Task-based lens design with application to digital mammography

Liying Chen and
Optical Sciences Center, University of Arizona, Tucson, Arizona 85721

Harrison H. Barrett
Optical Sciences Center, University of Arizona, Tucson, Arizona 85721, and Department of Radiology, University of Arizona, Tucson, Arizona 85724

Abstract

Recent advances in model observers that predict human perceptual performance now make it possible to optimize medical imaging systems for human task performance. We illustrate the procedure by considering the design of a lens for use in an optically coupled digital mammography system. The channelized Hotelling observer is used to model human performance, and the channels chosen are differences of Gaussians. The task performed by the model observer is detection of a lesion at a random but known location in a clustered lumpy background mimicking breast tissue. The entire system is simulated with a Monte Carlo application according to physics principles, and the main system component under study is the imaging lens that couples a fluorescent screen to a CCD detector. The signal-to-noise ratio (SNR) of the channelized Hotelling observer is used to quantify this detectability of the simulated lesion (signal) on the simulated mammographic background. Plots of channelized Hotelling SNR versus signal location for various lens apertures, various working distances, and various focusing places are presented. These plots thus illustrate the trade-off between coupling efficiency and blur in a task-based manner. In this way, the channelized Hotelling SNR is used as a merit function for lens design.

1. INTRODUCTION

The systematic method of lens design in use today is an iterative technique. By this procedure, a basic lens type is first chosen. A paraxial thin-lens and then a thick-lens solution are developed next. Then, with the great progress in computers and applied mathematics today, one can directly go to optimization by relatively high-speed ray trace. Any automatic computer optimization program actually drives an optical design to a local minimum, as defined by a merit function. Broadly speaking, the merit function can be described as a function of calculated characteristics, which is intended to completely describe the value or the quality of a given lens design. The typical merit function is the weighted sum of the squares of many image defects, among which are various aberration coefficients, spot sizes, and modulation transfer function at one or more locations in the field of view. Some merit functions consider the boundary conditions for the lens as well. These boundary conditions include items such as maintaining the effective focal length, center and edge spacings, and overall length. Choice of terms to be included in the merit function is a critical part of the design process. The basis for selecting the aberrations is the experience of designers and their understanding of the nature of lenses and aberrations. The exact choice of term weights and balancing aberrations is largely an empirical process. A large amount of effort has been expended to improve the optimization procedure, but the choice of merit function to provide the best image quality is still not so clear.

Corresponding author Liying Chen's e-mail address is lichen@u.arizona.edu..

Although the aberration coefficients give designers an idea of the performance of the system, the overall effects of these aberrations on the final image is hard to visualize. The question is, what effect does a given amount of aberration have on the performance of optical imaging systems? Phrased in this way, the question is rather vague, and more concrete definitions of image quality are needed. A scientific or medical image is always produced for some specific purpose or task, and the only meaningful measure of its quality is the appropriate observer's performance on that task. The basic difficulty with the weighted sums of the squared errors as measures of the image quality is that they do not take into account the observer and the task of images. A task-based approach to assessment of image quality must therefore start with a specification of the task followed by a quantitative determination of how well the observer performs the task. Generally, the tasks of practical importance can be either to classify the image into one of two or more categories or to estimate one or more quantitative parameters of the object from the image data. We call these two kinds of tasks classification and estimation, respectively. In astronomical imaging, for example, one might want to detect the presence of a small companion star near a brighter star, or one might want to estimate the relative magnitude of the two stars. The first problem is a classification task, and the second is an estimation task. The observer can be either a human or some mathematical model observer like a machine classifier or pattern recognition system. With the figure of merit for each task, we can quantitatively measure the task-based image quality of imaging systems.

