

Performance Evaluation of a Modular Gamma Camera Using a Detectability Index

John D. Sain, PhD¹; and Harrison H. Barrett, PhD²

¹Optical Sciences Center, University of Arizona, Tucson, Arizona; and ²Department of Radiology and Optical Sciences Center, University of Arizona, Tucson, Arizona

The performance of a modular gamma camera for the task of detecting signals in random noisy backgrounds was evaluated experimentally. The results were compared with a theoretical computer simulation. **Methods:** The camera uses a 10×10 cm thallium-doped sodium iodide crystal, a 2×2 array of 53×53 mm photomultiplier tubes, and a parallel-hole collimator (1.5-mm bore width, 23.6-mm bore length). The camera was positioned to look down into a 10-cm-deep water bath that filled its field of view (FOV). The top surface of the water was 5 cm from the front face of the camera. The camera has 3-mm intrinsic spatial resolution (SR) in the center of its FOV and 9-mm system SR for objects 5 cm below the top surface of the water. Uniform and nonuniform random background data were collected by imaging the bath containing 740 MBq (20 mCi) ^{99m}Tc. Nonuniformities were created by placing water-filled objects in the bath. Each signal dataset was collected by imaging a water-filled plastic sphere, injected with ^{99m}Tc and set at a specific depth (Z) in the bath. Data were collected for many signal diameters (D) (4, 7, 10, 13, 16, 28 mm) at 1 depth (5 cm) and for 1 signal diameter (10 mm) at several depths (1, 3, 5, 7, 9 cm). Sets of signal-present/signal-absent image pairs (380 pairs, 10^5 events per image) for known contrasts (C) were generated for use in ideal-observer studies in which the detectability (d') was calculated. Contrast-detail ($\log C$ vs. $\log D$) plots were created. The theoretical simulation, developed for uniform backgrounds, provided data for comparison. **Results:** The detectability increased linearly with C and decreased nonlinearly with decreasing D or increasing Z . The C required to achieve a specific d' increased sharply for $D < SR$. For $C = 5$, $D = 10$ mm, and $d' = 1.2$, the camera consistently detected signals for $Z < 6$ cm. Similar results were found for nonuniform backgrounds. The theoretical simulation verified the results for uniform backgrounds. **Conclusion:** The methodology presented here provides a way of evaluating gamma cameras on the basis of signal-detection performance for specified lesions, with particular application to scintimammography.

Key Words: detectability; signal detection; gamma camera

J Nucl Med 2003; 44:58–66

The detection of a signal in a noisy background is a classic problem that has been addressed in many areas of study. Successful detection of a signal depends significantly on the ability to understand features of the signal and its background and to select an imaging system that can exploit these features. Unfortunately, however, the probability of successfully detecting a signal is often limited by the capabilities of the imaging system. For example, the imaging system may not be able to detect objects with very low contrast relative to the background, or it may not be able to spatially resolve features of interest. Thus, an important prerequisite for using an imaging system is to understand the limits of its imaging capabilities.

In this article we present a general methodology for evaluating the signal-detection capabilities of gamma cameras. To illustrate the technique, a study was performed of a modular gamma camera for the task of detecting lesions within human breasts. The studies undertaken were not actual clinical studies but rather laboratory simulations in which real-life lesions were represented by small hollow plastic spheres and breast tissue was represented by a realistic physical phantom that mimicked the inhomogeneous uptake of real tissue. Initially, we outline the relevant fundamentals of signal-detection theory and define several metrics with which one may quantify the performance of an imaging system for the task of signal detection. Then we explain the experimental methods and procedures used to measure the camera performance. Finally, we report the results of the performance study with respect to a variety of signal parameters such as signal contrast, size, and depth. A comparison is also made between the experimental data and the output of a theoretical model for the case of uniform backgrounds.

SIGNAL-DETECTION THEORY

Measuring the performance of an imaging system for a particular task is a critical step in ensuring a proper system design. A well-established and rigorous method of measuring performance is signal-detection theory.

Received Mar. 4, 2002; revision accepted Aug. 8, 2002.
For correspondence or reprints contact: John D. Sain, PhD, Lawrence Livermore National Laboratory, Mail Stop L-229, P.O. Box 808, Livermore, CA 94551.
E-mail: sain1@llnl.gov

Background

Signal-detection theory (I) is a form of hypothesis testing in which the hypothesis H_1 is that a signal is present and the alternative hypothesis H_0 is that it is absent. Using the available data, an observer decides, with some degree of error, whether H_1 or H_0 is true. The observer's performance depends on the signal, the background, and the detector system. A more knowledgeable observer, or an observer that has a priori knowledge of one or more of these factors, can be expected to make decisions with a lower rate of error.

Signal-detection theory may be applied in a wide variety of situations that involve the task of determining whether or not a signal is present in a noisy background. For example, in clinical nuclear medical imaging, the signal could be represented by a lesion within a patient's body, the data could be a set of medical images of the patient, the observer could be the attending physician, and the decision could be the physician's diagnosis as to whether the lesion is present or not. The physical shape, histology, and location of the lesion as well as the nature of the surrounding tissue add complexity to the detection task. In this article we simulate the aforementioned example. The signal is a radioactive phantom consisting of an object in a noisy background, the data are a set of planar projections of the spatial distribution of radioactivity within the phantom, the observer is a machine (namely, a computer executing a data analysis program), and the decision as to whether or not the signal is present is based on the calculated value of a test statistic relative to a specified threshold.

Data

In this study the data consist of a set of real sample images, each of which contains a random noise background and a possible signal. Each sample image consists of a set of pixel values $\{g_m, m = 1, \dots, M\}$ that are ordered as an $M \times 1$ vector \mathbf{g} . The noise in the pixel values is Poisson in nature and, hence, dependent on the object. The signal, however, has very low contrast with respect to the background and only slightly perturbs the mean background pixel values, so the noise is assumed to be independent of the signal.

Observer

The observer used in this study is a channelized version of the ideal linear observer. An ideal observer is one who optimally uses all of the information in the data and is therefore one whose performance cannot be surpassed. The ideal linear observer, also known as the Hotelling observer (2–5), is one that is constrained to perform only linear operations on the data but is otherwise optimal in the sense of maximizing a detectability measure. A channelized Hotelling observer is one that performs optimal linear operations on data that have been processed by a relatively small number of linear operators called channels. The channels serve to greatly reduce the dimensionality of the dataset and make it much more feasible to find the optimal linear discriminant.

