mathbf{S}_i'$s for Tasks 2-5 are shown in Figure 4.5. The figure shows that the frequency spectra of $\mathbf{S}_i'$s for Tasks 2-4 are only different in high frequency, which is multiplied with small values of $[\boldsymbol{\Omega}]_{ii}$. Thus the difference in the high order terms $\left(\boldsymbol{\Sigma}_0^{-1/2}\Delta\boldsymbol{\Sigma}_i\boldsymbol{\Sigma}_0^{-1/2}\right)^k$ of

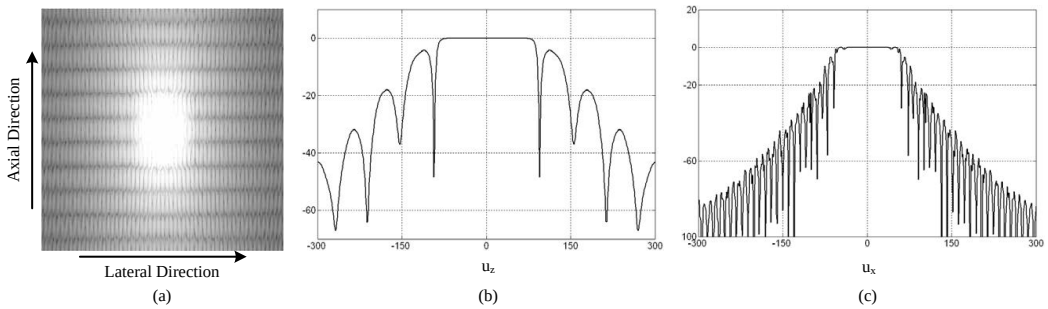

Figure 4.4: The frequency response of the filter $\sigma^2_{\text{obj}}\mathbf{H}^t\boldsymbol{\Sigma}_0^{-1}\mathbf{H}\mathbf{S}_i$ (a) in 2D, (b) a slice in the axial direction and (c) a slice in the lateral direction.

covariance matrix inverses are compressed by $[\boldsymbol{\Omega}]_{ii}$. These terms then are almost canceled by each other when calculating the decision variable, making the linear approximation by separating stationary/nonstationary good enough. It explains why the d'_As of Tasks 2-4 for WFB-mode images are identical with those for IWFB-mode images as illustrated in Figure 4.3, and the Wiener filter still works in the tasks (Tasks 2-4). But in Task 5, the frequency spectra of $\mathbf{S}'_i$s are different in low frequency, which is not compressed by $[\boldsymbol{\Omega}]_{ii}$, making the old linear approximation not good enough. In Figure 4.3 for Task 5, we see that the d_A's for WFB-mode and IWFB-mode images are plotted separately where the d_A for IWFB-mode has a greater slope. This means that IWFB-mode images contain more diagnostic information and explains why the iterative Wiener filter outperformed the Wiener filter.

4.3 Observer Efficiency

Human observer efficiency for the five visual tasks is plotted in Figure 4.6(a). Over the five tasks, the efficiency on the standard B-mode images varies from highest in Task 1 (6.37%) to lowest in Task 4 (0.29%). The WF increases the efficiency by 3 to 5 times in the first four tasks, but it fails in Task 5 where the efficiency goes down from 4.72% to 1.75%. Using IWF helps us double the efficiency in Task 5 (increase to 8.04%), while it keeps the efficiency in other tasks almost unchanged.

The efficiencies of SW observers for images are plotted in Figure 4.6(b). Assuming the SW observer is the ideal observer for B-mode images, we can consider the efficiency as the loss of information when taking the B-mode images. The loss is varying in the five tasks. The diagnostic information left is from 1.18% for Task 3 to 18.34% for Task 1. By filtering before taking the envelope, we increase the information for observation. In Task 5 where the WF fails

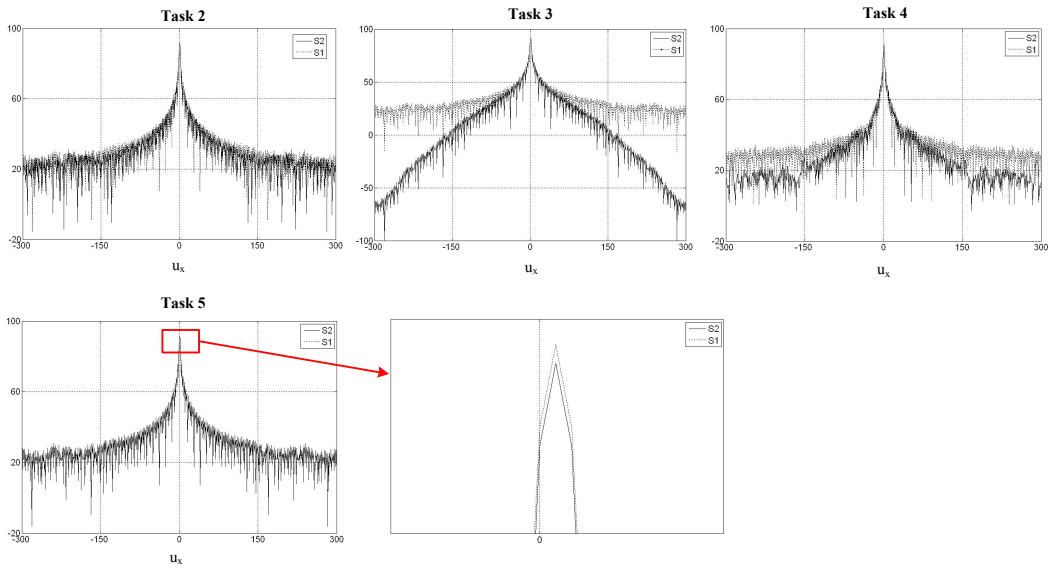

Figure 4.5: The frequency spectrum of $\mathbf{S}_i'$s for Tasks 2-5.

to make human observer improved, the increment of the IWF is double that of WF (28.38% versus 14.29%).

The relative efficiency of the human observers to the Smith-Wagner observers is plotted in Figure 4.6(c). Surprisingly, the human observer can capture the most information in B-mode images. The combination of the WF with an iterative process for tuning the lesion's contrast helps improve human observer efficiency, but it is still less than that of standard B-mode images. The plot suggests that the SW observer can capture some detailed information that is not observed well by the human eye-brain system. The limitation of the human response is beyond the scope of this study and remains to be investigated in the future.

