

Detectability index describes the information conveyed by sonographic images

Nghia Q. Nguyen,^{a,c} Craig K. Abbey,^d and Michael F. Insana^{a,b,c}

^aDepartment of Electrical and Computer Engineering, ^bDepartment of Bioengineering

^cBeckman Institute for Advanced Science and Technology,
University of Illinois at Urbana-Champaign, USA

^dDepartment of Psychology and Brain Sciences, University of California, Santa-Barbara, USA

Abstract— We have been developing the ideal observer formalism for sonography, which is based on the best-possible diagnostic performance. The ideal performance was compared to that of trained human observers to estimate the visual efficiency for discriminating lesion features. We find that humans are generally less than 10% efficient at accessing visual information essential for breast cancer diagnosis. In seeking ways to improve this process, we must first establish a connection between standard ROC observer metrics and instrument properties used in system design. In radiography, that relationship is made through the lesion signal-to-noise ratio SNR_I . SNR_I^2 , which describes task information, is simply related to contrast and spatial resolutions and noise power. Those relations break down for sonography due to the quadratic form of the ideal observer. Our goal in this paper is to establish a rigorous connection between ideal performance and engineering design metrics, which has directly applications for sonographic system design and optimization.

Index Terms - Breast sonography, ideal observer, image quality, Kullback-Leibler divergence, task-based design.

I. INTRODUCTION

In task-based design, the quality of a medical imaging system is measured in the context of a clinical task by observer performance, assessed through the gold standard method of ROC analysis [1]. The ROC curve depicts the probability of detection as a function of false-alarm rate. For binary decisions that are normally distributed with equal variance, the area under the curve (AUC) is adopted as a scalar metric of overall task performance.

Wagner and Brown [2] established a connection between ideal performance and instrument properties through the lesion signal-to-noise ratio for the ideal observer, SNR_I , which is monotonically related to AUC and is equal to the detectability index d_A if the test statistic is normally distributed. They showed that SNR_I defines the information in the visual tasks, and factorized into the spatial resolution, noise power, and large-area contrast resolution of the instruments. These ideas have formed the foundation of medical image quality analysis for the last 30 years as they applied to most medical imaging modalities including radiography, computed tomography, magnetic resonance, and position emission tomography. Their work was successfully extended to sonography and applied to B-mode images, but it was exact only under very stringent conditions [3].

Applying the analysis to sonography, but now in the radio-frequency (RF) domain, we could relax many of the stringent assumptions [4], [5]. However, the quadratic form of the ideal observer with respect to echo data is inconsistent with the Wagner-Brown theory. In this paper, we introduce the *Kullback-Leibler divergence*, J [6], a concept rooted in information theory, to define task information for sonography. We use Monte-Carlo studies involving RF echo data to compute J and connect it to the ideal observer performance. J is then related analytically to instrument properties. In this way, we can rigorously connect task information to both human observer performance and engineering design properties for typical visual tasks required to discriminate malignant from benign breast lesion features [4]. The goal of this paper is to further the development of a rigorous framework for medical sonography on par with other modalities.

II. IDEAL OBSERVER ANALYSIS

Details of the ideal observer analysis for breast lesion detection in sonography were provided previously [4], [5].

SIGNAL MODELING. In consultation with a radiologist, we selected five sonographic features characteristic of breast lesion diagnosis. These features define five visual tasks, and are graphically illustrated in low-right corner of Fig. 2. Task 1 is detection of a low-contrast circular lesion; Tasks 2-4 are discriminations between lesion boundary features; and Task 5 is to discriminate between high-contrast hypoechoic and anechoic lesions. More complex lesion features can be synthesized from these five elementary features.

Each task includes two classes of data specified by a pair of masks $S_i(x, y)$ that define the geometry and contrast of feature patterns for $i = 1$ malignant and $i = 0$ benign features. Task difficulty is controlled through the object contrast factor defined by the difference between two masks, $\Delta S(x, y)$ [4]. Object scattering functions for the i^{th} class of data $f_i(x, y)$ are formed by multiplying the mask by white Gaussian noise fields ($\text{WGN} \sim \mathcal{N}(0, \sigma_{\text{obj}}^2)$), such that $f_i(x, y) = \text{WGN}(x, y) \times \sqrt{S_i(x, y)}$. Consequently, the scattering functions for each class f_i are amplitude-modulated (nonstationary) random fields.