The original motivation for the channelized Hotelling observer was to mimic the spatial frequency-selective channels in the human visual system and thereby provide a mathematic model observer that accurately predicted human detection performance, even when the human could not perform as well as the true Hotelling observer. We have recently recognized, however, that with proper choice of channels a channelized Hotelling observer can give an accurate estimate of the performance of the ideal linear observer, in spite of the channels ($I, 6$). The key is to make use of prior knowledge of the signal and the background statistics in choosing the channels. In particular, we may know the signal location and know that both the signal and the background correlations are rotationally symmetric, and we can build that knowledge into the channels. Thus, even though channels are used in this article, the goal is to estimate the performance of the ideal linear or Hotelling observer, not a suboptimal linear observer that mimics humans.

To explain the mathematic form of the channelized observer used in this article, we begin with the ordinary Hotelling observer. This observer computes a linear functional of the data, often referred to as the Hotelling test statistic, that accounts for randomness in the signal or the background. This scalar test statistic λ can be written as:

$$\lambda = \sum_{m=1}^M w_m g_m = \mathbf{w}'\mathbf{g}, \quad \text{Eq. 1}$$

where \mathbf{w} , called the Hotelling template, is an $M \times 1$ vector of weights and the superscript t denotes transpose. For signal-known-exactly (SKE) studies in which the noise is considered to be independent of the signal, the Hotelling template is:

$$\mathbf{w} = \mathbf{K}^{-1}(\langle \mathbf{g} \rangle_1 - \langle \mathbf{g} \rangle_0) = \mathbf{K}^{-1}\mathbf{s}, \quad \text{Eq. 2}$$

where \mathbf{K} is an $M \times M$ matrix containing the data covariances, $\langle \mathbf{g} \rangle_1$ and $\langle \mathbf{g} \rangle_0$ are the mean signal-present and signal-absent data vectors, respectively, and \mathbf{s} is an $M \times 1$ vector containing the expected signal. In our model \mathbf{K} includes the effects of both measurement noise (i.e., Poisson noise) and randomness in the background (i.e., anatomic variations).

A channelized observer is often used when the dimensions of the sample images necessitate inverting a very large \mathbf{K} or the number of available sample images is too small to reasonably estimate \mathbf{K} . Suppose that a sample image has dimensions of $N \times N$. Then $M = N^2$ and \mathbf{K} has dimensions $N^2 \times N^2$. The process of forming a sample \mathbf{K} would require a minimum of N^2 sample images to ensure that the matrix is not singular and about 10–100 times N^2 sample images to ensure that the matrix elements are stable. The problem of performing calculations with an $N^2 \times N^2$ matrix and collecting 10–100 times N^2 sample images quickly leads to rather impractical computational situations.

Channels offer a solution to the problems outlined above by providing a means of reducing the size of the sample \mathbf{K} .

In essence, instead of working with the original image pixel values, one works with some functions of the pixel values. By definition, a channel is any template \mathbf{t}_k of the same size as an image. Analogous to linear features in the field of pattern recognition, a channel output is any linear functional of the data \mathbf{g} in the form:

$$g_{c,k} = \mathbf{t}_k^t \mathbf{g} = \sum_{m=1}^M t_{km} g_m, \quad \text{Eq. 3}$$

where $g_{c,k}$ is the output of the k th channel and c denotes “channel.” If K channels are used, then each image \mathbf{g} produces a set of K channel outputs that, when expressed as a $K \times 1$ vector, may be written as:

$$\mathbf{g}_c = \mathbf{T} \mathbf{g}, \quad \text{Eq. 4}$$

where \mathbf{T} is a $K \times M$ matrix with elements t_{km} .

We would like to choose channels such that g_c is a sufficient statistic for the signal-detection task, meaning that the channel outputs will serve as well as the original image pixel values for performing the signal-detection task. To help in this regard, we exploit a priori knowledge of the signal. In this study we know specifically (a) that the signal is rotationally symmetric and (b) the exact location of the signal. So we would like the channels to be a set of functions that are rotationally symmetric and have a known location.

The chosen channels were the 0th–5th orders of the Laguerre–Gauss family of functions centered on the signal location. A Laguerre–Gauss function is a Laguerre polynomial times a gaussian envelope. As required, the Laguerre–Gauss functions possess 2 features—rotational symmetry and a known location—that are also features of the signal. Again, as noted earlier, these channels were chosen not to mimic human performance but to estimate ideal performance. A detailed validation of Laguerre–Gauss channels as an estimate of Hotelling observer performance is provided by Gallas (6). Assuming that the background noise has no preferred orientation and, therefore, has a rotationally symmetric correlation function, we use the individual Laguerre–Gauss functions as rotationally symmetric basis functions to synthesize a Hotelling template that is rotationally symmetric, is centered at the signal location, and has a width approximately equal to that of the signal. The channelized Hotelling template is:

$$\mathbf{w}_c = \mathbf{K}_c^{-1} \mathbf{s}_c, \quad \text{Eq. 5}$$

where \mathbf{K}_c is the $K \times K$ covariance matrix of the channel outputs and \mathbf{s}_c is the signal as seen through the channels. The corresponding channelized Hotelling test statistic is:

$$\lambda_c = \sum_{m=1}^M w_{c,m} g_{c,m} = \mathbf{w}_c^t \mathbf{g}_c, \quad \text{Eq. 6}$$

where, as noted earlier, \mathbf{g}_c is a sample image as seen through the channels.

Thus, given the channelized template \mathbf{w}_c and channelized sample images \mathbf{g}_c , the channelized observer makes a decision by computing the scalar test statistic λ_c and comparing it to a threshold $\lambda_{c,th}$, deciding that H_1 is true if $\lambda_c > \lambda_{c,th}$ or that H_0 is true if $\lambda_c \leq \lambda_{c,th}$.