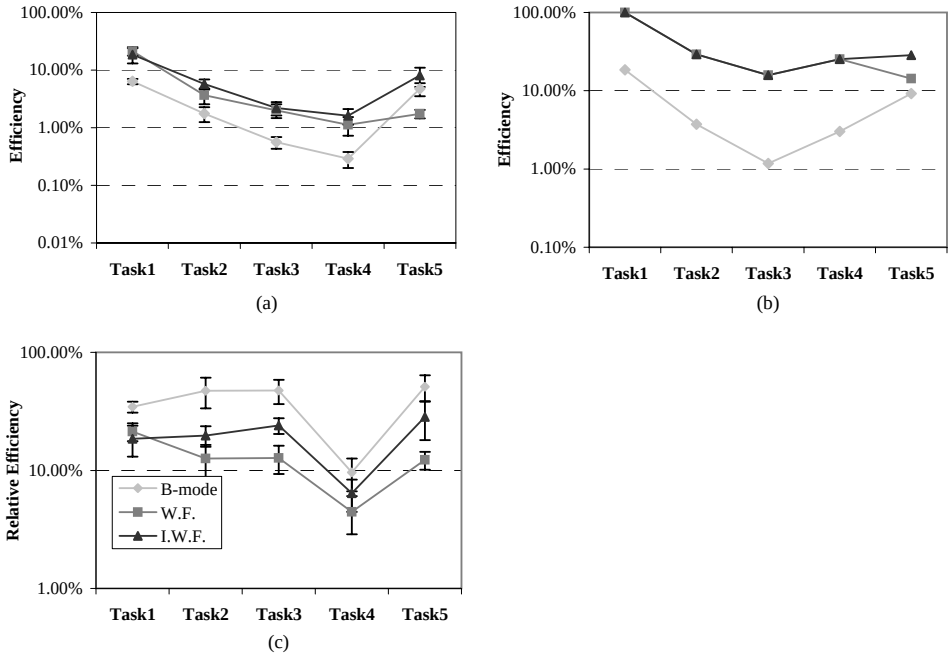

Figure 4.6: (a) Human observer efficiencies for B-mode images. (b) The efficiencies of SW observers. (c) The relative efficiencies of human observers to SW observers.

CHAPTER 5

ROBUSTNESS OF THE WIENER FILTERS

The ideal observer has been derived under a number of assumptions about the nature of scattering in breast tissue, and about the form of the system response. We also made simplifying assumptions to facilitate and speed computation. These assumptions may be violated in the more complex clinical environment. The Wiener filters, derived from linear approximations of the ideal observer, therefore may have some bearing on discrimination performance when applied to clinical data. In this chapter, we investigate these assumptions and their influences on the robustness of the Wiener filters.

5.1 Background Variability

In the synthesis of B-mode images, a more realistic model for a class of random objects is obtained by considering objects with two independent components: the signal to be detected or discriminated, and a background [30]. The background often includes natural structures that play no part in the task; for example, the ribs and other "non-nodule" structures in a chest film. These variabilities can be characterized by using a model of the *lumpy background*, introduced by Rolland and Barrett [31]. In this model, spatial elements, called *lumps* and denoted $l(\mathbf{x})$, are randomly distributed over some area, so the distribution has a form

$$f_\beta(\mathbf{x}) = \sum_{n=1}^{N} l(\mathbf{x} - \mathbf{x}_n). \tag{5.1}$$

A common choice for $l(\mathbf{x})$ is a Gaussian spatial distribution,

$$l(\mathbf{x}) = A\exp\left(-\frac{\mathbf{x}^t\mathbf{x}}{2\sigma_s^2}\right), \tag{5.2}$$

where A is a normalizing constant and positive. The positions $\mathbf{x}_n$ and also the total number N of lumps are random variables. If the positions $\mathbf{x}_n$ are uniformly distributed over the area, then

$f_\beta(\mathbf{x})$ is a stationary process. If N is large and the lump positions are statistically independent, the pdf of a lumpy background approaches a Gaussian by the central-limit theorem (see [1], Chapter 8 for details). Thus we can model the background by a vector column β, which has a Gaussian distribution given by $\beta \sim \text{MVN}(\mathbf{0}, \mathbf{C}_\beta)$.

With the introduction of a zero-mean Gaussian distribution process for the lumpy background, the RF data remains as a zero-mean Gaussian process under both hypotheses. However the covariance matrices become

$$\begin{aligned} \mathbf{\Sigma}_i &= \sigma^2_{\text{obj}}\mathbf{H}(\mathbf{I}+\mathbf{S}_i)\mathbf{H}^t + \sigma^2_{\text{n}}\mathbf{I} + \mathbf{H}\mathbf{C}_\beta\mathbf{H}^t \\ &= \mathbf{\Sigma}_0 + \Delta\mathbf{\Sigma}_i + \mathbf{\Sigma}_\beta \ , \text{ for } i = 1, 2, \end{aligned} \tag{5.3}$$

where $\mathbf{\Sigma}_0$ and $\Delta\mathbf{\Sigma}_i$ are defined in Eq (3.3), and $\mathbf{\Sigma}_\beta = \mathbf{H}\mathbf{C}_\beta\mathbf{H}^t$.

Since the background β is a stationary process, the covariance matrix $\mathbf{C}_\beta$ is a circulant matrix and can be diagonalized by the Fourier techniques. Combining with the decomposition of $\mathbf{H}$ in Eq (3.2), we find that $\mathbf{\Sigma}_\beta$ is also circulant, $\mathbf{\Sigma}_\beta = \mathbf{F}^{-1}\mathbf{N}_\beta\mathbf{F}$, and may be assigned to the stationary parts of the covariance matrices. Hence, the new stationary component in the covariance matrix is given by

$$\begin{aligned} \mathbf{\Sigma}_{0\beta} &= \mathbf{\Sigma}_0 + \mathbf{\Sigma}_\beta \\ &= \mathbf{F}^{-1}\mathbf{N}_{0\beta}\mathbf{F}, \end{aligned} \tag{5.4}$$

where $\mathbf{N}_{0\beta}$ is the diagonal matrix with its elements,

$$\begin{aligned} [\mathbf{N}_{0\beta}]_{ii} &= [\mathbf{N}_0]_{ii} + [\mathbf{N}_\beta]_{ii} \\ &= \sigma^2_{\text{obj}}\left|[\mathbf{T}]_{ii}\right|^2 + \sigma^2_n + [\mathbf{N}_\beta]_{ii}\,. \end{aligned} \tag{5.5}$$