We simulated RF echo data vectors for the i th class \mathbf{g}_i from sampled random object scattering vector $\mathbf{f}_i \sim \mathcal{N}(\mathbf{0}, \Sigma_{\text{obj}})$, with covariance matrix Σ_{obj} . Modeling the ultrasonic imaging system as a noisy linear transformation, the RF data are obtained through $\mathbf{g}_i = \mathbf{H}\mathbf{f}_i + \mathbf{n}$. System matrix \mathbf{H} transforms scattering vectors \mathbf{f}_i into RF echo data \mathbf{g}_i that includes additive, signal-independent, zero-mean acquisition noise $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma_n^2 \mathbf{I})$, where \mathbf{I} is the identity matrix. Noise variance σ_n^2 is adjusted relative to object background variance σ_{obj}^2 to produce a background echo $\text{SNR}_0 = 32$ dB; SNR_0 is the ratio of signal energy to noise energy. $N \times N$ object and $M \times M$ data fields are lexicographically reordered to form $N^2 \times 1$ and $M^2 \times 1$ vectors, respectively, \mathbf{f} and \mathbf{g} . \mathbf{H} has dimensions $M^2 \times N^2$, and has rows composed of pulse-echo impulse responses. If $M = N$ and the impulse response is shift invariant, \mathbf{H} is a block-Toeplitz matrix that we approximate as circulant for fast computation. System parameters were selected to model a 1-D linear array on a Siemens Antares system [5].

ENCODING TASK FEATURES IN SCATTERING FUNCTIONS. In sonography, received signals from random media are produced by the detection of mostly incoherent acoustic backscattered wave energy [4]. Feature contrast in random scattering objects therefore is modeled using a nonstationary covariance matrix, $\Sigma_{\text{obj}} = \sigma_{\text{obj}}^2(\mathbf{I} + \mathbf{S}_i)$. \mathbf{S}_i is a diagonal matrix formed by re-ordering the mask $S_i(x, y)$, which defines the shape and amplitude of lesions as distinct from background media. Σ_{obj} is diagonal with nonzero elements given background variance σ_{obj}^2 of the WGN process that are modulated by the addition of \mathbf{S}_i .

IDEAL OBSERVER. The optimal discriminator of these two classes of signals is given by the Bayesian ideal observer [1]. The covariance matrix of \mathbf{g}_i is found from two component covariances,

$$\Sigma_i = \mathbf{H}\Sigma_{\text{obj}}\mathbf{H}^t + \Sigma_n = \sigma_{\text{obj}}^2\mathbf{H}(\mathbf{I} + \mathbf{S}_i)\mathbf{H}^t + \sigma_n^2\mathbf{I}, \quad (1)$$

if we can model imaging as a linear system. Denoting the multivariate normal probability densities for \mathbf{g}_i as $p_i(\mathbf{g})$, the ideal observer operating on RF echo signals is given by the log likelihood ratio [1]

$$\lambda(\mathbf{g}) = \ln \frac{p_1(\mathbf{g})}{p_0(\mathbf{g})}. \quad (2)$$

Reducing (2) and eliminating terms unrelated to \mathbf{g} , because they do not modify detection performance, we adopt the simpler test statistic, T , given by

$$T(\mathbf{g}) = \frac{1}{2}\mathbf{g}^t(\Sigma_0^{-1} - \Sigma_1^{-1})\mathbf{g}. \quad (3)$$

The ideal observer, which makes decisions by comparing T to a threshold, achieves the largest AUC for any observer of this task. It is significant that T is a quadratic function of \mathbf{g} , which occurs because feature contrast in encoded in the object covariance Σ_{obj} rather than its mean.

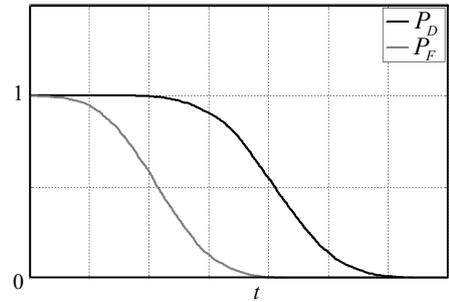


Fig. 1. Plots of the detection probability P_D and the false alarm probability P_F as functions of threshold t , where $-\infty < t < \infty$. J equals the area between the two curves.