Figures of Merit

One means of specifying how well an observer can distinguish between H_1 and H_0 is to calculate a signal-to-noise ratio (SNR) defined by:

$$\text{SNR}_{\lambda_c}^2 \equiv \frac{[\langle \lambda_c \rangle_1 - \langle \lambda_c \rangle_0]^2}{[0.5 \text{ var}_1(\lambda_c) + 0.5 \text{ var}_2(\lambda_c)]}, \quad \text{Eq. 7}$$

where $\langle \lambda_c \rangle$ is the conditional expectation of λ_c given that H_k is true and $\text{var}_k(\lambda_c)$ is the corresponding conditional variance. The quantity $\text{SNR}_{\lambda_c}^2$ is also referred to as the detectability index and is denoted by d' . Another means is to measure the area under the receiver-operator-characteristic (ROC) curve (7). If λ_c is normally distributed under both hypotheses, then $\text{SNR}_{\lambda_c}^2$ is related to the area under the curve, denoted as AUC, by:

$$\text{AUC} = 0.5 + 0.5 \text{ erf}(0.5 \text{ SNR}_{\lambda_c}), \quad \text{Eq. 8}$$

where $\text{erf}(\cdot)$ is the error function. Another means of specifying an observer's performance is the detectability index d_a , defined as:

$$d_a \equiv 2 \text{ erf}^{-1}(2 \text{ AUC} - 1), \quad \text{Eq. 9}$$

where $\text{erf}^{-1}(\cdot)$ is the inverse of the error function. The index d_a is equivalent to $\text{SNR}_{\lambda_c}^2$ if λ_c is normally distributed. This article will present the observer performance in terms of the detectability index for several cases in which λ_c is normally distributed.

MATERIALS AND METHODS

In this section we outline the details of a laboratory study designed to measure the performance of a modular gamma camera for the task of detecting signals against a noisy background. Because working with real breast tissue is not a realistic option, the intent of the study was to create an artificial phantom that, when injected with radioactivity and imaged by the camera, would lead to backgrounds similar to those seen in clinical mammoscintigrams (8). Real data obtained from imaging artificial lesions could be added to the backgrounds to create images with signals present.

Hardware Setup

The imaging system used was the stand-alone modular camera (SAMCAM), a single modular gamma-camera system built by the research group at our institution. The camera consists of a $10 \times 10 \times 0.5$ cm thallium-doped sodium iodide, or NaI(Tl), scintillation crystal optically coupled via a $10 \times 10 \times 1.9$ cm fused-quartz light guide to a 2×2 array of 53×53 mm square photomultiplier tubes (PMTs). A lead parallel-hole collimator (1.5-mm bore width, 23.6-mm bore length) is mounted in front of the camera, and the entire assembly is encased in a lead-lined housing. During data collection, the 4 PMT signals for each detected event are digitized to 8 bits and stored in list-mode format for future processing and

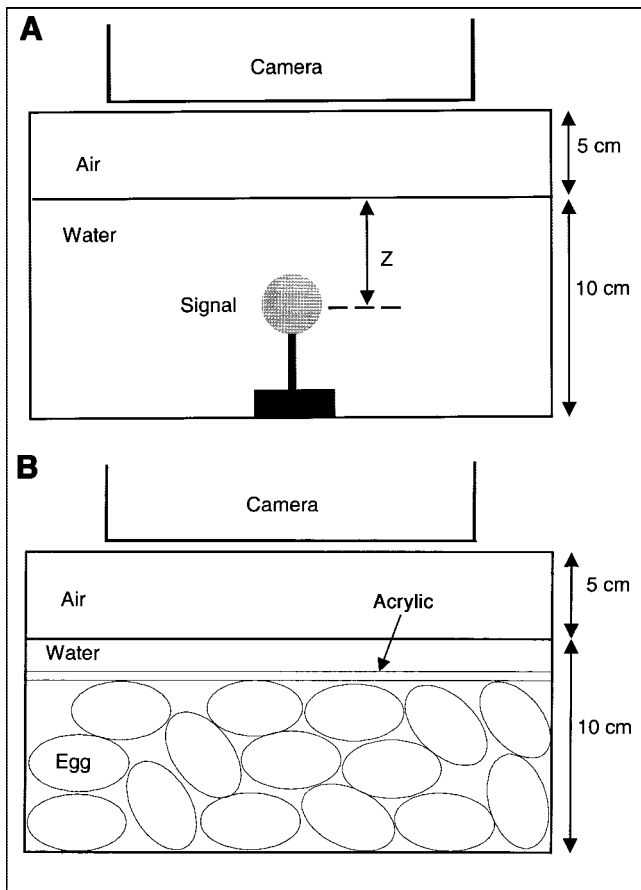


FIGURE 1. (A) Experimental setup for collecting signal data. (B) Experimental setup for collecting nonuniform random background data.

analysis. List-mode format is a data format in which the sets of raw PMT signals for every detected event are written to memory in the order that they were collected. The SAMCAM system was positioned to look down into a 10-cm-deep water bath that filled its field of view (FOV). The top surface of the water was 5 cm from the front face of the camera. The camera has 3-mm intrinsic spatial

resolution (SR) in the center of its FOV and a 9-mm system SR for objects 5 cm below the top surface of the water.

Signal Data

The presence of a lesion within normal fatty breast tissue was simulated by placing a water-filled plastic sphere, injected with ^{99m}Tc , at a specific depth in the water bath as shown in Figure 1A. Each sphere was held in place by a thin plastic rod secured to the bottom surface of the bath container and was located close to the center of the camera's FOV. The goal of each imaging procedure was to collect a high-count signal dataset from which many low-count sample signal datasets could be extracted. During each imaging procedure, the sphere contained about 370 MBq (10 mCi) ^{99m}Tc , whereas the surrounding water had no activity. The sphere diameter (D) and depth (Z) were varied. Six sphere diameters ($D = 4, 7, 10, 13, 16, 28$ mm) were used at 1 depth ($Z = 5$ cm) and 1 sphere diameter ($D = 10$ mm) was used at 5 depths ($Z = 1, 3, 5, 7, 9$ cm). Each signal dataset contained 20×10^6 events in list-mode format. A sample signal image ($D = 10$ mm, $Z = 5$ cm) is shown in Figure 2A. Each signal image was generated by running a set of list-mode data (containing 20×10^6 events) through a maximum-likelihood position-estimation look-up table (ML-PE LUT) (9,10). Nonuniformities in the camera response were removed with a high-count (8-million-event) reference flood image. An image consists of a 64×64 array of 1.5625×1.5625 mm pixels.

Background Data

Uniform Background Images. To generate uniform background images, 740 MBq (20 mCi) ^{99m}Tc were injected into the water bath, and then the water was stirred to ensure a uniform distribution of the radiotracer. A single dataset, containing 76×10^6 events, was collected and stored in list-mode format. A set of 760 images, consisting of independent realizations of uniform noise, was generated by running unique blocks of 10^5 events from the list-mode data through the ML-PE LUT.

