Generalized NEQ for assessment of ultrasound image quality

Roger J. Zemp*, Craig K. Abbey, and Michael F. Insana
Department of Biomedical Engineering, University of California, Davis, 95616

ABSTRACT

An information-theoretic framework for assessing and predicting ultrasound system performance for detection tasks is outlined. Current models of image quality for ultrasound detection tasks make some stringent assumptions, including large target area, that place limits on the applicability of the theory. New models of image quality for ultrasound systems are proposed based on the ideal observer that account for noise, system, and object properties. One result is an expression for the ideal observer detectability that is a generalization of Noise-Equivalent Quanta (NEQ), a measure used by photon imaging modalities. The detection signal-to-noise ratio is shown to be an integration of the generalized NEQ weighted by the spectral variance of the target signal (in contrast to the squared magnitude of the Fourier transform of the signal, as is the case with other modalities). This reflects that ultrasound systems are sensitive not to the magnitude of the medium parameters but rather the variance (spatial fluctuations) of these quantities. The resulting framework is amenable to measurement and prediction of system performance. The theory is used to predict the information content of the ultrasonic beam at various field points. New strategies are revealed for processing RF data that could improve detection of lesions.

Keywords: Ideal observer, noise-equivalent quanta, lesion, detection

1. INTRODUCTION AND MOTIVATION

At the frontier of cancer imaging is the identification and characterization of non-palpable neoplasms. Ultrasound systems are playing a more prominent role in the detection and management of cancer, as well as in understanding the basic mechanisms of disease progression. Central to effective imaging system design is the ability to rigorously and quantitatively characterize system performance. The concept of image quality can be made precise when associated with the ability of imaging systems to perform some task. We focus our attention on cancer applications, where the task is to say whether a neoplasm is present or absent, given an image or set of images.

The approach taken this paper is to provide a way to optimize ultrasound systems specifically for lesion detection tasks. In other words, we can tailor the system to the task to optimize relevant information content in the raw data. We do this by outlining a formalism to quantitatively characterize ultrasound system performance based on models of the system, the noise, as well as accounting for statistical variability in the body. We use an information-theoretic approach based on the ideal observer for ultrasound detection tasks.

To review the concept of an ideal observer, we first define what is meant by an observer. For detection tasks, we define an observer as an individual or entity that makes a binary decision given a set of image data represented by the vector $g$. An observer may often be modeled by a function $\lambda(g)$ of the image data. Given an image, the observer synthesizes a scalar test statistic $\lambda(g)$, then compares this number with a threshold to make a decision. Over an ensemble of images where a signal is present, $\lambda$ will be statistically distributed about some mean. Likewise, there will be a corresponding distribution when the signal is absent. Different observers will give rise to different distributions of $\lambda$. Upon thresholding, most observers will make decision errors. The most widespread way of characterizing observer performance for binary detection tasks is to plot the probability of a true positive for a given false positive probability in a Receiver Operating Characteristic (ROC) curve.

An ideal observer has the optimum ROC curve of all observers, that is, it has the maximum possible true positive probability for every false positive level. Classical statistical decision theory has shown that the optimal observer test statistic is the likelihood ratio, or a monotonic transformation of it, such as the log-likelihood ratio. Observers for signal

* E-mail: rjemp@ucdavis.edu; phone: (530) 754-7376
detection tasks, and indeed any task in medical imaging task can be influenced by a number of factors. First, there is object variability – or variations in the body that can obscure diagnostic interpretation. Second, system blur and noise, as well as confounding artifacts limit the type of information that can be collected. Third, the raw signal collected by the imaging system will be processed. This processing may incur some loss of information. Fourth, the processed data will be displayed in some media format, and fifth, the eye-brain system of the individual reading the images will influence interpretation. The ideal observer is not concerned with factors 4 and 5, but rather is concerned with how object variability, system design, and processing algorithms affect the detection task.

Understanding the ideal observer is important for system and algorithm design. The goal is to make ultrasound systems better than they now are at detecting cancerous lesions. Optimizing a system for this task can be viewed as a two-step procedure. First, one should design the system so as to acquire as much task-relevant information available as possible in the raw signal. Second, one should design processing algorithms and display the processed data in a way that makes information about the task optimally available to human observers. In other words, we want the human observer to perform as nearly as possible to the ideal observer upper bound. Although the ultimate goal is to improve performance of human observers by improving design and algorithms, this paper does not attempt to model human observer performance. We are mainly concerned with quantifying the upper bound on information content. Future work will attempt to integrate both human and ideal performance to optimize for the task. In some cases it may be necessary to sacrifice ideal observer performance to improve human observer performance. We should mention that true clinical tasks have many complexities, and that we only concern ourselves with modeling essential features of true clinical tasks.

Besides quantifying information content, the ideal observer is also important for another reason. The ideal observer indicates strategies for how to use raw data to make a decision that could help human observers. An image processing algorithm that could mimic the ideal observer’s strategy could be very valuable. Currently, B-mode ultrasound systems use envelope detection, a process that discards phase information. Does this procedure incur a loss of information relevant to the detection task? Are there any other practical ways of processing the data? These questions need a fundamental framework to be answered.