One challenging lens design area is medical x-ray imaging systems, which are usually designed to produce images having the highest quality possible for diagnosis within a specified dose. One type of digital x-ray imaging system uses optics to couple x-ray fluorescent screens onto detection devices or image receptors such as CCD arrays. The lens design must meet a very stringent requirement because the radiation dose applied to patients is limited. Within the safe dose level, x-ray imaging systems should be quantum limited, which means that the dominant noise is the quantum noise from the fluctuation in the x-ray photons. Although x-ray phosphor screens may create hundreds of visible or UV photons per x-ray photon absorbed, the imaging optics cannot couple all of them onto an image receptor. In fact, the number of transmitted light quanta is so much less than the number of those created that the imaging optics can easily become the quantum sink in the whole x-ray imaging system. In its simplest interpretation, when we present an imaging system as a series of cascaded processes or stages, the stage with the fewest quanta per incident x-ray photon is called the quantum sink. The cascaded model is one form of analysis often used by system designers. In this model, the quanta leaving one stage contribute an effective input to the subsequent stage. The starting process usually results from the Poisson statistics characterizing the incident x-ray photons. The randomness introduced by Poisson statistics is a basic noise in photon-limited imaging systems, but there always exists additive noise from external sources, such as electronic components. When imaging systems are designed or assessed, it is important to understand the processes that contribute to the image formation and the effects they have on image signal and noise. In the cascade model, the noise propagates through all stages to the final image as well. All sorts of blurring, especially from aberrations, introduce changes in the appearance and amplitude of the statistical x-ray fluctuations. Such noise takes on a characteristic mottled appearance. Only fast lenses can couple enough light onto image receptors to make recognizable contrast and satisfactory signal intensity over noise. At the same time, the inherent severe aberration of fast lenses can make the mottled appearance even worse.

Recent advances in model observers that predict human perceptual performance now make it possible to optimize optical imaging systems, such as lenses, for human task performance. We illustrate the procedure by investigating a design of a lens for use in an optically coupled digital mammographic imaging system. The cascaded imaging chain consists of a fluorescent screen, an imaging lens, and an electronic imaging device such as a CCD camera. Observers' performance is evaluated based on the statistical properties of the images from the system. The

statistics of images are determined by the statistics of objects and the statistics of the noise from imaging systems. A detailed model of the system to be optimized is described in Section 2. The general observer's performance and the specific analysis of this model are introduced in Section 3. Section 4 specifies the Monte Carlo strategy used for the numerical computation in evaluating our objective merit function for lens design. Some case studies and discussion are provided in Section 5, and conclusions are presented in Section 6.

2. SYSTEM MODEL

An optically coupled digital mammographic imaging system usually includes an x-ray source, a fluorescent screen, an imaging optical component, and a digital image receiver. A layout drawing of the model system is shown in Fig. 1. Optimizing the imaging optics in this context requires the knowledge of the full system so that the object and the noise propagation can be precisely decided. In Subsections 2.A and 2.B, each system component is investigated carefully and different statistical aspects are given as completely as possible.

A. Physical System

A remote x-ray generator is placed in front of a fluorescent screen, and the patients to be imaged are in between, usually close to the screen. The detection unit after the screen consists of an optical imaging lens and a CCD camera. The lens images the exit surface of the screen onto the detector plane of the CCD camera.

The most common type of x-ray source in diagnostic radiology is a vacuum tube in which high-energy electrons bombard a metal anode and create x-ray photons. The x-ray photons produced during the exposure are independent of each other, and the total number is Poisson distributed.[1] After the patient, the transmitted x-ray photons still keep their independence since their interactions with tissue are unrelated to each other. The x-ray source is usually far from the fluorescent screen, and there is an antiscattering grid after the patient, so the directions of the x-ray photons are approximately parallel to each other after the patient.

X-ray fluorescent screens are usually made of rare-earth-doped crystallites embedded in plastic binders. These tiny crystal grains emit greenish light upon x-ray excitation. The production rate of light photons is fairly high, over a thousand visible photons per absorbed diagnostic x-ray photon on average. Right after generation, visible light beams are scattered inside the screen by these same crystal grains and meander out to the exit surface of the screen. Once a visible light photon reaches the exit surface, it is most likely displaced from the original position of the absorbed x-ray photon on the surface, and it is also deviated from the original direction of the x-ray photon. The current screen types usually have a reflector on their incident sides to direct more visible light outward. However, the same reflector increases the scattering spread. The detailed specifications of the screens used in the model are based on Kodak products.