Nonuniform Random Background Images. The spatial distribution of radioactivity present within normal fatty breast tissue was simulated by imaging a bath of radioactive water containing randomly located nonradioactive objects. Fifty hollow plastic eggs (46×32 mm), filled with nonradioactive water, were placed in the bath and held 2.5 cm underwater by a sheet of 0.5-cm-thick acrylic

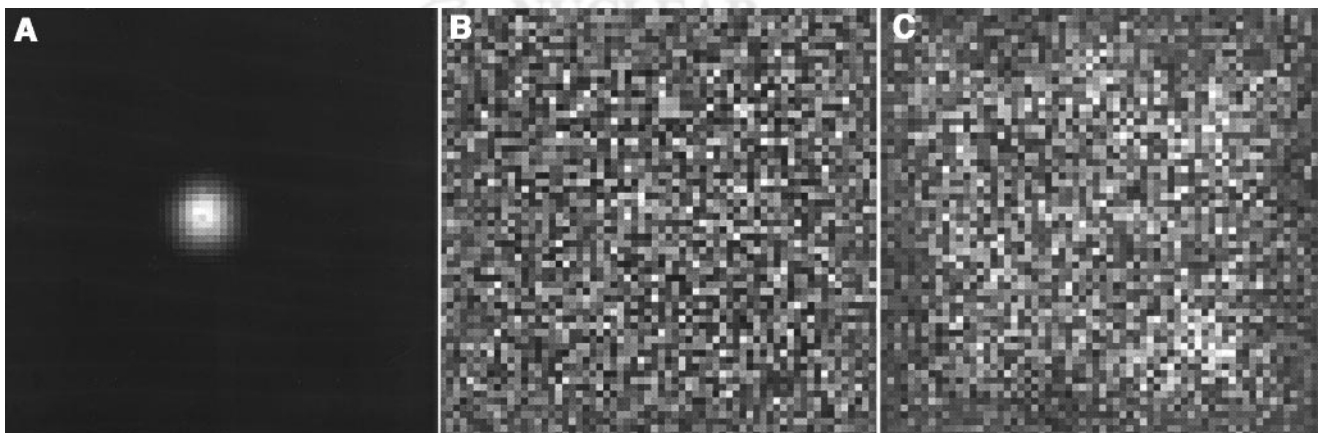


FIGURE 2. (A) Sample signal image ($D = 10$ mm, $Z = 5$ cm). (B) Sample uniform background image with signal present. (C) Sample nonuniform random background image with signal present.

during imaging as shown in Figure 1B. After 740 MBq (20 mCi) ^{99m}Tc were injected into the bath, the water and eggs were stirred to ensure that the radiotracer had spread uniformly throughout the water outside of the eggs. Seven hundred sixty datasets, each containing 10^5 events, were collected and stored in separate list-mode data files. Before collection of each dataset, the eggs were stirred so that their positions were unique and random. A series of 760 images, consisting of independent realizations of nonuniformly random noise, was generated by running each list-mode data file, containing 10^5 unique events, through the ML-PE LUT.

Average Radial Spectrum

The primary goal in generating the background images was to create realistic-looking nonuniform random backgrounds that had spatial frequency content similar to that of real clinical mammoscintigrams collected with the SAMCAM system. The spatial frequency content of an image was determined by calculating the angular average of the squared modulus of the discrete Fourier transform of the image. After the averaging, the remaining radial function is referred to as the average radial spectrum (ARS).

In general, the spatial frequency content of an image can be described by the normalized squared-modulus of the discrete Fourier transform of the image. Specifically, let a 2-dimensional, discrete image be denoted by $f_D(n_x, n_y)$, where the subscript D denotes discrete and n_x and n_y are the image pixel indices. The normalized spectrum $S_D(k_x, k_y)$ of $f_D(n_x, n_y)$ is defined as:

$$S_D(k_x, k_y) = \frac{|F_D(k_x, k_y)|^2}{|F_D(0, 0)|^2}, \quad \text{Eq. 10}$$

where k_x and k_y are the discrete spatial frequencies in the x and y directions, respectively. We specifically avoid the use of the term “power spectrum” in this definition because (a) the images under consideration are not stationary stochastic processes and (b) the total number of images is small and finite—not approaching infinity as assumed by the Wiener–Khinchin theorem. To compute the ARS, we define:

$$a_i(k_x, k_y) = \begin{cases} 1 & \text{for } \rho_i \leq \rho < (\rho_i + 1), \quad i = 0, 1, 2, 3, \dots \\ 0 & \text{elsewhere} \end{cases} \quad \text{Eq. 11}$$

$$b_i(k_x, k_y) = \begin{cases} S_D(k_x, k_y) & \text{for } \rho_i \leq \rho < (\rho_i + 1), \quad i = 0, 1, 2, 3, \dots \\ 0 & \text{elsewhere} \end{cases} \quad \text{Eq. 12}$$

and

$$A_i = \sum_{k_x, k_y} a_i(k_x, k_y). \quad \text{Eq. 13}$$

$$B_i = \sum_{k_x, k_y} b_i(k_x, k_y). \quad \text{Eq. 14}$$

For a given integral radial spatial frequency, represented by the integer i , the term A_i is the number of pixels in $S_D(k_x, k_y)$ that correspond to $\rho \in [\rho_i, \rho + 1)$, and the term B_i is the sum of the corresponding pixel values. Thus, the ARS may be expressed as:

$$\bar{S}_D(\rho) = \frac{B_i}{A_i}, \quad i = 0, 1, 2, 3, \dots, \quad \text{Eq. 15}$$

where, as i increases, the number of pixels over which the average is performed also increases. The mean ARS is obtained by averaging the ARS values for a large number (e.g., 100) of sample images.

In our analysis of clinical mammoscintigrams, we extracted square subregions, 32 pixels on a side, from the central regions of the images. The subregions were used to ensure that the spectral analysis was performed only on regions containing projections of breast tissue. We observed that log–log plots of ARS versus ρ were approximately linear with a negative slope. The mean ARS for the mammoscintigrams had a slope of approximately -1.2 . Experimentation in the laboratory with various object arrangements showed that the specified arrangement of eggs worked well when a layer of attenuating material—water, in this case—was placed between the eggs and the camera as shown in Figure 1B. The attenuation provided by the water acted to partially smooth out the nonuniformities in the spatial distribution of radioactivity created by the nonradioactive eggs. The corresponding slope of the mean ARS versus ρ plot for the sample nonuniform background images varied as a function of the thickness of the water layer, and we found that a 2.5-cm-thick layer of water above the eggs provided a slope of approximately -1.2 , quite close to that of the mammoscintigrams (8).