Thus, the eigenvalues of the stationary part $\mathbf{\Sigma}_{0\beta}$ increase. By repeating what we have derived for the upper bound of the norm $\left\|\mathbf{\Sigma}_0^{-1/2}\Delta\mathbf{\Sigma}_i\mathbf{\Sigma}_0^{-1/2}\right\|_2$ in Section 3.1, we find the new upper bound of $\left\|\mathbf{\Sigma}_{0\beta}^{-1/2}\Delta\mathbf{\Sigma}_i\mathbf{\Sigma}_{0\beta}^{-1/2}\right\|_2$ given by

$$\begin{aligned} \left\|\mathbf{\Sigma}_{0\beta}^{-1/2}\Delta\mathbf{\Sigma}_i\mathbf{\Sigma}_{0\beta}^{-1/2}\right\|_2 &\leq \max_i\left(\frac{\sigma^2_{\text{obj}}\left|[\mathbf{T}]_{ii}\right|^2}{\sigma^2_{\text{obj}}\left|[\mathbf{T}]_{ii}\right|^2 + \sigma^2_n + [\mathbf{N}_\beta]_{ii}}\right) \\ &\leq \max_i\left(\frac{\sigma^2_{\text{obj}}\left|[\mathbf{T}]_{ii}\right|^2}{\sigma^2_{\text{obj}}\left|[\mathbf{T}]_{ii}\right|^2 + \sigma^2_n}\right) = \frac{\sigma^2_{\text{obj}}\sigma^2_h}{\sigma^2_{\text{obj}}\sigma^2_h + \sigma^2_n}, \end{aligned} \tag{5.6}$$

where the second inequality comes from the fact that elements of $\mathbf{N}_\beta$ are non-negative. Hence,

the lumpy background reduces the norm of the matrix in the power series, making the linear approximations better.

In Section 4.2.2, we found a connection between the accuracies of the linear approximations with the performance of the filter derived from the approximation. The more accurate of the linear approximation to the covariance inversion, the better filter derived. Thus, Wiener filters are better with the lumpy background. Although the filter is better, the overall discrimination results may be reduced since the background variability may affect the visual performance.

5.2 Circulant Matrix Assumption

The circulant matrix assumption has been made for the system matrix $\mathbf{H}$ in order to have a tractable computational model. It is also known as the stationarity assumption, and allows us to diagonalize $\mathbf{H}$ by an appropriate Fourier transformation [1]. However, it is an unrealistic assumption since no practical imaging system can be accurately represented as a circulant matrix. In an investigation of image quality in digital radiography, Pineda and Barrett have shown that stationarity assumptions can give misleading results [32]. In this section, we investigate this assumption for an ultrasonic system to be sure that the Wiener filters do not fail when we approximate the system matrix as circulant.

In the context of the ultrasonic imaging system described in Appendix A, the RF data can be modeled as a 2D linear convolution of the scattering object and the system impulse response. The B-mode image generated without a stationarity assumption is shown in Figure 5.1(a). There is not much difference from the image generated using the stationarity h in Figure 3.2(a). With this model, we encounter the block Toeplitz matrix for the system matrix $\mathbf{H}$. Thus, assuming that $\mathbf{H}$ is circulant implies an appropriate digital wrap-around for the lower-left and upper-right corners of $\mathbf{H}$, and each block in $\mathbf{H}$. The wrapping causes the errors at the vertical edges of the WFB-mode image shown in Figure 5.1(b).

The error may be tolerable if the impulse response h is compact relative to the *field of view* (FOV). With the center frequency of 7.2 MHz, the impulse response is confined to a rectangle of 2 mm$\times$3.2 mm. Thus, for the FOV of approximately 40 mm, we can reduce the wrap-arround error by locating the lesion in the center of the image. The idea is illustrated in Figure 5.2 with an image including three lesions at different positions. On its WFB-mode images, the two lesions in the corners (in Figure 5.2(b)) have some artifacts caused by the errors, and they may reduce human observer detection performance. However, the lesion in the center is unaffected.

Although the errors can be reduced by increasing the image size, we cannot eliminate the errors in this way since the Wiener filter includes a matrix inversion. There is one way to

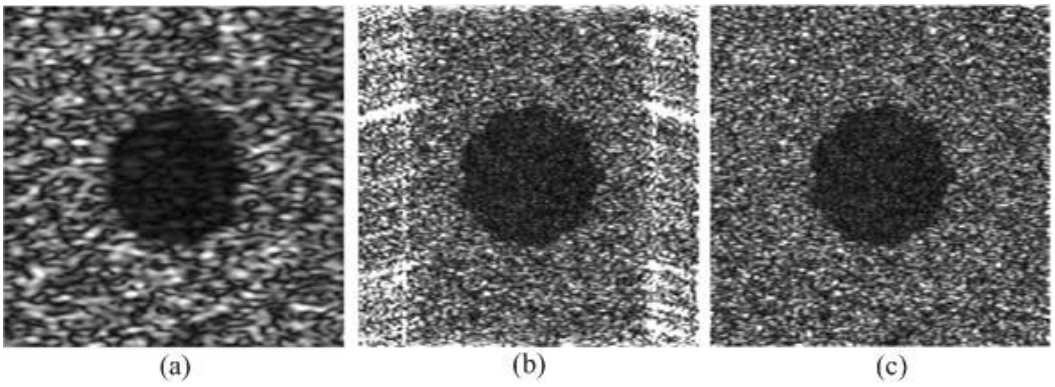

Figure 5.1: An example (Task 4, malignant) of envelope images generated without using stationarity assumption. (a) Standard B-mode. (b) Wiener filtered with stationarity assumption. (c) Wiener filter without stationarity assumption.

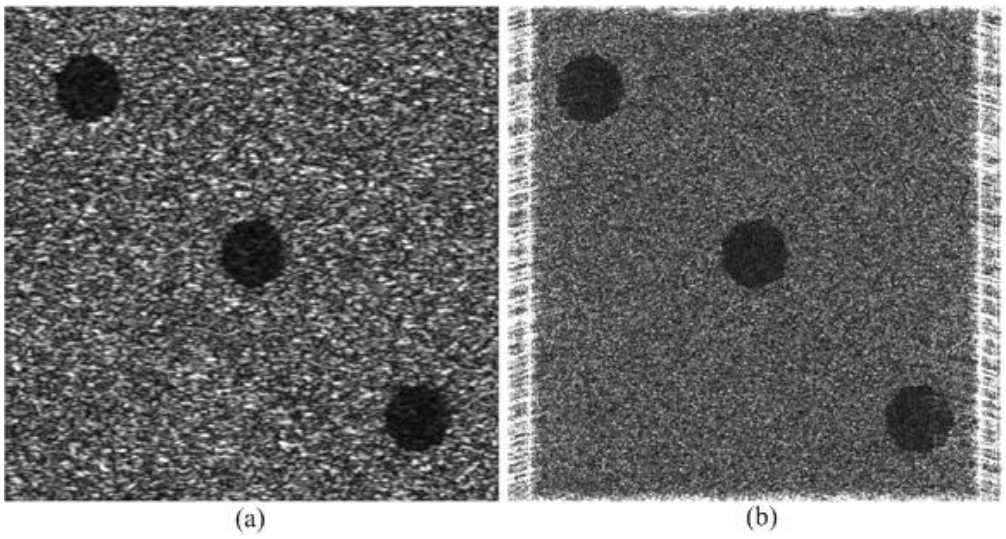

Figure 5.2: An example of the image including three lesions. (a) Standard B-mode image generated without stationarity assumption. (b) The corresponding WFB-mode image.

eliminate the artifact, and it is to use the *conjugate gradient* method [24]. The result of the new method is shown in Figure 5.1(c), and the method is discussed in the next section.