We begin our discussion with a review of current models of image quality in the literature. We show the need for a new model, and develop a framework applicable to more realistic situations. We then draw connections to current literature regarding Generalized Noise-Equivalent Quanta (GNEQ), a term first coined by Barrett et al. as a generalization of Noise-Equivalent Quanta (NEQ) first introduced by Shaw for photon imaging systems. The interpretation for Generalized NEQ for ultrasound systems is given, and an explanation is given regarding how performance metrics for ultrasound systems differ from photon imaging modalities. We show that GNEQ characterizes the ultrasound system for a given tissue type. Our results provide a practical way to not only predict, but to measure the effectiveness of design or beamforming strategies in various scattering media.

2. LITERATURE REVIEW

Some important contributions to the area of ultrasonic image quality have been made in past literature. Seminal papers by Wagner and Smith and colleagues have outlined fundamental models for signal statistics, and a simple model for the ideal observer given specialized circumstances. Here we give a didactic overview of the salient features of these findings, clarify assumptions made, and motivate the need for new models.

The Smith and Wagner approach began with an assumption that echo signals could be modeled as a random walk in the complex plane. For large numbers of random scatterers per pulse volume, envelope-detected image data \( \mathbf{v} \) (representing \( N \) pixels) was shown to exhibit Rayleigh statistics, characterized by a single parameter, \( \psi \).

\[
pdf(\mathbf{v} | \psi) = \prod_{m=1}^{N} \frac{v_m}{\psi} e^{-v_m^2/2\psi}
\]

This density is obtained from zero-mean circular Gaussian statistics by a simple transformation of variables. Notice that an assumption is made that data points are statistically independent. Thus, instead of counting pixels, Smith and Wagner chose to deal with ultrasonic speckle spots, which have dimensions defined by correlation lengths. Statistical decision theory was then put to use. The problem they were interested in was the task of target detection, where the target was for
example, a disk signal represented by an increased region of variance. Signal present and signal absent variances may be denoted as $\psi_+$ and $\psi_-$. The test statistic of the ideal observer is the log-likelihood ratio

$$\lambda(v) = \log \left[ \frac{pdf(v | \psi_+)}{pdf(v | \psi_-)} \right]$$

(2)

Here $pdf(v | \psi_+)$ and $pdf(v | \psi_-)$ are the likelihoods (probabilities) of the signal being present and absent respectively. Ignoring irrelevant constants and scaling factors, this test statistic is shown to be

$$\lambda(v) = \sum_{m=1}^{N} v_m^2$$

(3)

Using first and second moments of the Rayleigh distribution, the ideal observer lesion detection signal-to-noise ratio

$$SNR_j^2 = \frac{[\bar{\lambda}_+ - \bar{\lambda}_-]^2}{\sigma_+^2 + \sigma_-^2 / 2}$$

(4)

is given as

$$SNR_j^2 = N \frac{[\psi_+ - \psi_-]^2}{\psi_+^2 + \psi_-^2} = \frac{A}{S_c}SNR_o^2C^2.$$  

(5)

$\bar{\lambda}_+$ and $\sigma_+^2$ are the mean and variance of the test statistic $\lambda$ assuming the lesion is present, and $\bar{\lambda}_-$ and $\sigma_-^2$ are the mean and variance assuming the signal is absent. $A$ is the lesion area, $S_c$ is the area of a speckle spot, $SNR_o$ is the pixel signal to noise ratio (1.91 for Rayleigh distribution), $C$ is the contrast of the lesion, and $N$ is the number of independent samples (speckle spots). The speckle spot area is the integrated normalized autocovariance of the image field, and depends on the system design. When there is a correspondence between spatial resolution and speckle spot size, the $SNR$ equation tells us that detectability is better for high-resolution systems because there is more independent information available for making decisions. Task performance is also improved with increasing lesion contrast.

The Smith-Wagner theory was a breakthrough for ultrasonography because it connected for the first time it established a rigorous connection between engineering properties that we measure in the laboratory and clinical task performance. Conveniently, system properties could be uncoupled from object properties. The Smith-Wager theory used some stringent assumptions including the following: (1) Electronic noise was neglected in the analysis. (2) The theory used envelope detected signals instead of the raw RF data. Further insights into optimal detection might be available by using unprocessed data. (3) Speckle spots instead of pixels were used. A modified theory that allowed for correlated pixels would be more satisfying. (4) There was an implicit assumption made that only large lesions were being considered, as the variance for both hypotheses were constant across image data. Hence clutter due to beam sidelobes, and related edge effects are not described in this theory. A detection theory that addresses smaller lesions may prove very important. (5) An assumption of linear shift-invariance was made. We propose a theory that will extend the Smith-Wagner theory, and relax some of the assumptions upon which it is based.

To address these questions we turn to the toolbox of literature available for image quality assessment. In particular, three papers on objective assessment of image quality by Barrett and colleagues provide us with a rigorous foundation on which to make further developments. These papers deal with estimation and detection task performance assessment. They make connections to the Rose theory of contrast and the Wagner-Brown model for assessment of detection tasks.