Contrast

The 3-dimensional contrast was calculated using the source activity per unit volume (A_s) and background activity per unit volume (A_b). The contrast can be expressed in terms of the tumor-to-background ratio (TBR) as:

$$C = \frac{A_s - A_b}{A_b} = \frac{A_s}{A_b} - 1 = \text{TBR} - 1. \quad \text{Eq. 16}$$

Image Pair Generation

For both uniform and nonuniform random backgrounds and for each unique combination of signal contrast, diameter, and depth, a set of 380 image pairs was generated for use in detectability studies. A single image pair consisted of a signal-present image (an image containing signal data) and a signal-absent image (an image containing no signal data). Signal-absent images were generated by running 10^5 background events through the ML-PE LUT. Signal-present images were similarly generated by running 10^5 events through the ML-PE LUT, but both signal and background events were used. A unique set of background and, if necessary, signal event data were used in every single image. For each combination of signal diameter and depth, the number of signal events added to a background image was determined by the desired contrast. Recall that C is related to the TBR by the expression $C = \text{TBR} - 1$. Because signal and background datasets were collected independently of each other, the amount of activity per unit volume of water in each signal dataset was scaled with respect to that of the background dataset to simulate a desired TBR. The range of TBR used in each set of image pairs ranged from 1.1 to 20.0. Sample signal-present images for both the uniform and the nonuniform random cases are shown in Figures 2B and 2C, respectively. The signal parameters in these images are $D = 13$ mm, $Z = 5$ cm, and $C = 5$.

Laguerre–Gauss Functions

The Laguerre–Gauss functions used to synthesize the channelized Hotelling template \mathbf{w}_c were Laguerre polynomials multiplied by a gaussian envelope. On the basis of the results of Gallas (6),

we concluded that 6 polynomials would be sufficient to accurately represent the true Hotelling template. We defined the expansion for the template, expressed in continuous radial coordinates with the origin at the signal location, to be:

$$w(r) = \sum_n \alpha_n \exp\left(-\frac{\pi r^2}{a^2}\right) L_n\left(\frac{2\pi r^2}{a^2}\right), \quad n \in [0, 5], \quad \text{Eq. 17}$$

where r is the radial coordinate, n is the order number, α_n is the order coefficient, and a is the parameter controlling the full width at half maximum (FWHM) of the 0th-order function. The final digital template was obtained by sampling $w(r)$ on the pixel grid. The values of the parameter a were chosen to optimize the detectability for each sphere diameter. Specifically, the values of a used for the spheres of diameter 4, 7, 10, 13, 16, and 28 mm at a depth of 5 cm were 10.0, 10.8, 10.9, 13.0, 14.1, and 21.6 mm, respectively. Similarly, the values of a used for the sphere of diameter 10 mm at depths of 1, 3, 5, 7, and 9 cm were 9.8, 10.6, 11.7, 13.0, and 14.7 mm, respectively.

Theoretical Model of Laboratory Experiment

A theoretical computer model was developed to simulate the laboratory experiment for the case of a uniform background. A representative diagram is shown in Figure 3. The camera was modeled by a square (10 cm per side), pixellated (64×64) detector in the xy -plane. The water bath was represented by a cube (10 cm per side) of uniform background activity per unit volume $A_b(x,y,z)$. The signal was modeled by a sphere of diameter D and uniform activity per unit volume $A_s(x,y,z)$ and was located such that its center was above the center of the detector and at a depth of Z below the surface of the water. The attenuation factor for both the signal and the background was assumed to be that of water, or 0.15 cm^{-1} .

The activity within the modeled water bath was projected, along the z -axis, onto the detector. Each projection was considered to be

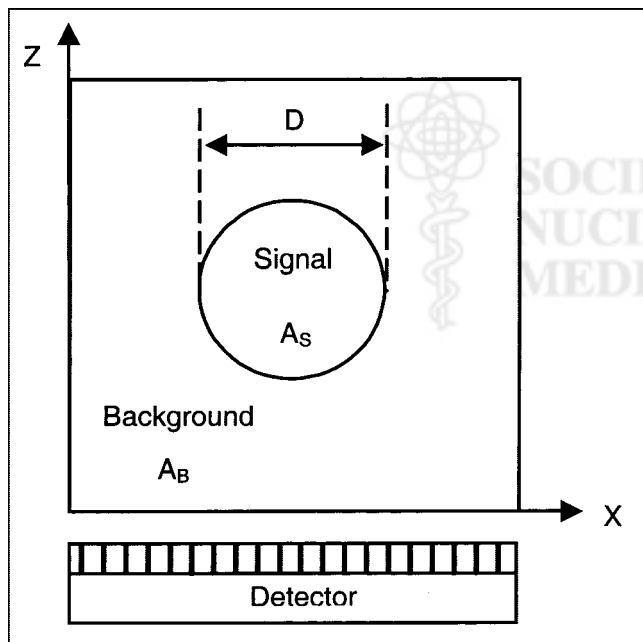


FIGURE 3. Diagram of theoretical model for uniform background experiment.

the mean of a very large number of sample noisy projections, so count-dependent Poisson noise in the individual pixel values is assumed to have averaged to zero. Specifically, we defined the projected activity per unit area, denoted by $P(x,y)$, to be:

$$P(x,y) = \int_0^L [A_s(x,y,z) + A_b(x,y,z)] dz, \quad \text{Eq. 18}$$

the projected signal activity per unit area, denoted by $P_s(x,y)$, to be:

$$P_s(x,y) = \int_0^L A_s(x,y,z) dz, \quad \text{Eq. 19}$$

and the projected background activity per unit area, denoted by $P_b(x,y)$, to be:

$$P_b \equiv P_b(x,y) = \int_0^L A_b(x,y,z) dz, \quad \text{Eq. 20}$$

where we have noted that $P_b(x,y)$ is constant. Scattering within the imaging process and blurring within the projection images due to the intrinsic spatial resolution of the imaging system were both simulated by convolving $P(x,y)$ with a gaussian kernel having a FWHM equal to 9 mm, the system SR for objects 10 cm from the camera face and, simultaneously, 5 cm deep in water. The blurred functions of the projected activities per unit area were designated with a prime, or by $P'(x,y)$ and $P'_s(x,y)$. The detectability, d' , was calculated from the 2-dimensional projection data as:

$$d' = \sqrt{\frac{\sum_{x,y} P_s'^2(x,y)}{P_b}} = \sqrt{\frac{\sum_{x,y} [P'(x,y) - P_b]^2}{P_b}}. \quad \text{Eq. 21}$$

These expressions for d' are valid for the Hotelling observer in SKE studies with uniform backgrounds. In this model, $P_b = 1.48 \text{ kBq/mL}$ ($0.04 \text{ } \mu\text{Ci/mL}$), $Z = 5.0 \text{ cm}$, $A_s \in [1.48, 88.8] \text{ MBq/mL}$ ($[0.04, 2.40] \text{ mCi/mL}$), and $D \in [3, 30] \text{ mm}$.