5.3 Linear Shift Variant

Another assumption that has been made for the system impulse response is the shift-invariance. This assumption allows us to characterize the system matrix by just one impulse response that is linearly convoluted with f to generate RF data. In this section, we will investigate how this assumption affects Wiener filters. For doing that, we need a model to generate the RF data without using the assumption, and then use the filter derived in Section 3.2 to process it.

First, the shift-invariant model for the system impulse response is only valid in a small region. In [12], Zemp et al. have shown that the assumption can be made in the axial direction, but it is reasonable only in a small area along the lateral direction (2 mm with dynamic focusing and 1 mm with fixed focusing receiver). Hence, to generate the RF data without the shift-invariant assumption, we divide the image region into small patches. The linear convolution is still applied, but locally, with the impulse response system updated for each isoplanatic patch. The B-mode image corresponding to the RF data is shown in Figure 5.3(a). Unlike the B-mode image in Figure 5.2(a) generated with shift-invariant assumption, its contrast is not uniform over the whole image plane. The contrast is greater in the middle field than in the near field and far field. That is explained by the impulse response being strongest at the focal region when we use a fixed focus for both transmitting and receiving.

By using the impulse response at the focal region to construct the filter, we obtain the WFB-mode image corresponding to the B-mode image; it is shown in Figure 5.3(b). Besides the artifacts on its edges caused by the stationarity assumption, the contrast and the speckles are not uniform over the image because it is the result of a deconvolution with an incorrect impulse response in the near field and far field. In order to have an exact result, we need to modify the derivation of the Wiener filter in Section 3.2.

Generating the RF data under linear shift variant conditions, we derive a new ideal observer and a new Wiener filter in Appendix B. The filter still has the form of $\mathbf{H}^t\boldsymbol{\Sigma}_0^{-1}\mathbf{g}$, but now $\mathbf{H}$ is a partial sum of block-Toeplitz matrices, and $\boldsymbol{\Sigma}_0$ is no longer a covariance matrix of a stationary process, which creates a computational challenge to calculate the new Wiener filter.

From Eq. (B.11), we note that $\boldsymbol{\Sigma}_0$ is a positive definite matrix, so the conjugate gradient method can be applied to calculate the product of $\boldsymbol{\Sigma}_0^{-1}\mathbf{g}$. The multiplication of $\mathbf{H}^t$ with a column vector is taken through a partial sum of 2D-linear convolutions. Thus the calculation of the new Wiener filter can be implemented but at the cost of computational time to achieve

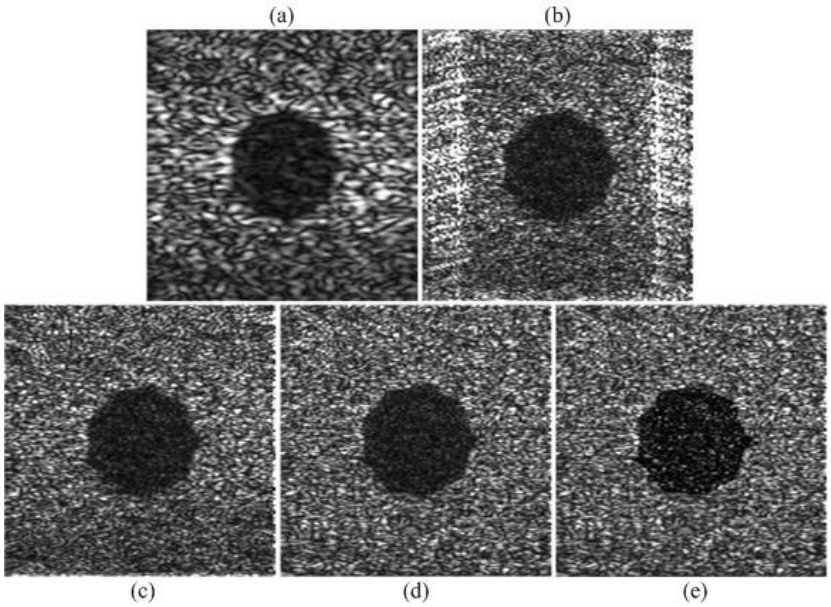

Figure 5.3: An example (Task 4, malignant) of (a) standard B-mode with linear shift variant, (b) WFB-mode image using both assumptions, (c) WFB-mode image using linear shift invariant only, (d) WFB-mode image by filtering the image locally and (e) IWFB-mode image with the WF from (d).

convergence using the conjugate gradient method. To investigate the effect of the linear-shift-invariant assumption only, we use this method for one impulse response at the focal region to calculate the Wiener filter. The corresponding WFB-mode image is shown in Figure 5.3(c). Comparisons with Figure 5.3(b) show the edge errors have been removed, but computational time increases. It takes ~ 25 sec to calculate one realization of the Wiener filter, while calculation using the stationarity assumption takes only a few multiplications and Fourier transformations of column vectors. Besides, the contrast and the speckles size are not uniform over the image plane. Although this error is small, it may affect the human observer performance of Task 1 when detecting the lesion with very low contrast.

Figure 5.3(d) shows the WFB-mode image without using the two assumptions. Compared with the image in Figure 5.3(c), the image is filtered locally with the impulse responses corresponding to each isoplanatic region. The contrast and speckle size are uniform over the whole image, and it looks the same as the WFB-mode image for the ideal case in Figure 3.2(b).

The two assumptions also cause errors in the iterative process of the IWF. However, the circulant assumption mostly affects the edges of images. It does not much affect the region specified by the lesions. Hence, we can use both assumptions for the iterative process to simplify the computation. The IWFB-mode image is shown in Figure 5.3(e). We can see that the two

assumptions do not much affect the quality of the image. Thus we can use them for calculating the iterative process, which helps reduce the time of computation.

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

In this research, we have been developing a rigorous analytical framework for designing and evaluating sonographic systems for breast cancer imaging. By translating clinical situations into Gaussian probability density functions and combining that information with a linear system model for an ultrasound machine, we have derived the test statistic for the ideal observer for the beamformed RF data. The test statistic is quadratic in the data and involves inverses of the associated covariance matrices. Because of the large size of the covariance matrix, the ideal observer's responses have been approximated by applying a series approach expansion and retaining terms up to the first order. From combinations of ideal observer analysis and psychophysical studies on a panel of five typical tasks, we have identified a loss of diagnostic information in demodulation RF data. Based on the exploration of the ideal observer, we have proposed some imaging formation algorithms, interpreted from ideal observer equations, to recover the diagnostic information. The exploration also helps us find the optimal strategy in specific tasks, characterizing for discrimination features of breast tumors.