An imaging lens couples the exit surface of the screen onto the front surface of a digital imaging detector. To accommodate the whole screen area onto the detector, this lens has to work in a small absolute magnification. The consequently small numerical aperture in the object space cuts off so much visible light emitted from the screen that the ultimate x-ray-to-light-photon conversion rate at the detector stage is around 1, even though 1000 light photons are produced by one x-ray. Making the lens faster while maintaining the field height introduces larger aberrations. Larger aberrations (except distortion) further worsen the blur in the final image. Vignetting can help reduce the off-axis aberrations, but at the same time it causes light loss. The amount of the aberration and the amount of light flux through the lens are coupled with each other.

The last part in the imaging system is a digital imaging detector to produce a discrete set of pixel values representing a continuous image. CCD cameras now have very low noise and very high sensitivity. The electronic noise from CCDs has physical sources independent from other noise previously introduced. In Section 3 we show that it is fairly easy to include any independent noises in the analysis. For now, we analyze an ideal CCD camera with an identical response function for all pixels but no electronic noise. The fill factor of the CCD camera is assumed to be 1.

B. Mathematical Model

The above detailed descriptions and approximations of the physical model system provide us a solid ground to describe mathematically this imaging chain as cascaded stochastic point processes. This type of model for medical imaging is discussed in detail in Ref. 1. The chain begins just in front of the entrance surface of the fluorescent screen and ends at the output of the CCD camera.

1. Point Process of Absorbed X-Ray Photons—In radiography the object is the three-dimensional distribution of the x-ray attenuation coefficient, and the image is the two-dimensional (2D) projection of the x-ray distribution transmitted through that three-dimensional object. To facilitate our analysis of the point processes later in the paper, we use the x-ray distribution transmitted through the patient as the base for the object to our model system. When x-ray beams pass through the fluorescent screen, the crystallites inside the screen absorb a portion of the x rays. Only this portion of the x rays can excite the luminescent crystals to emit visible light and contribute to the image in the CCD camera. We therefore take this portion as the object to the model system and call it $f(\mathbf{R})$, where $\mathbf{R}$ is the 2D coordinate on the entrance surface of the screen. Although a realistic x-ray spectrum will introduce energy-dependent absorption coefficients, we think that this detail will not change the results qualitatively: The trends of the signal-to-noise ratio (SNR) curves and the orders of them in the various figures should remain the same, although the exact values can be different. An artificially monochromatic x-ray source is used in our study for the sake of computational simplicity. When we assume that the x-ray source is monochromatic, this object $f(\mathbf{R})$ can also be interpreted as the x-ray number fluence in units of mean number of photons per unit area. The energy fluence for a source is defined as the radiant energy per unit area in radiological imaging, and this quantity is called exposure in photography. The object is thus calculated from the energy fluence divided by the energy of each x-ray photon. Our object is in fact the noise-free x-ray pattern. If we take into account all the patients to be imaged, the object is a random function sampled from the ensemble of patients. Each object produces a noisy x-ray pattern. This x-ray pattern can be described as a 2D Poisson point process, where each x-ray photon contributes in the process a randomly located delta function on the entrance surface of the screen. The photons all propagate in the same direction that we denote as the $+z$ axis in the three-dimensional coordinate system. The entrance surface of the screen therefore lies in the x, y plane. Similar to the situation in patients, every x-ray photon interacts independently of all the other photons within the screen. The absorbed portion of the x-ray photons comprises another random point process that follows Poisson statistics. Mathematically speaking, the absorption mechanism is a binomial selection of the incident point process, and the binomial selection of a Poisson process is still a Poisson with a modified mean.[1] The mean of the Poisson process resulting from the absorption is the object.