### 3. LINEAR SYSTEM MODELS FOR ULTRASOUND SIGNALS

Ultrasound systems transmit acoustic waves into tissue and form images from the scattered signals that return to the transducer. Ultrasound radio-frequency (RF) echo signals can be modeled as a shift-variant linear system of the form

$$g(t) = \int dx h(x, t) y(x) + n(t).$$

(6)
Here $g$ is the voltage trace echo signal, $n$ is the noise, and $h$ is the imaging system response to a delta function object. $h$ represents all (linear) aspects of the system including the transmit waveform, electromechanical coupling of the transducer, diffractive propagation, scattering, reception, and beamforming. It does not include nonlinear processing steps such as envelope detection. $\gamma$ is the object function, which represents spatial perturbations in density and compressibility. Unlike other modalities, ultrasound systems are not sensitive to the shape or distribution of homogeneous materials but rather spatial fluctuations in medium properties. Consequently, a scattering medium is most usefully characterized statistically. The variance of $\gamma$ is associated with echo backscatter strength. The spatial vector $x$ defines the location in the object space, and the temporal vector $t$ spans the data space. For a single A-scan line, $t$ may be a 1-D vector. The echo data for a 2-D B-mode image uses a 2-D matrix for $t$. For higher dimensional acquisitions, e.g., flow and elastography, $t$ further increases in dimension. For discrete images of $M$ pixels the linear system can also be written in vector-operator notation as:

$$g = H \gamma + n$$

(7)

where $H$ is a continuous-to discrete operator and the data $g$ is now a vector or matrix with $M$ elements.

4. IDEAL OBSERVERS OF THE RF SIGNAL

It is the statistical properties of the echo signal that give us information about the object imaged. The probability density functions for signal present and signal absent RF echo signals can be modeled as zero-mean multivariate normal densities

$$pdf(g | +) = \frac{\exp\left[-\frac{1}{2}g^T K_+^{-1} g\right]}{(2\pi)^{M/2} \sqrt{\det K_+}}$$

and

$$pdf(g | -) = \frac{\exp\left[-\frac{1}{2}g^T K_-^{-1} g\right]}{(2\pi)^{M/2} \sqrt{\det K_-}}.$$  

(8)

Here the covariance matrices $K_+ = \langle gg' | + \rangle$ and $K_- = \langle gg' | - \rangle$ describe the image textures for the signal present and signal absent hypotheses respectively. A lesion present, thus, for example, may have a region of increased variance relative to the surrounding tissue. The log-likelihood ratio is given as

$$\lambda(g) = \log \left[ \frac{pdf(g | +)}{pdf(g | -)} \right] = \frac{1}{2} g^T \left[ K_+^{-1} - K_-^{-1} \right] g + \frac{1}{2} \log \left[ \det \left( K_+^{-1} \right) \right] + \frac{1}{2} \log \left[ \det \left( K_-^{-1} \right) \right] + \frac{1}{2} \left( g^T K_+^{-1} g + g^T K_-^{-1} g \right) - \frac{1}{2} \left( g^T K_+^{-1} g + g^T K_-^{-1} g \right).$$

(9)

We can ignore the second term as it is a data-independent constant. The ideal observer test statistic gives the strategy for optimal detection. The strategy is to compute the difference between the squared magnitudes of the pre-whitened data. In section 8 we comment on the importance of the ideal observer strategy for improving human observer performance by image processing. Eq. 9 represents a signal known statistically (SKS) quadratic task. By quadratic task we mean the test statistic is a quadratic form of the data vector $g$. At this point we can compute the SNR of the test statistic using (4):

$$SNR_\lambda^2 = \frac{1}{8} \left[ tr\left[ \left( K_+^{-1} - K_-^{-1} \right) \left( K_+ - K_- \right) \right] \right]^2 \left[ tr\left[ \left( K_+^{-1} \right)^2 \right] + tr\left[ \left( K_-^{-1} \right)^2 \right] \right]$$

(10)

where $tr[\cdot]$ is the trace of the matrix. The expression in the denominator made use of Isserlis’s formula for fourth order moments of Gaussian distributed signals. Eq. 10 describes how echo signal textures due to the system and object inter-relate to affect task performance. It is important because it extends the Smith-Wagner theory to include pixel correlations, and in fact the SNR reduces to the Smith-Wagner model when covariances are diagonal and stationary, and pixels are sampled on the same scale as correlation lengths. It is not a very intuitive model, however, and could be computationally expensive to evaluate due to the inverse covariances needed. Our ultimate goal is to gain more intuition of ideal observer performance and so we consider how other approaches could complement this analysis.