The first-order approximation to covariance matrices suggests application of the Wiener filter to the RF echo data before computing the B-mode images [6]. By investigating the effects of Wiener filtering on the five tasks, we found a substantial improvement in four of the five tasks. In Task 5, related to discrimination of hypoechoic interior from anechoic interior, we found that human observer performance was reduced. This failure has guided us to extend the ideal observer exploration to form new signal processing strategies.

By analyzing the matrices in a power series approach, we found the assumption for the old linear approximation is violated under the clinical conditions of Task 5. The Wiener filter is tuned to the background of the image and not the lesion interiors that contain task information. By using the new covariance separation into sum and difference terms, we formed a new linear approximation of the ideal observer, which leads us to the iterative Wiener filter for the RF data. It is a Wiener filter, but it is now combined with an iterative process for tuning the filter to match properties of the lesions. The new filter requires image segmentation to roughly specify lesion boundaries. Performing the segmentation on the WFB-mode images speeds the

process with tolerable errors. The results show that the iterative Wiener filter makes a big improvement in performance of Task 5, in comparison with the Wiener filter, while it still keeps the performance on par with other tasks.

With the introduction of the Smith-Wagner observer, approximated as an ideal observer on the envelope images, we decompose the efficiency of human observer relative to the ideal observer into two factors. The first factor is the loss of diagnostic information in computation of an envelope image, while the second is the inaccessibility of human observers to diagnostic information available on the image. The analysis shows that the envelope computation is the significant limiting step for standard B-mode processing, and can be improved by Wiener filtering of the RF data. However, human observers appear to be less efficient at viewing the filtered B-mode images than standard B-mode images. The inconsistency in human observer efficiencies motivates further studies as well as development of observer models for predicting the human observer.

In this thesis, we investigated the consequences of applying spatial filters to clinical data, including the presence of background variability and assumptions made to derive the ideal observer test statistic. By using the ideal observer approach given the statistical distribution of the background, we have shown that Wiener filters increase information content in the presence of background variability. However, the overall visual performance may be reduced since the added variability can affect the diagnostic process. First we assume the system impulse response is stationary, which simplifies computation of the matrix inverses. Inconsistencies caused by violations of this assumption cause artifacts at the edges of the filtered envelope images. If the lesion is small (its diameter is less than 10 mm) and located near the center of the image, the errors can be tolerated since they do not reduce performance. The second assumption is that the system impulse response is shift-invariant. It allows us to characterize the matrix system by one impulse response, which we choose at the focal region. Failure to maintain shift invariance can vary the average speckle brightness and size throughout the image, which lowers detection performance. We have derived the Wiener filter without the shift-invariant assumption. The Wiener filtered envelope images are as good as their versions in the ideal case (when using both assumptions to generate the data), but the computational costs increase significantly.

There are still several issues that we may encounter when applying the Wiener filters to data acquired in the clinical environment. The first issue is *beam steering*. During the acquisition of RF data, beam steering may be applied at both ends of the transducer array to broaden the field of view, or during the scan process to increase the line density. The Wiener filter is derived based on the convolution between the scattering object and the impulse response function. It means that we use one impulse response to characterize all the data in the image plane (shift-invariant) or for one isoplanatic patch (shift-variant). The impulse response is measured in the direction

perpendicular to the transducer's surface. Therefore, beam steering causes some direction errors and they can affect the quality of the filtered images. The second issue is the *correlated noise*. One of the difficulties is the need to accurately estimate the system response. If the impulse response is not estimated exactly, there is some correlated noise in the RF data and the ideal observer does not represent the true situation. Thus Wiener filters derived for uncorrelated noise may work incorrectly. These two issues need to be investigated in the future.

In the next step of the research, we will expand the ideal observer approach for RF data before summation that forms beams. The ideal observer will be derived for RF data from each channel and then explored. Based on the exploration, we will form an optimal beamforming strategy for each specific feature of breast tumor.

APPENDIX A

LINEAR MODEL DESCRIPTION

In this appendix, we describe how to model RF data of the ultrasound machine as an output of a linear system. The acquisition process of RF data has been analyzed and modeled by our research group in [12]. The RF data generation follows the imaging equation, which has a form

$$g(\mathbf{t}) = \int d\mathbf{x} h(\mathbf{x}, \mathbf{t}) \gamma(\mathbf{x}) + n(\mathbf{t}), \tag{A.1}$$

where g is time-varying RF echo data. The function h is the pulse-echo spatio-temporal impulse response that maps objects in space to data in time $\mathbf{t}$. The term γ is called the object scattering function, and n is signal-independent zero-mean, additive, Gaussian system noise. Vector $\mathbf{t}$ represents measurement time for data acquired along the axial and lateral dimensions of the object.

To represent the RF data as in Eq. (A.1), we start with the homogeneous wave equation describing propagation of a small-amplitude acoustic wave in inhomogeneous media [29, 33]. With the scattering geometry illustrated in Figure A.1, the wave equation is as follows:

$$\nabla^2 p - \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} = \gamma_\kappa(\mathbf{x})\frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} + \nabla \cdot (\gamma_\rho(\mathbf{x})\nabla p), \tag{A.2}$$

where

$$\begin{aligned} \gamma_\kappa &= (\kappa_e(\mathbf{x}) - \kappa_0)/(\kappa_0) = \Delta\kappa(\mathbf{x})/(\kappa_0) \\ \gamma_\rho &= (\rho_e(\mathbf{x}) - \rho_0)/(\rho_e(\mathbf{x})) = \Delta\rho(\mathbf{x})/(\rho_e(\mathbf{x})) \\ c^2 &= 1/\kappa_0\rho_0 \ . \end{aligned} \tag{A.3}$$

The quantities κ_e and ρ_e are the compressibility and density, respectively, of the medium at position $\mathbf{x}$, while κ_0 and ρ_0 are the corresponding mean values.