RESULTS

The performance of the modular gamma camera was evaluated for the task of detecting the presence of an artificial lesion within noisy backgrounds. The results include both experimental and theoretical data. The experimental performance data from the laboratory include comparisons of signal detectability versus signal contrast, signal contrast versus signal diameter, and signal contrast versus signal depth for the cases of uniform and nonuniform random backgrounds. The theoretical performance data from the computer model included a comparison of signal contrast versus signal diameter for the case of uniform backgrounds.

Laboratory Performance Data

Detectability Versus Contrast. The d' increased linearly with respect to C as shown in Figure 4. The C required to achieve a specified d' was greater for the nonuniform random background than for the uniform background. The magnitudes of the slopes of the uniform and nonuniform plots are 0.37 and 0.20, respectively, indicating that the

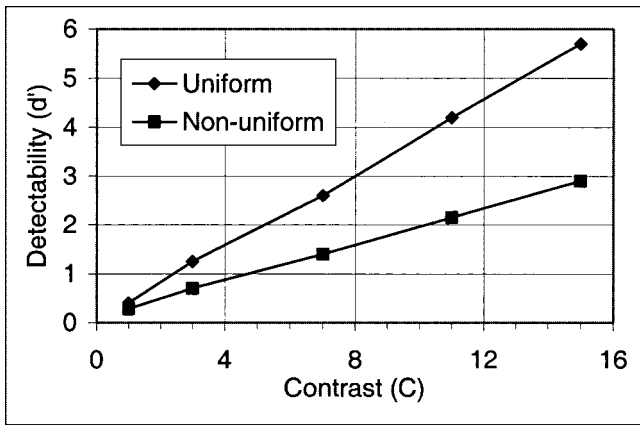


FIGURE 4. Detectability (d') vs. contrast (C) plots for uniform and nonuniform random background experiments ($D = 10$ mm, $Z = 5$ cm).

source activity in the nonuniform case must be approximately 1.85 times greater than the activity in the uniform case to achieve the same d' .

Contrast Versus Diameter. The C required to maintain a constant d' increased nonlinearly with decreasing D as shown in Figure 5. The rate of increase in C was significantly larger if D was less than the system SR (9 mm). For the uniform case shown in Figure 5A, the average slope of the contrast-detail plots increased from -2.5 for larger

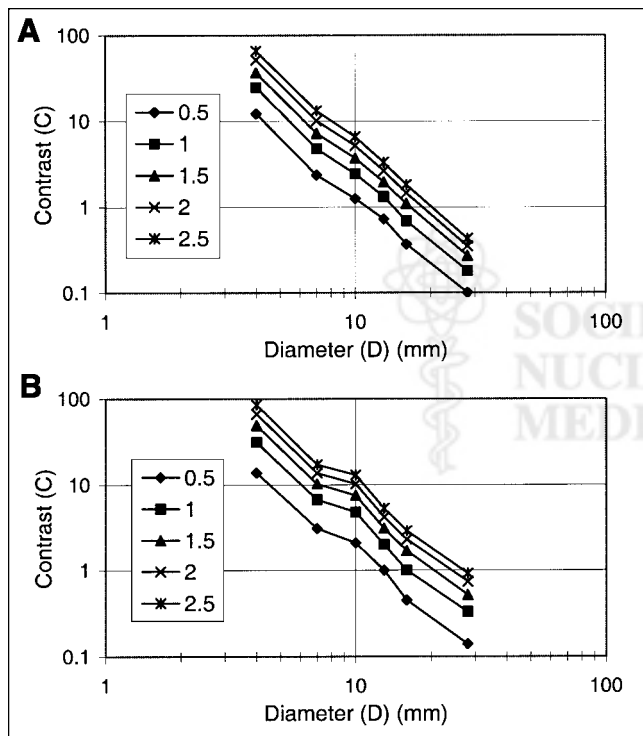


FIGURE 5. (A) Contrast (C) vs. diameter (D) plot for uniform background experiments as function of detectability (d'). (B) Contrast (C) vs. diameter (D) plot for nonuniform random background experiments as function of detectability (d').

diameter signals to -2.8 for smaller diameter signals. Similarly, for the nonuniform case shown in Figure 5B, the average slope increased from -2.0 to -2.9 . The amount of change in slope was not significant, and the limited number of data points ($n = 6$) may be masking some of the true behavior of the relationship. For $d' = 1.2$ (or $AUC \approx 0.8$), the average increase in C for all D , from the uniform to the nonuniform case, was 37%. The largest rise was 76% for $D = 10$ mm.

Contrast Versus Depth. The C required to maintain a constant d' increased nonlinearly with increasing Z as shown in Figure 6. The rate of increase in C was larger for larger Z . For $d' = 1.2$, the average increase in C for all Z , from the uniform case shown in Figure 6A to the nonuniform case shown in Figure 6B, was 38%. The percentage of increase rose as Z increased—from 19% for $Z = 1$ cm to 60% for $Z = 9$ cm. Thus, as Z increased, higher levels of activity had to be present to be detected with the same level of performance.

Theoretical Performance Data

The theoretical model output and the experimental data agree. A contrast-detail plot for a detectability of 2.0 is shown in Figure 7. The contrast required to achieve a detectability of 2.0 increases sharply for signal diameters of <9 mm, the system SR for objects in water at a depth of 5 cm.