Now we are ready to describe the initial random process in the imaging chain. It is the Poisson point process for any given object $f(\mathbf{R})$. This random process models both the object and the noise before the whole imaging system. A sample function of this first random process can be written as

$$y_x(\boldsymbol{R}) = \sum_{n=1}^{N} \delta(\boldsymbol{R} - \boldsymbol{R}_n), \tag{2.1}$$

where $\mathbf{R}_n$ is the 2D position vector in the entrance surface of the screen at which the nth x-ray photon hits, and each delta function describes an x-ray photon passing through the patient and absorbed by the screen. The x-ray object $f(\mathbf{R})$ is a random function drawn from the ensemble of objects to be imaged, so this process is actually a doubly stochastic Poisson process.[1] For any given object, the total number of absorbed x-ray photons N is a Poisson random variable, whose mean is the integral of the process mean $f(\mathbf{R})$. The integral is over the support of the object, which is usually the whole area of the screen: N

$$\bar{N}(f) = \iint_{\text{supp}} \mathrm{d}^2 R f(\boldsymbol{R}). \tag{2.2}$$

Because of the independence of the photons in the Poisson process, the probability density function (PDF) of $\mathbf{R}_n$ is just the properly normalized process mean function:

$$\mathrm{pr}_R(\boldsymbol{R}_n \,\big|\, f) = \frac{f(\boldsymbol{R}_n)}{\bar{N}(f)} = \frac{f(\boldsymbol{R}_n)}{\iint_{\text{supp}} \mathrm{d}^2 R f(\boldsymbol{R})}. \tag{2.3}$$

This 2D point process does not include the third spatial coordinate of the absorbed x-ray photons z, which is the absorption depth. The screen should have a uniform response over its area when well manufactured. The absorption depth z_n of the nth absorbed x-ray photon is thus independent of its 2D position $\mathbf{R}_n$. Given the characteristic x-ray absorption depth d_x, the absorption depth z_n follows the exponential law within the screen thickness d:

$$\mathrm{pr}_Z(z_n) = \frac{1}{d_x} \frac{\exp\left(-z_n \big/ d_x\right)}{1 - \exp\left(-d \big/ d_x\right)}, \quad 0 \le z_n \le d. \tag{2.4}$$

The proportion of the absorbed x rays in the x-ray pattern after the patient is $1 - \exp(-d/d_x)$.

The 2D point process of the absorbed x-ray photons, together with the PDF of the absorption depth, forms the complete mathematical description of the first stage in the cascaded imaging system model.

2. Point Process of Incident Visible Photons—Once an x-ray photon is absorbed, some of its energy is then re-emitted in the form of fluorescent light photons. Each absorbed x-ray photon can produce more than a thousand visible light photons on average. The initial x-ray interactions with the fluorescent crystals most likely produce high-energy photoelectrons. A high-energy electron migrating through inorganic scintillator crystallites inside the screen will form a large number of electron–hole (e–h) pairs. An e–h pair quickly recombines and independently emits a visible photon with high probability. The e–h pair sites are all so close together near the primary absorption site of the x-ray photon that all the excited visible lights can be considered as emitting from the absorption site. In fact, the e–h pair range is of the order of 2 μm in the mammographic energy range. After demagnification by the lens, the images of the e–h pair ranges are much smaller than the pixel size in common CCD cameras, which is of the order of tens of micrometers. The visible photons propagate equally probably in all directions upon creation. This simplified light conversion model is the starting point of the visible light scattering. The visible light photons are scattered by the tiny crystals in the screen and may be reflected by the diffuse reflectors on the back of the screen. Before they can get out of the screen, they have a slight chance of being absorbed along the way. The result of the scattering analysis follows next.