There are several approaches to solving the wave equation (A.2). In [7], Insana and Brown used a weak scattering Green's function approach to provide a time-independent solution, given

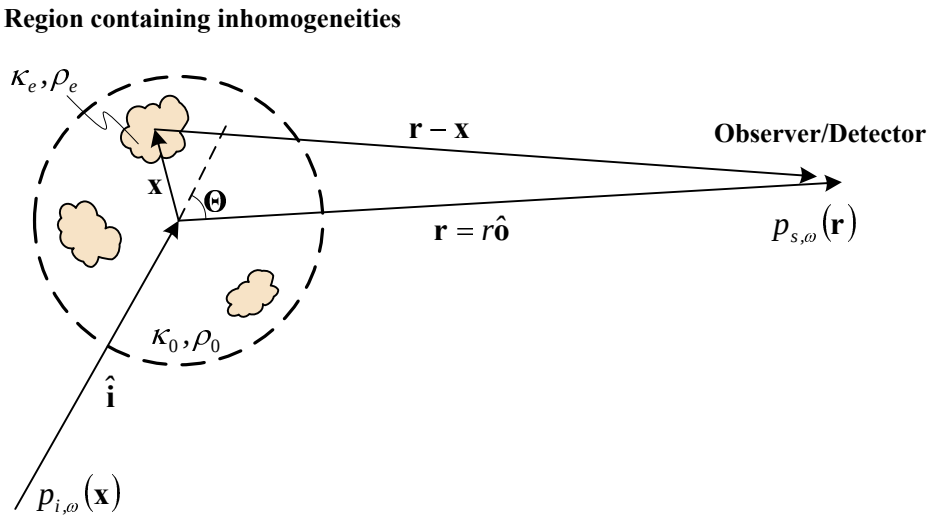

Figure A.1: The wave scattering geometry: $p_{s,w}(\mathbf{r})$ is the scattered pressure field at observed location $\mathbf{r}$ and frequency w, $p_{i,w}(\mathbf{x})$ is the incident pressure field at location $\mathbf{x}$. The backscattering pressure field corresponds to $\Theta = 180^{\circ}$ [7].

as

$$p_{s,w}(\mathbf{r}) = \int_V d\mathbf{x}(k^2\gamma_\kappa(\mathbf{x})p_{i,w}(\mathbf{x})G(\mathbf{r}|\mathbf{x}) + \gamma_\rho(\mathbf{x})\nabla p_{i,w}(\mathbf{x})\nabla G(\mathbf{r}|\mathbf{x})), \tag{A.4}$$

where $p_{s,w}(\mathbf{r})$ is the scattered pressure field at observed location $\mathbf{r}$ and frequency w, $p_{i,w}(\mathbf{x})$ is the incident pressure field at location $\mathbf{x}$ and frequency w, and k is the wave number which can be complex to allow for attenuation. $G(\mathbf{r}|\mathbf{x})$ is the Green's function, describing propagation of the scattered wave at source point $\mathbf{x}$ to the observational point $\mathbf{r}$.

A net instantaneous force, which is a coherent sum of backscattered energy at frequency w propagating to transducer surface [34], is given as

$$f_w(t) = \int_S dS p_{s,w}(\mathbf{r}), \tag{A.5}$$

where S is the area of the transducer surface.

From [33], the incident field pressure for a plane wave can be written in terms of

$$p_i(\mathbf{x}) = i\rho_0 ckU(w)\exp(iwt)A(\mathbf{x}, k), \tag{A.6}$$

where

$$\begin{aligned} A(\mathbf{x},k) &= \frac{1}{2\pi}\int_S dS\xi(\mathbf{r})\frac{\exp(-ik|\mathbf{r}-\mathbf{x}|)}{|\mathbf{r}-\mathbf{x}|)} \\ &= 2\int_S dS\xi(\mathbf{r})G(\mathbf{r}|\mathbf{x}) \end{aligned} \tag{A.7}$$

is the velocity potential, and $\xi(\mathbf{r})$ is the transducer apodization function, which may be complex to allow for focusing. The first expression of Eq. (A.7) is also known as the Rayleigh integral [13], a linear superposition of excitations from elements of the transducer surface. Combining with Eq. (A.5), the net instantaneous force is given as

$$\begin{aligned} f_w(t) &= \frac{1}{2}i\rho_0 ckU(w)\exp(iwt) \\ &\quad \times \int_V d\mathbf{x}\left\{k^2\left(\frac{\Delta\kappa(\mathbf{x})}{\kappa_0}\right)A^2(\mathbf{x},k) + \left(\frac{\Delta\rho(\mathbf{x})}{\rho(\mathbf{x})}\nabla A(\mathbf{x},k)\cdot\nabla A(\mathbf{x},k)\right)\right\}. \end{aligned} \tag{A.8}$$

By ignoring the output noise, the RF echo-signal voltage can be considered as a weighted superposition of net forces at each frequency in the transmitted pulse, which is given as

$$g(t) = \int dwY(w)f_w(t). \tag{A.9}$$

$Y(w)$ is the complex electromechanical coupling coefficient of the transducer.

Substituting Eq. (A.8) into Eq. (A.9), we can write the echo-signal voltage as a product in (A.1) of two functions (neglecting output noise): the first includes all items related to the system (spatiotemporal impulse response) and the second is the object function.

$$h(\mathbf{x},t) = \int dwY(w)\left[\frac{1}{2}i\rho_0 ckU(w)\right]k^2A^2(\mathbf{x},k)\exp(iwt), \tag{A.10}$$

and

$$\gamma(\mathbf{x}) = \left(\frac{\Delta\kappa(\mathbf{x})}{\kappa_0} - \frac{\Delta\rho(\mathbf{x})}{\rho(\mathbf{x})}\right) \cong -2\frac{\Delta z(\mathbf{x})}{z_0}. \tag{A.11}$$

The second equation of (A.11) comes from $z(\mathbf{x}) = \sqrt{\rho(\mathbf{x})/\kappa(\mathbf{x})}$, the acoustic impedance, where z_0 is the mean impedance. Thus the object function γ is the relative change in acoustic impedance. The expression only holds for a small change in density and compressibility. Typically, $\kappa(\mathbf{x}) < 0.1$ and often it is less than 0.01 for biological tissues [7].

From Eq. (A.10), we write the spatiotemporal impulse response in terms of convolution as

$$h(\mathbf{x},t) = -\frac{1}{c^2}\frac{\partial^2}{\partial t^2}\left\{h_y(t) * h_y(t) * v(t) * \frac{\partial}{\partial t}h_a(\mathbf{x},t) * h_a(\mathbf{x},t)\right\}, \tag{A.12}$$

where $h_y(t) = \mathcal{F}^{-1}Y(w)$, which is the electromechanical impulse response, and $v(t)$ is the driving force. The expression $h_a(\mathbf{x}, t) = \mathcal{F}^{-1}A(\mathbf{x}, k)$, which is the transmit impulse response that characterizes the pressure field for a particular transducer geometry, is given as

$$h_a(\mathbf{x}, t) = \frac{1}{2\pi} \int_S dS\xi(\mathbf{r}) \frac{\delta(t - |\mathbf{r} - \mathbf{x}|/c)}{|\mathbf{r} - \mathbf{x}|)}. \tag{A.13}$$

We see that $h_a(\mathbf{x}, t)$ is the velocity potential caused by a temporal delta function excitation. In [10, 29], $h_a(\mathbf{x}, t)$ is also named the *spatial impulse response*. Thus we have described the ultrasonic imaging system as a linear system with the impulse response given in Eq. (A.12).