PERFORMANCE METRICS. Visual discrimination performance for different processing approaches was quantified from the results of two-alternative force-choice (2AFC) experiments [4]. From the percentage of correct responses, we measure AUC. The standard detectability index d_A of the ideal observer is converted from AUC via [1]

$$d_A = 2\text{erf}^{-1}(2\text{AUC} - 1), \quad (4)$$

where $\text{erf}(\cdot)$ is the error function and $\text{erf}^{-1}(\cdot)$ is its inverse. d_A^2 is a psychophysical measure for the ideal observer that is compared with that for a test observer to estimate absolute discrimination efficiency. d_A^2 ranges from 0 to ∞ as AUC moves from 0.5 to 1.

When observer responses are normally distributed, AUC can be calculated through SNR_I , which is defined from moments of the test statistic (3),

$$\text{SNR}_I^2 = \frac{(E_1\{T\} - E_0\{T\})^2}{(\text{var}_1\{T\} + \text{var}_0\{T\})/2}, \quad (5)$$

where $E_i\{T\}$ and $\text{var}_i\{T\}$ are means and variances of the test statistic conditioned on class i being true. SNR_I is the separation between response means in units of their average variance. For normally distributed response data, $\text{SNR}_I = d_A$. Also for normally distributed T , AUC is related to d_A and SNR_I , which then connects task information, discrimination efficiency, and instrument properties defining image quality.

III. TASK INFORMATION IN SONOGRAPHY

T is a quadratic function of \mathbf{g} in sonography, and thus it follows a noncentral chi-squared distribution [1] that deviates from normal more for edgy tasks 2 and 4. Thus relationships between d_A , AUC and SNR_I in sonography are more complex than in radiography. Which of these figures of merit characterizes performance? The answer is d_A , which equals the Kullback-Leibler divergence J [6]. J measures task information under 2AFC experimental conditions and therefore it may be used to establish the validity of other figures of merit for conditions where T is not normally distributed.

In terms of the probability density function for the two classes,

$$J = \int d\mathbf{g} (p_1(\mathbf{g}) - p_0(\mathbf{g})) \ln \frac{p_1(\mathbf{g})}{p_0(\mathbf{g})}. \quad (6)$$

The difficulty in calculating J is that it involves the inverses and determinants of large-size covariance matrices that are difficult to compute. While inverses have been computed using a power series expansion [4], determinant calculations remain challenging. However, it can be shown that

$$J = \int_{-\infty}^{\infty} dt [P_D(t) - P_F(t)], \quad (7)$$

where $P_D(t) = \Pr(\lambda(\mathbf{g}) > t | H_1)$ is the probability of detection and $P_F(t) = \Pr(\lambda(\mathbf{g}) > t | H_0)$ is the probability of a false-alarm measured at decision threshold t .

Eq. (7) is illustrated graphically in Fig. 1. J can be found numerically by integrating the difference between the two probability curves. Covariance determinants are included as additive constants in the log-likelihood ratio. Their influence is to shift both curves along t by the same interval, and therefore they do not affect the area calculation. Consequently, we can plot P_D and P_F using $T(\mathbf{g})$ instead of $\lambda(\mathbf{g})$ to quickly and accurately calculate J .

We computed d_A^2 , SNR_I^2 , and J from (4), (5), and (7), respectively, and plotted the results in Fig. 2 as functions of object contrast. J , which defines task information, is statistically equivalent to d_A^2 for all tasks. Difference are maximized at 6.5% in Task 5 and within the computational errors. SNR_I^2 , however, is lower than the other metrics in Tasks 2 and 4. The difference is up to 14.5% in Task 2 and larger than computational errors.

The differences between SNR_I^2 and other figures of merit for Tasks 2 and 4 might be expected because the areas encompassed by the task difference $\Delta\mathbf{S}$ is small, suggesting the central-limit theorem does not produce a normally distributed T . The difference also suggests the relationship between SNR_I and d_A is not exact. Thus it is better to begin with J when attempting to related task information to image quality metrics, as we do next.