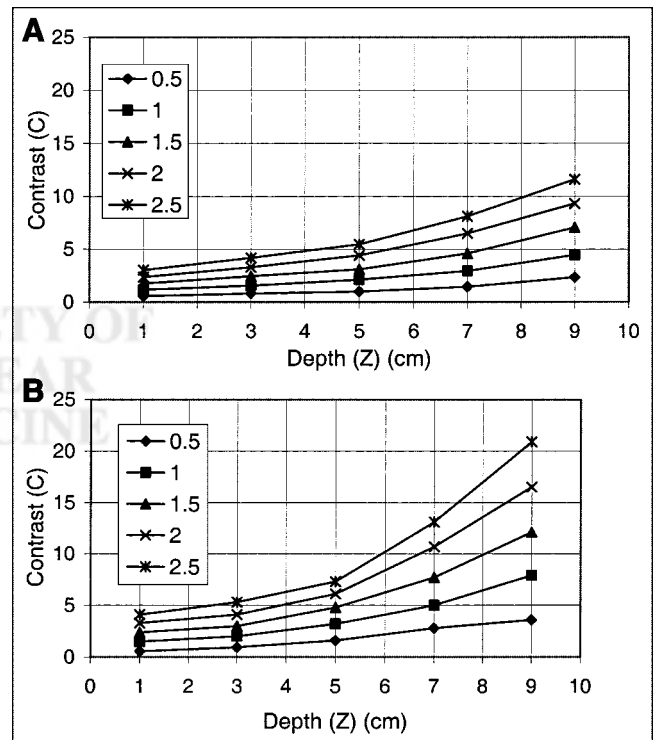


FIGURE 6. (A) Contrast (C) vs. depth (Z) plot for uniform background experiments as function of detectability (d'). (B) Contrast (C) vs. depth (Z) plot for nonuniform random background experiments as function of detectability (d').

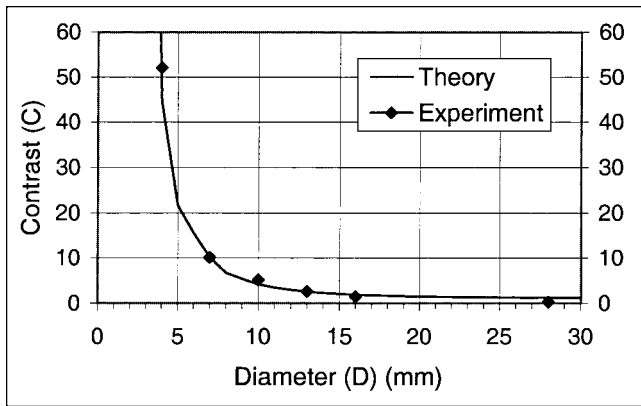


FIGURE 7. Comparison of theoretical and experimental contrast-detail plots for uniform background experiment ($d' = 2.0$).

DISCUSSION

In this section we discuss (a) the rationale for our choice of parameters for the signal and background phantoms and (b) both the origin of the contrast-detail diagram and the significant details of the contrast-detail diagrams presented in this study.

The signal and background phantoms were designed to represent—in a simple manner—what the camera might see in a clinical imaging situation. For signal contrast, we knew that different types of tissue exhibit varying degrees of radiotracer uptake, and the degree of uptake can be affected by different characteristics of the tissue. For the case of breast lesions, we noted through a literature survey that ductal and lobular carcinoma consistently exhibited abnormal tissue growth and relatively high uptakes of ^{99m}Tc -sestamibi, whereas normal and fatty tissues exhibited relatively low uptakes. Indeed, ductal and lobular carcinoma located in normal or fatty breast tissue exhibited relative uptake ratios, or TBRs, of nearly 6:1 with respect to the surrounding tissue (11). Some other types of tissues also exhibited abnormal cell growth and relatively high uptakes of ^{99m}Tc -sestamibi but did not do so on a consistent basis. Thus, C was allowed to range from 1 to >12 —more than twice the observed value of about 5:1 for ductal and lobular carcinomas. For signal size, D ranged from 28 mm down to 4 mm in an effort to test the camera's capabilities as D passed below 10 mm, a size below which some nuclear medical imaging systems begin to experience an increased loss in sensitivity for detection. For signal depth, a Z of up to 9 cm was allowed in anticipation that most breasts—whether pendant or compressed—would be <10 -cm thick along the camera's line of sight. We acknowledge that real breast lesions seldom possess perfectly spherical shapes or contain perfectly uniform radiotracer distributions, but these experiments do provide an initial estimate of how well the modular gamma camera might reliably detect the presence of lesions within normal fatty breast tissue.

The ability of an imaging system to detect a signal depends, in part, on the signal contrast, signal diameter, and

exposure quanta (or time). For a specified level of detection performance, a relationship between these 3 parameters can, in theory, be derived. The theoretical relationship for an ideal observer has been formulated and used extensively in the analysis of imaging systems by Wagner and Brown (12,13). These authors illustrated this relationship in the form of a contrast-detail diagram, or a log-log plot of C versus D . They observed that, in general, (a) the plot had a negative slope and was essentially linear for the asymptotic cases of $D \rightarrow 0$ and $D \rightarrow \infty$, (b) the slope magnitude was greater for the asymptotic case of $D \rightarrow 0$, and (c) the slope magnitude underwent a smooth and gradual transition from the greater to the lesser value near the value of D that matched SR. Wagner and Brown reasoned that as D decreased, the corresponding rate of increase in C —on a log-log scale—required to maintain a specified d' would increase once $D < \text{SR}$. They noted that the slope magnitude for each part of the plot depends, in part, on the type of imaging system being used. They also cautioned that although the existence of such relationships—essentially power laws—for imaging systems is well known, the slope value is often incorrect or at least oversimplified. In this study we plotted contrast-detail diagrams for $d' \in [0.5, 1.0, 1.5, 2.0, 2.5]$ for both uniform and nonuniform random backgrounds, as shown in Figures 5A and 5B, respectively. We did not observe a significant change in the slope near a value of $D = \text{SR} = 9$ mm. Several potential reasons exist to explain the apparent lack of change in slope. First, each plot contains only 6 data points, and the data points span only 1 order of magnitude of D —3 mm to 30 mm. The number of data points is too small to make reasonable estimates of the slopes for $D \ll \text{SR}$ and $D \gg \text{SR}$. A larger number of data points, spread out over a wider range of D , is required to ensure reasonable estimates of slope. Second, the physical dimensions of the nonuniformities introduced into the image backgrounds by the plastic eggs were similar in size to D when $D \approx 9$ mm. In the nonuniform random background case, shown in Figure 6B, the imaging system clearly required a higher C to detect the signal when $D \approx 9$ mm. In essence, as seen by the imaging system, the signal may have appeared to be one of many random nonuniformities present in the background. Third, we must acknowledge the possibility that the contrast-detail diagram for this particular imaging system simply may not exhibit an observable change in slope as suggested by Wagner and Brown. Considering all 3 potential reasons, a more comprehensive set of laboratory experiments could generate more informative contrast-detail diagrams.