APPENDIX B

WIENER FILTER FOR SHIFT VARIANCE

In a previous study [6] and so far in this research, RF data are represented as a linear convolution of the scattering object and the impulse response at the focal region. This linear shift-invariant model is valid only for the isoplanatic region where the system impulse response is unchanged. In the ultrasonic system, the impulse response is changed very little along the lateral direction, but rapidly in the axial direction because of diffraction and ultrasound attenuation. Thus, we need a more realistic model for the RF data and a corresponding Wiener filter.

In [12], Zemp et al. found that the isoplanatic region can be made for a small patch (in the axial direction), where the size of the patch is 2 mm for dynamic focused and 1 mm for fixed focused at the receiver. Therefore, we proposed a new model for the RF data, in which it is still a linear transformation of the scattering object but the impulse response is applied locally and updated for each isoplanatic patch. The equation for RF data $\mathbf{g}$ is given as

$$\mathbf{g} = \sum_{j=1}^{k} \mathbf{H}_j \mathbf{f}_j + \mathbf{n} = \mathbf{H}\mathbf{f} + \mathbf{n}, \tag{B.1}$$

where k is the number of the divided patches and $\mathbf{H}_j$ is the block Toeplitz matrix constructed from the impulse response h_j for the j^{th} patch. $\mathbf{f}_j = \mathbf{E}_j\mathbf{f}$, where $\mathbf{f}$ is a column vector of scattering objects, and $\mathbf{E}_j$ is a diagonal matrix with the elements of the diagonal of 1's in the region corresponding to patch j, and 0's for elsewhere. Then we obtain the equation for $\mathbf{E}_j$, given as

$$\sum_{j=1}^{k} \mathbf{E}_{\text{j}} = \mathbf{I}, \tag{B.2}$$

where $\mathbf{I}$ is an identity matrix of dimension n, and $\mathbf{H} = \sum_{j=1}^{k} \mathbf{H}_j\mathbf{E}_j$. Note that although each of $\mathbf{H}_j$ is a block Toeplitz matrix, $\mathbf{H}$ is not.

The covariance matrices for the RF data $\mathbf{g}$ now become

$$\mathbf{\Sigma}_i = \sigma_{\text{obj}}^2 \sum_{j=1}^{k} \mathbf{H}_j \mathbf{E}_j (\mathbf{I} + \mathbf{S}_i) \mathbf{E}_j^t \mathbf{H}_j^t \tag{B.3}$$

and the decision variable is still given by (2.17).

Separating the covariance matrices into stationary and nonstationary terms as in (3.3), we obtain the equations for $\mathbf{\Sigma}_0$ and $\Delta\mathbf{\Sigma}_i$ given as

$$\begin{aligned} \mathbf{\Sigma}_0 &= \sigma_{\text{obj}}^2 \sum_{j=1}^{k} \mathbf{H}_j \mathbf{E}_j \mathbf{E}_j^t \mathbf{H}_j^t + \sigma_n^2 \mathbf{I} \\ \Delta\mathbf{\Sigma}_i &= \sigma_{\text{obj}}^2 \sum_{j=1}^{k} \mathbf{H}_j \mathbf{E}_j \mathbf{S}_i \mathbf{E}_j^t \mathbf{H}_j^t. \end{aligned} \tag{B.4}$$

Using the power series expansion of covariance matrix and truncating at the first order, we obtain the linear approximation of the decision variable

$$\lambda(\mathbf{g}) \cong \frac{1}{2} \mathbf{g}^t \mathbf{\Sigma}_0^{-1} (\Delta\mathbf{\Sigma}_2 - \Delta\mathbf{\Sigma}_1) \mathbf{\Sigma}_0^{-1} \mathbf{g}, \tag{B.5}$$

where

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

Noting that for $j \neq l$, $\mathbf{E}_j (\mathbf{S}_2 - \mathbf{S}_1) \mathbf{E}_l^t = 0$, $\Delta\mathbf{\Sigma}_2 - \Delta\mathbf{\Sigma}_1$ can be written as

$$\begin{aligned} \Delta\mathbf{\Sigma}_2 - \Delta\mathbf{\Sigma}_1 &= \sigma_{\text{obj}}^2 (\sum_{j=1}^{k} \mathbf{H}_j \mathbf{E}_j)(\mathbf{S}_2 - \mathbf{S}_1)(\sum_{j=1}^{k} \mathbf{E}_j^t \mathbf{H}_j^t) \\ &= \sigma_{\text{obj}}^2 \mathbf{H} (\mathbf{S}_2 - \mathbf{S}_1) \mathbf{H}^t. \end{aligned} \tag{B.7}$$

By replacing (B.7) into (B.5), we obtain the linear approximation of the test statistic given as

$$\lambda(\mathbf{g}) \approx \frac{\sigma_{\text{obj}}^2}{2} \mathbf{g}^t \mathbf{\Sigma}_0^{-1} \mathbf{H} (\mathbf{S}_2 - \mathbf{S}_1) \mathbf{H}^t \mathbf{\Sigma}_0^{-1} \mathbf{g}. \tag{B.8}$$

Thus, the Wiener filter has a form of $\mathbf{H}^t \mathbf{\Sigma}_0^{-1} \mathbf{g}$, where

$$\mathbf{H} = \sum_{j=1}^{k} \mathbf{H}_j \mathbf{E}_j \tag{B.9}$$

and

$$\mathbf{\Sigma}_0 = \sigma_{\text{obj}}^2 \sum_{j=1}^{k} \mathbf{H}_j \mathbf{E}_j \mathbf{E}_j^t \mathbf{H}_j^t + \sigma_n^2 \mathbf{I}. \tag{B.10}$$

We note that for $j \neq l$, $\mathbf{E}_j \mathbf{E}_l^t = 0$, combining with (B.9), $\mathbf{\Sigma}_0$ can be written as

$$\mathbf{\Sigma}_0 = \sigma_{\text{obj}}^2 \mathbf{H}\mathbf{H}^t + \sigma_n^2 \mathbf{I}. \tag{B.11}$$

Thus the Wiener filter still has the same form as the Wiener filter we derived in [6], except that the system matrix $\mathbf{H}$ now is a partial sum of block-Toplitz matrices and $\mathbf{\Sigma}_0$ is not a covariance of a stationary process.

REFERENCES

[1] H.H. Barrett and K.J. Myers, *Foundations of Image Science*. Hoboken, NJ: John Wiley & Sons, 2004.

[2] H.H. Barrett, "Objective assessment of image quality: Effects of quantum noise and object variability," *Journal Optical Society of America A*, vol. 7, no. 7, pp. 1266-1278, 1990.