5. CLARKSON-BARRETT APPROACH

Here we summarize an approach to SNR analysis taken by Clarkson and Barrett\(^17\). One contribution of our paper is to point out the applicability of the Clarkson-Barrett theory to ultrasound performance assessment. We shall also use this theory later on to understand a more intuitive picture of image quality.
The ideal observer is dependent on the probability distributions of the test statistic – conditioned on the hypothesis that the signal is present or absent. Alternatively, instead of pdf’s, moment-generating functions for the log-likelihood may be used. They are essentially the Laplace transform of the pdf’s,

\[ M_+(\beta) = \langle \exp(\beta \lambda) \rangle_+ \quad \text{and} \quad M_-(\beta) = \langle \exp(\beta \lambda) \rangle_- \]  

and hence are also related to the characteristic functions (Fourier transform) of the pdf’s (likelihoods). Because probability must be conserved (probability of lesion present + probability of lesion absent = 1), all the information about the area under the ROC curve is contained in one of the moment-generating functions. The moment-generating functions can consequently be written in terms of a single function, called the likelihood generating function \( G(\beta) \):

\[ M_+(\beta) = \exp[\beta(\beta - 1)G(\beta - 1/2)] \quad \text{and} \quad M_-(\beta) = \exp[\beta(\beta + 1)G(\beta + 1/2)] \]  

Both moment-generating functions and likelihood generating functions are useful for quantifying detectability of the signal with stochastic calculations. Different kinds of metrics can be used to quantify detectability. We have already discussed SNR\(_L\). The area under the ROC curve (AUC) is also a useful measure of detectability. Clarkson and Barrett show that SNR\(_L\) has the unusual behavior that it does not increase without bound as signal strength increases. This SNR is also not invariant to monotonic transformation of the decision variable, whereas the AUC is. Another SNR is given by

\[ \text{SNR}_{G(0)} = \left[ 2G(0) \right]^{1/2}, \]  

where

\[ G(0) = -4 \log M_-(1/2). \]  

This SNR is invariant under monotonic transformations of the decision variable, and is related to the Bhattacharyya distance, a metric quantifying the “distance” between two general probability density functions \( p_1(x) \) and \( p_2(x) \):

\[ d_\theta(p_1, p_2) = -\log \left\{ \int dx [p_1(x) p_2(x)]^{1/2} \right\}. \]  

When the two distributions have no overlap, there is very good separability. In this case, the integral tends to zero, and the Bhattacharyya distance becomes infinite. When there is complete overlap, the distributions are identical and normalization constrains the integral to 1. Consequently the Bhattacharyya distance tends to zero. For our task, the essential moment generating function is

\[ M_-(\beta) = \frac{\sqrt{\det \mathbf{K}_-^{\beta-1}} \sqrt{\det \mathbf{K}_+^{-\beta}}}{\sqrt{\det ([1 - \beta] \mathbf{K}_-^{-1} + \beta \mathbf{K}_+^{-1})}} \]  

and the corresponding signal-to-noise ratio is

\[ \text{SNR}_{G(0)}^2 = 2 \log \left[ \frac{(\det(\mathbf{K}_+ + \mathbf{K}_-))^{2}}{2^{2N} \det \mathbf{K}_+ \det \mathbf{K}_-} \right]. \]  

Intriguingly and conveniently this SNR does not require inversion of any covariance matrices. It’s computational evaluation may nevertheless be prohibitive for non-stationary covariance matrices due to expensive determinant operations.

6. ADDITIVE SIGNAL HOTELLING (IDEAL) OBSERVER

One unsatisfying aspect of the derived figures of merit for the quadratic task of detection is that the imaging system parameters are buried inside covariance matrices. Hence interpretation and intuition for system optimization is not obvious. Moreover, for the problems we are interested in, the object statistics are non-stationary (statistical properties of the image are location-dependent), hence the echo signals themselves are non-stationary. The advantage of stationarity is that diagonalization is natural and algorithmically efficient with a discrete Karhunen-Loève expansion. Lacking such stationarity, computational evaluation of image quality metrics can be prohibitive.

Another approach to modeling the signal detection problem is as follows. Consider an object function for the signal absent that is a zero-mean stationary stochastic process \( \gamma_b(x) \) with covariance \( \mathbf{K}_b \). For the signal present case consider
that the object function can be modeled as the random background $\gamma_b(x)$ plus an additive signal $\Delta\gamma(x)$ that is one deterministic realization of a random process (the statistical properties of this random process may vary from point to point in space). In essence, we are using the concept that a multiplicative signal in the variance can be represented as an additive signal. The advantage of this is that we are now back to a problem that is well known in the image quality literature: an additive signal $s = H(\Delta\gamma(x))$ and stationary covariance $K = H K_s H + K_n$, where $K_n$ is the noise covariance. The problem is descriptive of hyper-echoic lesions (lesions where the backscattered intensity is greater than the surrounding tissue background). The optimum linear observer is the Hotelling observer which has test statistic

$$\lambda_{Hot} = s^t K^{-1} g$$

with SNR given by

$$[\text{SNR}_{Hot}^2] = tr\left[K^{-1}ss^t\right].$$

It has been shown that this SNR is identical to the ideal observer SNR for non-random signals in normally distributed random backgrounds. The strategy for the Hotelling observer is to perform matched filtering with a pre-whitening step. Assuming local shift-invariance, $K$ is stationary, and a frequency space description of the SNR is possible via a Fourier transform. Additionally, we average over signal realizations to obtain:

$$\left<\text{SNR}_t^2\right> = \int d\xi \frac{\left<|\Delta\Gamma(u)|^2\right>}{|H(u)|^2 S_{\gamma_b}(u) + S_n(u)}.$$  

Similar averaging of $\text{SNR}^2$ has been done before except over location uncertainty. Here $\Delta\Gamma(u)$ is the Fourier description of the object function signal $\Delta\gamma(x)$, $S_{\gamma_b}(u)$ is the power spectrum of the background object function process, and $S_n(u)$ is the noise power spectrum. $H(u)$ is the system response function. The integration is done out to the extent of the sampling frequency. This is an ensemble average over a generalized Wagner-Brown model of detectability.