The whole process can be described mathematically as a random amplification process with x-ray photons as primaries and visible photons that get out of the exit surface of the screen as secondaries.[1–3] The amplification number of each absorbed x-ray photon is random. The nth primary produces K_n secondaries, and we can write down a sample function of the 2D stochastic process after this amplification mechanism:

$$y_{v,\,\mathrm{in}}(\boldsymbol{r}) = \sum_{n=1}^{N} \sum_{k=1}^{K_n} \delta(\boldsymbol{r} - \boldsymbol{r}_{nk})$$

$$= \sum_{n=1}^{N} \sum_{k=1}^{K_n} \delta(\boldsymbol{r} - \boldsymbol{R}_n - \Delta \boldsymbol{r}_{nk}), \tag{2.5}$$

where $\mathbf{r}_{nk}$ is the location on the exit surface of the screen for the kth secondary produced by the nth primary and $\Delta \mathbf{r}_{nk}$ is the random displacement. The position vectors here and in the rest of the paper are all 2D. They may be on different planes along the optical axis, but all the planes are parallel to each other and have identical coordinate systems with the origins on the optical axis. When projected along the optical axis on a certain plane, position vectors follow the common 2D vector arithmetic correctly.

In radiometry, a function called the bidirectional transmission distribution function [(BTDF) $(\mathbf{r}, \hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}')$] is used to quantify the directional property of light transmission through a thin transmissive layer.[1] It is defined as the ratio of transmitted radiance in the direction $\hat{\boldsymbol{s}}$ to the radiant incidence at the position $\mathbf{r}$ on the layer. Radiant incidence is the irradiance of a highly collimated beam traveling in the direction $\hat{\boldsymbol{s}}'$ so the BTDF actually specifies the transmitted radiance generated by such a beam. The BTDF is in units of inverse steradians, and we can relate the transmitted radiance with the incident radiance by the following integral:

$$\boldsymbol{L}_{\mathrm{trs}}(\boldsymbol{r}, \hat{\boldsymbol{s}}) = \iint_{2\pi} \mathrm{d}\Omega' \, \mathrm{BTDF}(\boldsymbol{r}, \hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}') \boldsymbol{L}_{\mathrm{inc}}(\boldsymbol{r}, \hat{\boldsymbol{s}}') \cos(\theta'), \tag{2.6}$$

where the integral is over the hemisphere of unit vector $\hat{\boldsymbol{s}}'$ directed toward the surface, and $\hat{\boldsymbol{s}}$ lies in the same hemisphere. Here we assume that the visible photons all have the same energy and, with the above assumption of monochromatic x-ray sources, the radiometric quantities of x ray and visible light can be defined in units of photon numbers.

Now we generalize the conventional BTDF to describe the entire light transmission through the fluorescent screen. For thick translucent objects like screens, the generalized BTDF is also a function of the position vector on the exit surface $\mathbf{r}'$, namely, BTDF $(\mathbf{r}, \hat{\boldsymbol{s}}; \mathbf{r}', \hat{\boldsymbol{s}}')$. The light-transmission process includes two steps: (1) the amplification process in which the absorbed x-ray photons are converted into a larger number of visible light photons and (2) the scattering process in which the visible light photons are displaced from the original x-ray absorption sites on the exit surface of the screen. In this analysis, the incident light is the absorbed x rays, and the transmitted light is the visible light out of the screen. Because the transmitted light displacement distribution can be different at different incident x-ray absorption depths, the generalized BTDF should also be a function of depth z. The generalized BTDF is thus similarly defined as the ratio of transmitted radiance in the direction $\hat{\boldsymbol{s}}$ at the position $\mathbf{r}$ on the exit plane to the radiant flux on the entrance plane when the incident x-ray is absorbed at the depth z.

Radiant flux is the flux of highly collimated x rays traveling in direction $\hat{\boldsymbol{s}}'$ with a pointlike size at the position $\mathbf{r}'$ on the entrance plane. Hence the generalized BTDF, by definition, is in units of inverse steradians per square meters, suggesting that it can be used in an integral over a solid angle and a 2D spatial vector. In fact, it is the kernel of the integral relating incident radiance to the transmitted radiance if the absorption happens at depth z:

$$L_{\text{trs}}(\boldsymbol{r}, \hat{\boldsymbol{s}} \big| z) = \iint_{2\pi} d\Omega' \int_{\infty} d^2 R \text{BTDF}$$
$$\times (\boldsymbol{r}, \hat{\boldsymbol{s}}, \boldsymbol{R}, \hat{\boldsymbol{s}}' \big| z) L_{\text{inc}}(\boldsymbol{R}, \hat{\boldsymbol{s}}') \cos(\theta'), \tag{2.7}$$

where all the radiometric quantities are in units of photon numbers. Therefore the mean gain of the amplification process is automatically incorporated into the generalized BTDF.