[3] D.M. Green and J.A. Swets, *Signal Detection Theory and Psychophysics*. New York, NY: Wiley, 1966.

[4] R.F. Wagner, S.W. Smith, J.M. Sandrik, and H. Lopez, "Statistics of speckle in ultrasound B-scans," *IEEE Transactions on Sonics and Ultrasonics*, vol. 30, no. 3, pp. 156-163, 1983.

[5] S.W. Smith, R.F. Wagner, J.M. Sandrik, and H. Lopez, "Low contrast detectability and contrast/detail analysis in medical ultrasound," *IEEE Transactions on Sonics and Ultrasonics*, vol. 30, no. 3, pp. 164-173, 1983.

[6] 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 Transactions on Medical Imaging*, vol. 25, no. 2, pp. 198-209, 2006.

[7] M.F. Insana and D.G. Brown,"Acoustic scattering theory applied to soft biological tissues," in *Ultrasonic Scattering in Biological Tissues*, K.K. Shung and G.A. Thiemes, Eds. Boca Raton, FL: CRC Press, 1993.

[8] D. Boukerroui, O. Basset, N. Guerin, and A. Baskurt, "Multiresolution texture based adaptive clustering algorithm for breast lesion segmentation," *European Journal of Ultrasound*, vol. 8, no. 2, pp. 135-144, 1998.

[9] G. Xiao, M. Brady, J.A. Noble, and Y. Zhang, "Segmentation of ultrasound B-mode images with intensity inhomogeneity correction," *IEEE Transactions on Medical Imaging*, vol. 21, no. 1, pp. 48-57, 2002.

[10] J.A. Jensen and N.B. Svendsen, "Calculation of pressure fields from arbitrarily shaped, apodized, and excited ultrasound transducers," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 39, no. 2, pp. 262-267, 1992.

[11] J.A. Jensen, "Field: A program for simulating ultrasound systems," *Medical and Biological Engineering and Computing*, vol. 34, Supplement 1, Part 1, pp. 351-353, 1996.

[12] R.J. Zemp, C.K. Abbey, and M.F. Insana, "Linear system model for ultrasonic imaging: Application to signal statistics," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 50, no. 6, pp. 642-654, 2003.

[13] J.W. Goodman, *Introduction to Fourier Optics*, 2nd ed. New York, NY: McGraw-Hill, 1996.

[14] R.F. Wagner and D.G Brown, "Overview of a unified SNR analysis of medical imaging systems," *IEEE Transactions on Medical Imaging*, vol. 1, no. 4, pp. 210-213, 1982.

[15] R.J. Zemp, C.K. Abbey, and M.F. Insana, "Detection performance theory for ultrasound imaging system," *IEEE Transactions on Medical Imaging*, vol. 24, no. 3, pp. 300-310, 2004.

[16] T. Tot,"General morphology of benign and malignant breast lesions: Old parameters in new perspectives," in *Emerging Technologies in Breast Imaging and Mammography*, J.S. Suri, R.M. Rangayyan, and S. Laxminarayan, Eds. Stevenson Ranch, CA: American Scientific Publishers, 2008, pp. 1-23.

[17] M. Insana and M. Oelze,"Advanced ultrasonic imaging techniques for breast cancer research," in *Emerging Technologies in Breast Imaging and Mammography*, J.S. Suri, R.M. Rangayyan, and S. Laxminarayan, Eds. Stevenson Ranch, CA: American Scientific Publishers, 2008, pp. 141-160.

[18] American College of Radiology, *Breast Imaging Reporting and Data System Atlas*. Reston, VA: American College of Radiology, 2003.

[19] H.L. Van Trees, *Detection, Estimation, and Modulation Theory*, Part I. New York, NY: Wiley, 1968.

[20] H.H. Barrett, J.L. Denny, R.F. Wagner, and K.J. Myers, "Objective assessment of image quality. II. Fisher information, Fourier crosstalk, and figures of merit for task performance," *Journal of the Optical Society of America A*, vol. 12, no. 5, pp. 834-852, 1995.

[21] H.H. Barrett, C.K. Abbey, and E. Clarkson, "Objective assessment of image quality. III. ROC metrics, ideal observers, and likelihood-generating functions," *Journal of the Optical Society of America A*, vol. 15, no. 6, pp. 1520-1535, 1998.

[22] W.P. Tanner and T.G. Birdsall, "Definition of *d'* and η as psychophysical measures," in *Signal Detection and Recognition by Human Observer: Contemporary Readings*, J.A Swets, Ed. New York, NY: Wiley, 1964, pp. 147-163.

[23] R.F. Wagner, M.F. Insana, and D.G. Brown, "Statistical properties of medical ultrasound with applications to medical imaging," *Journal of the Optical Society of America A*, vol. 4, no. 5, pp. 910-922, 1987.

[24] G.H. Golub and C.F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: The Johns Hopkins University Press, 1996.

[25] T.N. Pappas, "An adaptive clustering algorithm for image segmentation," *IEEE Transactions on Signal Processing*, vol. 40, no. 4, pp. 901-914, 1992.

[26] B.J. Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society: Series B*, vol. 48, no. 3, pp. 259-302, 1986.

[27] J.S. Geman and D. Geman, "Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721-741, 1984.

[28] E.A. Ashton and K.J. Parker, "Multiple resolution Bayesian segmentation of ultrasound images," *Ultrasonic Imaging*, vol. 17, no. 4, pp. 291-304, 1995.

[29] R.S.C. Cobbold, *Foundations of Biomedical Ultrasound*. New York, NY: Oxford University Press, 2006.

[30] K.J. Myers,"Ideal observer models of visual signal detection," in *Handbook of Medical Imaging*, Vol. 1, J. Beutel, H.L. Kundel, and R.L. Van Metter, Eds. Bellingham, WA: The Society of Photo-Optical Instrumentation Engineers (SPIE) Press, 2000, pp. 559-592.

[31] J.P. Rolland and H.H. Barrett, "Effect of random background inhomogeneity on observer detection performance," *Journal of the Optical Society of America A*, vol. 9, no. 5, pp. 649-658, 1992.

[32] A. R. Pineda and H.H. Barrett, "What does DQE say about lesion detectability in digital radiography?" in *Proceedings of Society of Photo-Optical Instrumentation Engineers (SPIE)*, vol. 4320, pp. 561-569, 2001.

[33] L.E. Kinsler, A.R. Frey, A.B. Coppens, and J.V Sanders, *Fundamentals of Acoustics*, 4th ed. New York, NY: Wiley, 2000.

[34] M.F. Insana, R.F. Wagner, D.G. Brown, and T.J. Hall,"Describing small-scale structure in random media using pulse-echo ultrasound," *Journal of the Optical Society of America A*, vol. 87, no. 1, pp. 179-192, 1990.