Before examining the interesting properties of this SNR, we first examine the expectation value of the magnitude of the Fourier domain target signal in the numerator of the integral. Consider that the spatial domain target signal is given by $\Delta\gamma(x) = w(x)\xi(x)$, where $w$ is a deterministic window function and $\xi$ is stochastic process. The purpose of writing $\Delta\gamma$ in this form is to characterize the location dependence of target signal statistical properties. For example, $w$ could be a disk signal that is unity inside the lesion and zero outside the lesion. In general, the Fourier domain stochastic average over target signal variations can be written as

$$S_{\Delta\gamma}(u) \equiv \frac{1}{X} \left<|\Delta\Gamma(u)|^2\right> = \frac{1}{X} \left<|W(u)*\Xi(u)|^2\right> = \frac{1}{X} \int d\xi d\eta W(p)W^*(p)\Xi(u-p)\Xi^*(u-q)$$

where $W$ and $\Xi$ are the Fourier transforms of $w$ and $\xi$ respectively, and $X$ is the spatial region of support over the image (the image area or volume).

At this point, we look at some simplifying assumptions to compare our results with those of the Smith-Wagner theory. If in particular $\xi$ is a white Gaussian noise (WGN) process with variance $\sigma_\xi^2$, then $\Xi$ is also WGN. The WGN assumption is a good one when spatial variations of density and compressibility are on a much smaller scale than the resolution of the imaging system and smaller than the sampling intervals used. With this assumption,

$$S_{\Delta\gamma}(u) = \sigma_\xi^2 \int dw(x)|w(x)|^2,$$

by Parseval’s theorem. If, for example, we model disk signals as Smith and Wagner did, such that $w(x)$ is unity where the lesion is present and zero where it is absent, we have $S_{\Delta\gamma}(u) = \sigma_\xi^2 A$, where $A$ is the area of the lesion (or volume of the lesion if considering 3-dimensions). Thus, when the object statistics can be treated as WGN, $S_{\Delta\gamma}(u)$ is a constant, that is proportional to the object variance and the area of the lesion. Hence,
\[
\text{SNR}_{\text{GNEQ}}^2 \equiv \left\langle \text{SNR}_f^2 \right\rangle = \sigma_z^2 AX \int \frac{|H(u)|^2}{|H(u)|^2 S_b(u) + S_n(u)} = \sigma_z^2 AX \int dGNEQ(u). \tag{23}
\]

We have written this quantity as an integration over what has been termed the **Generalized Noise-Equivalent Quanta**\(^1\) or **GNEQ**, defined as the integrand of the third term. **Noise-Equivalent Quanta (NEQ)** has historical origins with Shaw and others, as well as the Wagner-Brown theory of detectability\(^3\). For photon imaging systems **NEQ** represents the frequency-specific density of quanta at the input of an ideal detection system that would yield the same output noise as the real system under evaluation. **Generalized NEQ** as described by Barrett and colleagues\(^6,11\) provides provision for a stochastic background texture of the object. For ultrasound systems the **GNEQ** quantity is a measure of the spatial frequency sensitivity of detecting a signal in a background texture and in the presence of electronic noise. For photon imaging modalities, the ideal observer detectability is given by an integral over the frequency-domain of **NEQ** (or **GNEQ**) times the squared magnitude of the Fourier transform of the signal. For ultrasound systems, the **GNEQ** is not weighted by the magnitude of the Fourier transform of the signal shape, but rather by the spectral variance of the target signal. This is to be expected since ultrasound systems are not sensitive to the magnitude of density or compressibility but rather depend on the variance (spatial fluctuations) of these quantities. The entire **GNEQ spectrum** can be used to characterize ultrasound systems for a particular tissue type in a target-independent manner. Defining \( \sigma_b^2 \) as the integrated background power spectral density, we can also re-write the **SNR** as:

\[
\text{SNR}_{\text{GNEQ}}^2 = C_{\text{obj}} A \int dGNEQ(u) \frac{MTF^2(u)}{MTF^2(u) S_b(u) + \frac{1}{X\sigma_b^2 S / N_b(u)}} \tag{24}
\]

where \( C_{\text{obj}} \) is the object function contrast, defined as \( C_{\text{obj}} = \sigma_z^2 / \sigma_b^2 \) and **MTF** is the modulation transfer function defined as

\[
MTF(u) = \left|H(u)\right| / H_{\text{max}} \tag{25}
\]

where \( H_{\text{max}} \) is the maximum value of \( |H(u)| \). The background signal to noise ratio is defined as:

\[
S / N_b(u) = \left|H_{\text{max}}\right|^2 X\sigma_b^2 / S_n(u), \tag{26}
\]

and is a measure of the backscattered signal strength of the object medium relative to the noise level. Note the connections between our expression for **SNR** and the Smith-Wagner theory: both \( \text{SNR}^2 \) measures are proportional to lesion area. Curiously, \( [\text{SNR}_{\text{GNEQ}}^2] \) is proportional to contrast, yet in the Smith-Wagner theory, the \( \text{SNR}^2 \) is proportional to the square of object contrast. Both metrics, however, reveal that task performance is improved with improved contrast.