IV. SONOGRAPHIC IMAGE QUALITY

Wagner and Brown [2] derived the following expression that relates ideal observer performance to laboratory measurements of image quality such as spatial resolution, noise power, and signal contrast,

$$\text{SNR}_I^2 = \int_{-\infty}^{\infty} du \int_{-\infty}^{\infty} dv |\Delta\tilde{S}(u, v)|^2 \text{NEQ}(u, v), \quad (8)$$

where $\text{NEQ}(u, v) = |\tilde{H}(u, v)|^2 / \sigma_n^2$. $|\tilde{H}(u, v)|$ is the magnitude of the system transfer function and σ_n^2 is the variance of photon noise process. The expression is derived for a linear shift-invariant (LSI) system and stationary noise, and for low-contrast lesion detection (Task 1). NEQ characterizes the ability of the instrument to transfer object contrast into recorded data.

To derive an expression analogous to (8) for sonography, we express J in terms of class covariances and then image quality parameters. For the two class distributions with zero mean and covariance matrices Σ_i [7],

$$J = \frac{1}{2} \text{Tr} [(\Sigma_0^{-1} - \Sigma_1^{-1})(\Sigma_1 - \Sigma_0)]. \quad (9)$$

Inverting these large matrices can be achieved by a power series expansion. Since we are examining the low-contrast detection task we can truncate the expansion after the first term to approximate (9) in closed form as

$$J \approx \frac{1}{2} \text{Tr} [\Sigma_s^{-1} (\Delta\Sigma_1 - \Delta\Sigma_0) \Sigma_s^{-1} (\Sigma_1 - \Sigma_0)], \quad (10)$$

where Σ_s and $\Delta\Sigma_i$ given by

$$\begin{aligned} \Sigma_s &= \sigma_{\text{obj}}^2 \mathbf{H}\mathbf{H}^t + \sigma_n^2 \mathbf{I}, \\ \Delta\Sigma_i &= \sigma_{\text{obj}}^2 \mathbf{H}\mathbf{S}_i\mathbf{H}^t. \end{aligned} \quad (11)$$

Since Fig 2 shows that $d_A^2 \simeq J$ for all tasks and contrasts, we substitute d_A^2 for J and obtain

$$d_A^2 \approx \frac{1}{2} \text{Tr} [\mathbf{K}_s \Delta\mathbf{S} \mathbf{K}_s \Delta\mathbf{S}]. \quad (12)$$

where $\mathbf{K}_s = \mathbf{H}^t \Sigma_s^{-1} \mathbf{H}$ and $\Delta\mathbf{S} = \sigma_{\text{obj}}^2 (\mathbf{S}_1 - \mathbf{S}_0)$ defines task contrast.

Under the LSIV/stationary assumptions, \mathbf{K}_s can be diagonalized using Fourier techniques as $\mathbf{K}_s = \mathbf{F}^{-1} \tilde{\mathbf{K}}_s \mathbf{F}$, where \mathbf{F} is the forward DFT matrix [1]. Since $\tilde{\mathbf{K}}_s$ is diagonal, its elements can be represented by a single index, $\tilde{\mathbf{K}}_s(k, k) = \tilde{K}_s(k)$. Similarly, $\Delta\mathbf{S} = \mathbf{F}^{-1} \Delta\tilde{\mathbf{S}} \mathbf{F}$, and therefore (12) is expressed as a double sum over frequency indices

$$\begin{aligned} d_A^2 &\simeq \frac{1}{2} \text{Tr} [\tilde{\mathbf{K}}_s \Delta\tilde{\mathbf{S}} \tilde{\mathbf{K}}_s \Delta\tilde{\mathbf{S}}] \\ &= \frac{1}{2} \sum_k \sum_l \tilde{K}_s(k) \Delta\tilde{S}(k, l) \Delta\tilde{S}(l, k) \tilde{K}_s(l). \end{aligned} \quad (13)$$

Since $\Delta\mathbf{S}$ is diagonal, $\Delta\tilde{\mathbf{S}}$ is Hermitian and stationary, i.e. $\Delta\tilde{S}(l, k) = \Delta\tilde{S}^*(k, l) = \Delta\tilde{S}(l - k)$, in which $\Delta\tilde{S}(k)$ is the Fourier transform of $\Delta\mathbf{S}$ but re-arranged into a column vector before taking the transform.