CONCLUSION

We have presented an experimental methodology for evaluating gamma cameras on signal-detection tasks. The approach allows consideration of realistic random backgrounds as well as Poisson noise.

There are 2 novel features of the methodology. The first is that it is essentially experimental. By using plastic models for backgrounds and signals and measured image samples, it avoids theoretical and potentially oversimplified models for the object and imaging system. The drawback to this approach is that an impossibly large number of sample images would be required for direct construction of the Hotelling observer. We avoid this problem with the second novel feature, the use of a channelized Hotelling observer with the channels chosen in such a way that their outputs form an approximate sufficient statistic for the detection problem. In particular, radially symmetric Laguerre–Gauss channels centered on the known signal location are able to capture the same information with respect to the detection task as the actual Hotelling template. Thus, we are computing the performance of the full Hotelling observer, not a suboptimal model that might account for limitations of the human observer, yet we can do so with a realistic number of sample images.

A detailed validation of the use of Laguerre–Gauss channels in this way was performed by Gallas (6). He considered a wide range of lumpy backgrounds and studied the effects of varying the number of channels and their width. The general conclusion from this work is that the channelized-observer detectability asymptotically approaches the true Hotelling detectability as the number of channels increases, but 5–7 channels are sufficient except in a few extreme cases not representative of the kinds of background used in this study. Gallas also found that there was a broad maximum in detectability as a function of the width parameter for a fixed number of channels. Neither the width nor the number of channels was at all critical. On the basis of this work, we chose to use 6 channels and to vary the width for maximum detectability. The drawback to using >6 channels is that then sampling error from the finite number of image samples increases. This topic was also treated in detail (6) by means of both Monte Carlo simulation and bootstrap methods.

We used this methodology to assess the performance of our modular gamma cameras for the task of detecting small low-contrast signals in uniform and nonuniform backgrounds. For uniform backgrounds, we found that detectability increases linearly with signal contrast and nonlinearly with signal diameter and that it decreases nonlinearly with increasing signal depth. An average increase in contrast of about 38% was required to maintain a specified detectability when changing from uniform to nonuniform random backgrounds. In the case of nonuniform random backgrounds, for a signal contrast of 5, the camera was able to reliably detect (a) signals with diameters of >10 mm at a depth of 5 cm and (b) signals with 10-mm diameter at depths of <6 cm.

Though these results apply strictly only to the ideal Hotelling observer, they should be indicative of the performance of human observers. In earlier work we have found that human and Hotelling observers correlate fairly well (3,5,14–16) if there are no high-pass noise correlations of the sort introduced by tomographic image reconstruction. For example, the Hotelling observer has been found to give the correct rank ordering of collimators for planar liver imaging (14) and to predict correctly the optimal pinhole size for the best human detection in lumpy-background models similar to those studied here (5). More quantitative predictions of human performance, especially in tomographic noise, require anthropomorphic model observers with channels similar to those in the visual system and observer internal noise (16). An advantage of the methodology presented in this study is that it can be adapted to both of these situations—estimating ideal performance and predicting human performance—simply by proper choice of channels.

ACKNOWLEDGMENT

This study was supported by U.S. Army grant DAMD17-97-1-7190 and National Institutes of Health grants P41RR14304 and RO1CA52643.

REFERENCES

1. Barrett H, Abbey C, Gallas B, Eckstein M. Stabilized estimates of Hotelling observer detection performance in patient-structured noise. *Proc SPIE*. 1998; 3340:27–43.
2. Smith W, Barrett H. Hotelling trace criterion as a figure of merit for the optimization of imaging systems. *J Opt Soc Am A*. 1986;3:717–725.
3. Fiete R, Barrett H, Smith W, Myers K. The Hotelling trace criterion and its correlation with human observer performance. *J Opt Soc Am A*. 1987;4:945–953.
4. Myers K, Rolland J, Barrett H, Wagner R. Aperture optimization for emission imaging: effect of a spatially varying background. *J Opt Soc Am A*. 1990;7:1279–1293.
5. Rolland J, Barrett H. Effect of random background inhomogeneity on observer detection performance. *J Opt Soc Am A*. 1992;9:649–658.
6. Gallas B. *Signal Detection in Lumpy Backgrounds* [PhD dissertation]. Tucson, AZ: University of Arizona; 2001.
7. Metz C. Basic principles of ROC analysis. *Semin Nucl Med*. 1978;8:293–298.
8. Sain J. *Optical Modeling, Design Optimization, and Performance Analysis of a Modular Gamma Camera for the Detection of Breast Cancer* [PhD dissertation]. Tucson, AZ: University of Arizona; 2001.
9. Aarsvold JN, Barrett HH, Chen J, et al. Modular scintillation cameras: a progress report. *Proc SPIE*. 1988;914:319–325.
10. Milster TD, Aarsvold JN, Barrett HH, et al. A full-field modular gamma camera. *J Nucl Med*. 1990;31:632–639.
11. Maublant J, de Latour M, Mestas D, et al. Technetium-99m sestamibi uptake in breast tumor and associated lymph nodes. *J Nucl Med*. 1996;37:922–925.
12. Wagner R, Brown D. Overview of a unified SNR analysis of medical imaging systems. *IEEE Trans Med Imaging*. 1982;MI-1:210–214.
13. Wagner R, Brown D. Unified SNR analysis of medical imaging systems. *Phys Med Biol*. 1985;30:489–518.
14. Barrett HH, Gooley TA, Girodias KA, Rolland JP, White TA, Yao J. Linear discriminants and image quality. *Image Vis Comput*. 1992;10:451–460.
15. Barrett HH, Yao J, Rolland JP, Myers KJ. Model observers for assessment of image quality. *Proc Natl Acad Sci USA*. 1993;90:9758–9765.
16. Abbey CK, Barrett HH. Human- and model-observer performance in ramp-spectrum noise: effects of regularization and object variability. *J Opt Soc Am A*. 2001;18:473–488.