One elegant feature of this formalism is the direct dependence of system and noise properties on the detectability. Also, the description is in the Fourier domain, a natural choice for analyzing ultrasound imaging systems because focused fields in the Fresnel region and unfocused apertures in the far-field (Fraunhofer) region have k-space descriptions equivalent to the Fourier transform of the aperture. Also, the quantities involved in the **GNEQ(u)** are all measurable quantities, so this could provide a convenient way to characterize a system design experimentally! To do so we would need to measure the **MTF(u)** and the normalized background plus noise power spectrum (**NBNPS**), which we define as:

\[
\text{NBNPS}(u) \equiv \left|H(u)\right| S_b(u) + S_n(u) / |H_{\text{max}}|, \tag{27}
\]

then take the ratio of the two:

\[
\text{GNEQ}(u) = \frac{\text{MTF}(u)}{\text{NBNPS}(u)}. \tag{28}
\]

Wire or bead phantoms can be used to estimate the point-spread functions – and consequently the **MTF**. It may be necessary to acquire a small number of images to average over noise properties. Assuming local ergodicity, a patch of one or a few (RF) images may be used to estimate the background+noise power spectrum for a particular field region. Once we know these properties, the integrated **GNEQ** or **IGNEQ** as we shall call it may be computed:

\[
\text{IGNEQ} = \int dGNEQ(u). \tag{29}
\]
This metric may apply to media where object statistics can be treated as nearly white, such as standard phantoms and many types of tissues. Conveniently, this metric is independent of the signal size and contrast, and characterizes the performance of the imaging system with a single number. It may be normalized by $\sigma_b^2$ to obtain a number useful for simulation studies, which we shall call the normalized $IGNEQ$. Tissues that have complex striations or variable patches of brightness may make target discrimination more difficult. Object texture can be accounted for in the theory and measured experimentally.

### 7. MONOTONICITY WITH THE FULL QUADRATIC TASK

Recall that when we modeled both the lesion and its surroundings as Gaussian stochastic processes the resulting ideal observer test statistic was a quadratic form in the data. In contrast, when we modeled the lesion as an additive deterministic signal, the test statistic was linear in the data. We may very well expect that performance metrics for these two detection tasks may not exhibit identical properties. Consequently, we now examine whether the figures of merit derived for both tasks give similar information about diagnostic performance. To do so we choose to examine the trends of both the Clarkson-Barrett $SNR$ and the $GNEQ$ theory with varying system and object parameters. Consider the case where both $K_+$ and $K_-$ are simultaneously diagonalizable with a Karhunen-Loeve transformation. In assuming this we consider the case where the size of the lesion is the same size as the window $[w(x)=1 \text{ everywhere}]$, and also that the system is locally shift-invariant. The result is a spectral description:

$$SNR_{G(u)}^2 = 2\log \left\{ \frac{\prod_{m=1}^{M} [H(u_m)]^2 \left( S_+(u_m) + S_-(u_m) \right) + 2S_n^2(u_m)}{2^{2M} \prod_{m=1}^{M} [H(u_m)]^2 S_+(u_m) + S_n^2(u_m) \prod_{m=1}^{M} [H(u_m)]^2 S_-(u_m) + S_n^2(u_m)} \right\}, (30)$$

where $S_+(u)$ and $S_-(u)$ are the eigenvalues (spectrum) of $K_+$ and $K_-$ respectively, and $u$ is a (spatial) frequency vector of length $M$. Note that $S_-(u) = S_b(u)$ and $S_+(u) = S_b(u) + S_{S_n}(u)$, where $S_b(u)$ is the spectrum of the background object function, and $S_{S_n}(u)$ is the spectrum of the target. Also, $S_n(u)$ is the noise-power spectrum (NPS). The spectral description now gives a means for comparing with $SNR_{GNEQ}$. In particular, we shall examine trends of both $SNR_{G(0)}$ and $SNR_{GNEQ}$ for parameters of bandwidth, noise power, and target signal power.

To investigate the behavior of the $SNR$ with system bandwidth ($BW$), consider for simplicity an ideal low-pass system response:

$$H(u) = \begin{cases} 1 & 0 \leq |u| \leq u_m \\ 0 & \text{otherwise} \end{cases}. (31)$$

Suppose also that object and noise statistics are white: $S_+(u) = S_+$, $S_-(u) = S_-$, and $S_n(u) = S_n$ for all $u$. Then

$$SNR_{G(u)}^2 = 2m \log \left\{ \frac{\left[ (S_+ + S_-) + 2S_n^2 \right] S_+^2}{2^2 \left[ S_+ + S_n \right] \left[ S_- + S_n \right]} \right\} + 2(M-m-1) \log \left\{ \frac{2S_n^2}{2^2 S_n^2} \right\}. (32)$$

The second term vanishes since the log of unity is zero. Consequently, the $SNR^2$ is proportional to $m$, which is simply proportional to $u_m$, the system bandwidth.