Expressing (13) as a continuous function of 2-D spatial frequency variable, $\mathbf{u} = (u, v)$, we have

$$\begin{aligned} d_A^2 &\simeq \frac{1}{2} \int_{-\infty}^{\infty} d\mathbf{u} \int_{-\infty}^{\infty} d\mathbf{u}' \tilde{K}_s(\mathbf{u}') \left| \Delta\tilde{S}(\mathbf{u} - \mathbf{u}') \right|^2 \tilde{K}_s(\mathbf{u}) \\ &= \int_{-\infty}^{\infty} d\mathbf{u} \left| \Delta\tilde{S}(\mathbf{u}) \right|^2 \left\{ \frac{1}{2} \tilde{K}_s(\mathbf{u}') * \tilde{K}_s(-\mathbf{u}') \right\}(\mathbf{u}) \\ &= \int_{-\infty}^{\infty} d\mathbf{u} \left| \Delta\tilde{S}(\mathbf{u}) \right|^2 \text{ACF}(\mathbf{u}). \end{aligned} \quad (14)$$

where \tilde{K}_s is given by

$$\tilde{K}_s(u, v) = \frac{|\tilde{H}(u, v)|^2}{|\tilde{H}(u, v)|^2 \sigma_{\text{obj}}^2 + \sigma_n^2}. \quad (15)$$

Comparing (14) to (8), we find performance in both cases depends on task contrast $|\Delta\tilde{S}|^2$. However the autocorrelation

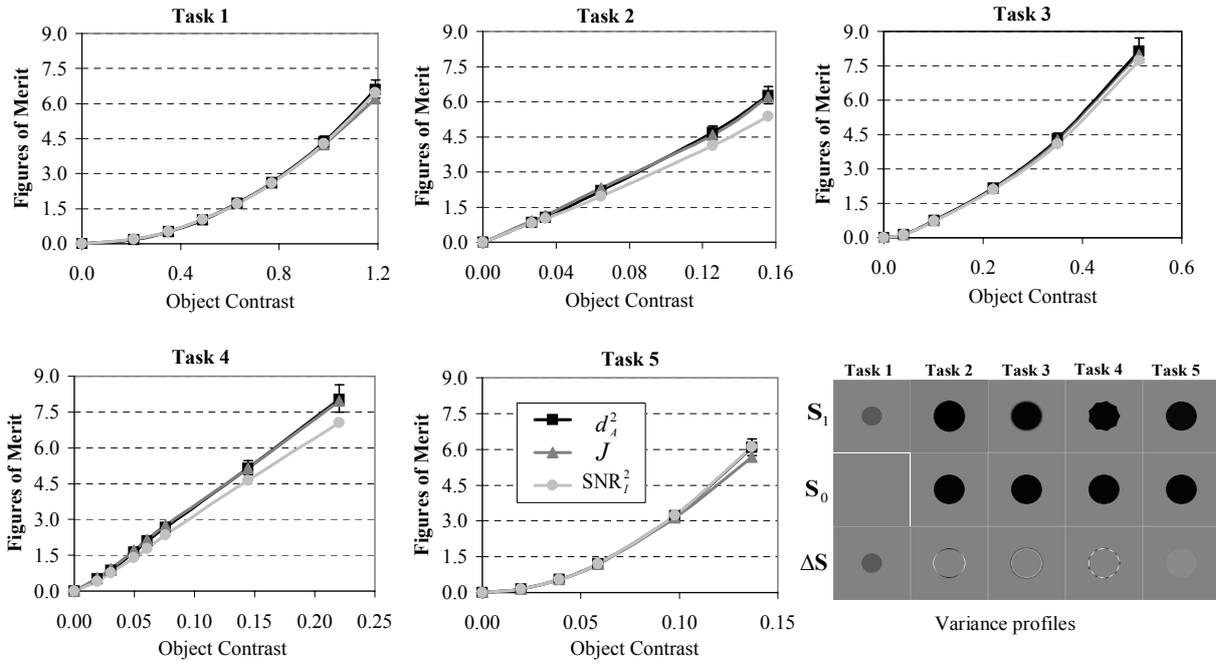


Fig. 2. Comparison of ideal observer performance metrics d_A^2 , J -divergence, and SNR_I^2 measured for five visual tasks. Curves are plotted as a function of object contrast. The legend for Task 5 applies to all. The lower-right corner shows variance profiles for five tasks related to lesion discrimination.

function $\text{ACF}(\mathbf{u}) = \frac{1}{2} \tilde{K}_s(\mathbf{u}) * \tilde{K}_s(-\mathbf{u})$ in sonography is analogous to NEQ in radiography. It provides an avenue for relating image quality metrics to task information.