To show that $SNR_{G(0)}$ decreases with increasing noise, we took the derivative with respect to the noise power and showed that the result was a strictly negative quantity, indicating that the slope of the graph of $SNR_{G(0)}$ with respect to noise was negative (calculation not shown). As the noise grows larger without bound, both $SNR_{G(0)}$ and $SNR_{GNEQ}$ both tend to zero. As noise power tends to zero, $SNR_{G(0)}$ and $SNR_{GNEQ}$ increase to a limit that is only bound by the density of sampling, and the Nyquist limit of $H$. If $H$ is non-zero for frequency channels $u_m \in \Omega$, then as noise tends to zero:
\[ \text{SNR}_{G(0)}^2 \xrightarrow[\text{noise} \to 0]{} \propto \int_{\Omega} d\mathbf{u} \frac{|H(\mathbf{u})|^2}{|H(\mathbf{u})|^2 S_b(\mathbf{u})} = \int_{\Omega} d\mathbf{u} \frac{1}{S_b(\mathbf{u})}. \tag{33} \]

As the sampling frequency grows, so does the SNR as long as \( H \) has even a small amount of signal. The same behavior is seen with \( \text{SNR}_{G(0)} \):

\[ \text{SNR}_{G(0)}^2 \xrightarrow[\text{noise} \to 0]{} \sum_{m,n \in \Omega} 2\log \left\{ \frac{\left| S_0(\mathbf{u}_m) + S_+(\mathbf{u}_m) \right|^2}{2\left| S_0(\mathbf{u}_m) + S_-(\mathbf{u}_m) \right|^2} \right\}. \tag{34} \]

As the number of frequency channels in \( \Omega \) increases, this quantity will in general increase until both target and background object power spectra tend to the same number, for example, as they die off to zero. In the Smith-Wagner theory, noise-power can be totally neglected, and the detectability was bounded not by the sampling density, but by the number of independent speckle spots. By using the RF signal instead of the envelope signal, the ideal observer has the potential to compensate for the blurring kernel of the imaging system. Noise prevents the ideal observer from doing this completely. One interpretation of the ideal observer strategy is that it deconvolves the effects of the imaging system before performing template matching. Deconvolution cannot necessarily recover precise information lost in the null-space but it can restore image statistics useful for the detection task. An interesting observation related to deconvolution is that the argument of the \( \text{SNR}_{G(0)} \) integral looks very much like a Weiner filter. In practice, there will always be noise, and the ideal observer detectability will be bounded. The imaging system should be designed so as to sample at minimum at the Nyquist rate, determined by the system’s MTF. Sampling more densely than the Nyquist rate will add noise power to the images (which may degrade task performance for human observers). For ideal observers, sampling at greater than the Nyquist limit will not degrade performance, but will not provide any benefit either. This is because the ideal observer will know to truncate spectral noise where there is no signal.

Lastly, we note the increasing trends of both \( \text{SNR}_{G(0)} \) and \( \text{SNR}_{G(0)} \) with increasing signal strength. \( \text{SNR}_{G(0)} \) is proportional to target signal variance. For \( \text{SNR}_{G(0)} \), the relationship with target strength is more subtle, however, by differentiating with respect to signal strength, one can show that the relevant slope is positive (calculation not shown).

The conclusion of the previous analysis is that \( \text{SNR}_{G(0)} \), a figure of merit for a slightly different task than the full quadratic task modeled in sections II, IV, and V gives similar information as do metrics for the quadratic task. The benefits are that small targets (lesions) can be considered, noise and system parameters can be included directly and intuitively into the analysis, and the SNR can be both easily measured and predicted with a target-independent metric.

### 8. Shift-variant Information Maps of the Ultrasonic Field

Realistic ultrasound systems are not shift-invariant. But we can assume that they are shift-invariant over some local region. By doing so, we can compute the potential information content of imaging a lesion at each field location. We considered a 128-element linear array transducer with 64 active elements and fixed focus at 60 mm. No aperture growth, dynamic receive focusing, elevation focusing, or apodization were employed. Attenuation was neglected, and we only consider 2 dimensions of a truly 3-D ultrasound beam sensitivity. All such effects could easily be integrated into future simulations. Point spread functions (psf’s) simulated using FIELD II\(^{18}\), a public domain ultrasound simulator, are shown below in Fig. 1(a). This is a B-mode image of point scatterers located at successive axial depths separated by 5 mm intervals. At each point the MTF was computed, corresponding to the normalized 2-D FFT of the RF point-spread functions. Assuming white background and noise processes, and \( S/N_b = 1/30 \) at the focus, a normalized IGNEQ value was computed at each axial point, and the results plotted in Fig. 1(b).