V. DISCUSSION

When $|\Delta\tilde{S}|^2$ (see Fig. 2, low-right corner) is not much larger in area than the 2-D speckle correlation area, the central-limit theorem does not drive the observer response distribution toward normal. Therefore SNR_I in (8) is no longer directly related to the ideal observer AUC. We replaced SNR_I with the Kullback-Leibler divergence J , a fundamental measure of task information, and found through Monte Carlo studies that $J = d_A^2$, which is related to AUC, for the five discrimination tasks commonly found in breast lesion diagnosis. Thus AUC measured in 2AFC experiments can be related to task information J that can also be expressed in terms of image quality parameters. These are the relationships we sought.

$\text{ACF}(\mathbf{u})$ provides a rigorous connection between task information and image quality parameters. It may be interpreted as the number of independent samples of task information being offered to the observer at spatial frequency \mathbf{u} . $\tilde{K}_s(\mathbf{u})$ in Eq. (15) resembles a generalized NEQ quantity [1] for photon imaging in a variable background. Here acoustic speckle in the RF signal is considered to be a random background.

Interpretation is easier when \tilde{K} is written in the form

$$\tilde{K}_s(\mathbf{u}) = \frac{\text{SNR}_0 \times \text{MTF}^2(\mathbf{u})}{\sigma_{\text{obj}}^2 (\text{SNR}_0 \times \text{MTF}^2(\mathbf{u}) + 1)}. \quad (16)$$

$\text{SNR}_0 \triangleq |H(0,0)|^2 \sigma_{\text{obj}}^2 / \sigma_n^2$ is the pixel SNR outside the target area and $0 \leq \text{MTF} \leq 1$ is the modulation transfer function.

Consider task 1, which has contrast energy concentrated at low spatial frequencies. If the product $\text{SNR}_0 \text{MTF}^2$ is large (good quality imaging system), then $\tilde{K}_s(\mathbf{u}) \simeq 1/\sigma_{\text{obj}}^2$ and $\text{ACF}(0) = B/\sigma_{\text{obj}}^4$. In this case, lesion detection is limited by system bandwidth B . Center frequency, bandwidth, beam-forming, and noise conditions are all represented by SNR_0 and $\text{MTF}(\mathbf{u})$ but not in a simple way. Nevertheless (14)-(16) provide a foundation for image quality investigations for task 1. Generalizing the analysis for all tasks and contrast requires application of numerical techniques in (9) to evaluate covariance matrices.

ACKNOWLEDGMENT

This work is supported by NIH under award No. CA118294.

REFERENCES

- [1] H.H. Barrett and K.J. Myers, *Foundations of Image Science*. Hoboken, NJ: John Wiley & Sons, 2004.
- [2] R.F. Wagner and D.G. Brown, "Unified SNR analysis of medical imaging systems," *Phys. Med. Biol.*, vol 30, pp.489-518, 1985.
- [3] S.W. Smith, R.F. Wagner, J.M. Sandrik, and H. Lopez, "Low contrast detectability and contrast/detail analysis in medical ultrasound," *IEEE Trans Son Ultrason*, vol. 30, no. 3, pp. 164-173, May 1983.
- [4] C.K. Abbey, R.J. Zemp, J. Liu, K.K Lindfors, 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, Feb. 2006.
- [5] N.Q. Nguyen, C.K. Abbey, M.F. Insana, "An adaptive filter to approximate the Bayesian strategy for sonographic beamforming," *IEEE Trans Med Imag*, vol. 30, no. 1, pp. 28-37, Jan. 2011.
- [6] S. Kullback, *Information Theory and Statistics, 2/e*. New York, NY: Dover, 1997.
- [7] K. Fukunaga, *Introduction to Statistical Pattern Recognition, 2/e*. New York, NY: Academic Press, 1990.