The results are somewhat counter-intuitive. We would have expected the focal region to be optimally informative since it has the most focused resolution size. Instead we find that the nearfield has remarkably more information potentially available. To understand this unexpected result, we compared the nearfield and focal region psf’s and MTF’s, as shown in Fig. 2. Intuition tells us that 2(a) has better resolution, but upon examining the spatial frequency domain, we need to give the question some more thought. The effective lateral spatial bandwidth in (d) is 74% greater than in (b). Thus – the lateral spatial resolution of the large curved wavefront (c) is potentially greater than the lateral resolution of the smaller focal region field (a)! To understand this better, notice that a lateral slice through the psf’s looks very much like a chirp...
Figure 1. (a) Point spread functions due to 64 active elements of a 128-element linear array transducer of height 5 mm, element width $\lambda$, and gap spacing of 0.1 mm, with fixed focus at 60 mm, and no elevation lens. An attenuationless medium was considered for simplicity. 2 cycles of a 3 MHz sinusoid weighted by a Hanning window were used to simulate the excitation pulse. The same function was used to simulate the electromechanical coupling impulse response of the transducer. (b) The normalized IGNEQ values corresponding to field points along the beam axis. The larger the normalized IGNEQ value the more informative the system is for the given field point.

Figure 2. (a) Focal and (c) nearfield RF psf's due to a 3 MHz fixed focus linear array of height 0.5 cm, with 64 active elements of width $\lambda$ separated by distances of 0.1 mm. Azimuthal focus was 6 cm. The beam was electronically swept laterally across the point target located at (a) 6 cm and (b) 3 cm. (b) and (d) are the k-space (MTF) representations of (a) and (c) respectively.

function. The resolution must be recovered by processing similar to current coded excitation schemes\textsuperscript{19}. Similar to Jensen\textsuperscript{20} we propose a matched filter technique, whereby a time-reversed copy of an RF psf (c) is used as a filter. When convolved with an RF image of a point target 2(c), a much narrower lateral distribution 3(b) is the result. The signal
amplitude in 3(b) is also more than 20 times greater than that in 3(a), and 2.2 times greater than the focal gain. A marked improvement in visibility is seen when matched filtering is applied to simulated images of a cyst phantom, shown in Figs. 3(c) and (d). The matched filtering procedure is consistent with the strategy for the ideal or Hotelling observer, which is to whiten and match filter. In the nearfield, matched filtering without whitening performs well because the resulting image statistics are nearly white. It remains to be seen how robust the technique is to phase aberrations and other artifacts.

Regardless of whether spatial matched filtering is a practical technique, it shows an important point – that image processing can make information in the raw RF data more accessible to human observers, as seen in Fig. 3(d). Processing does not, however, increase the information content. The raw RF signal contains all the task-relevant data. Processing can at best maintain information content, and often, information will be lost. We should comment that beamforming schemes (at least for reception) can be viewed as either part of the system design or as part of the processing algorithms. Both views may have merit, and further research in this area could be pursued.

Before concluding, we comment on the general downward trend of detectability with increasing axial distance as shown in Fig. 1(b). Diffractive losses mean that the farfield has less signal available than does the nearfield, and consequently the $S/N_b$ decreases, dropping the overall detectability.

One result of our analysis is an expression for the ideal observer detectability that is a generalization of Noise-Equivalent Quanta ($NEQ$), a measure widely used by other imaging modalities. The detection signal to noise ratio is shown to be an integration of the generalized $NEQ$ weighted by the spectral variance of the target signal (rather than the squared magnitude of the Fourier transform of the signal, as is the case with other modalities). This reflects that ultrasound systems are sensitive not to the magnitude of the medium parameters but rather depend on the variance (spatial fluctuations) of these quantities. Moreover, when object statistics are white over the sampling domain, the frequency-dependent weighting of $GNEQ$ becomes constant, and a target-independent picture of detectability can be summarized in a single number. The resulting framework behaves similarly to the quadratic task, is intuitive, and is amenable to measurement and prediction of system performance. The theory has been used to predict the information content of the ultrasonic beam at various field points. Similar analysis of location-dependent $IGNEQ$ may prove useful for evaluating

9. DISCUSSION AND CONCLUSIONS

Figure 3. (a) The envelope-detected $psf$ in 2(c). (b) The resulting image obtained after matched filtering 2(c) with a time-reversed replica filter. (c) Pre-filtered image of a lesion due to $psf$ 2(c). (d) The corresponding post-filtered image.
the effectiveness of beamforming strategies. Currently lesion detectability is assessed with time-consuming Monte-Carlo simulations, or by using phantoms with real systems. Backscattered power inside and outside the lesion are measured in ratio to give a detectability metric. The figures of merit presented here are instead based on statistical models and have a different intention: they give an upper bound on the information content available with a particular system design. The theory does not attempt to predict human observer performance. However, new strategies mimicking the ideal observer are revealed for processing RF data that could improve detection of lesions for human observers. Experimental characterization of ultrasound systems is planned in the near future.

ACKNOWLEDGEMENTS

This work was supported in part by the National Institutes of Health R01 CA 82497.

REFERENCES