In our model, the incident radiance is an angular delta function scaled by the absorbed x-ray irradiance. Because we assume that x rays have normal incidence on the screen, the direction cosine is 1. The delta function is along the propagating direction $\hat{\boldsymbol{s}}'_0$ of all the x-ray photons, which is parallel to the optical axis, defined in the $+\hat{z}$ direction:

$$L_{\text{inc}}(\boldsymbol{R}, \hat{\boldsymbol{s}}') = \frac{M_{\text{inc}}(\boldsymbol{R})\delta(\hat{\boldsymbol{s}}' - \hat{\boldsymbol{s}}'_0)}{\cos(\theta')} = M_{\text{inc}}(\boldsymbol{R})\delta(\hat{\boldsymbol{s}}' - \hat{\boldsymbol{z}}), \tag{2.8}$$

where $\mathbf{M}_{\text{inc}}$ is the absorbed incident x-ray irradiance on the entrance surface at the position $\mathbf{R}$. By applying Eq. (2.7), it is easy to get the transmitted visible light radiance at the absorption depth z:

$$L_{\text{trs}}(\boldsymbol{r}, \hat{\boldsymbol{s}} \big| z) = \iint_{\infty} d^2 R M_{\text{inc}}(\boldsymbol{R}) \text{BTDF}(\boldsymbol{r}, \hat{\boldsymbol{s}}, \boldsymbol{r}', \hat{\boldsymbol{z}} \big| z). \tag{2.9}$$

The subsequent irradiance on the exit surface can be found by integrating the irradiance over all directions in the forward hemisphere:

$$M_{\text{trs}}(\boldsymbol{r} \big| z) = \iint_{2\pi} d\Omega L_{\text{trs}}(\boldsymbol{r}, \hat{\boldsymbol{s}} \big| z) \cos(\theta)$$
$$\iint_{\infty} d^2 R M_{\text{inc}}(\boldsymbol{R}) \iint_{2\pi} d\Omega \text{BTDF}$$
$$\times (\boldsymbol{r}, \hat{\boldsymbol{s}}, \boldsymbol{R}, \hat{\boldsymbol{z}} \big| z) \cos(\theta). \tag{2.10}$$

The kernel that relates incident irradiance to transmitted irradiance can also be called the irradiance point response function (PRF), which is

$$\text{PRF}(\boldsymbol{r}, \boldsymbol{R} \big| z) = \iint_{2\pi} d\Omega \text{BTDF}(\boldsymbol{r}, \hat{\boldsymbol{s}}, \boldsymbol{R}, \hat{\boldsymbol{z}} \big| z) \cos(\theta). \tag{2.11}$$

When in photon number units, the irradiances are equivalent to the mean processes f of the corresponding random point processes:

$$M_{\text{inc}}(\boldsymbol{R}) = f(\boldsymbol{R}), \quad M_{\text{trs}}(\boldsymbol{r} \big| z) = f'(\boldsymbol{r}). \tag{2.12}$$

The amplification PRF p_d can be derived from the irradiance PRF in Eq. (2.11):

$$p_d(\boldsymbol{r}, \boldsymbol{R} \big| z) = \iint_{2\pi} d\Omega \text{BTDF}(\boldsymbol{r}, \hat{\boldsymbol{s}}, \boldsymbol{R}, \hat{\boldsymbol{z}} \big| z) \cos(\theta)$$
$$= \bar{\mathcal{K}}(\boldsymbol{R} \big| z) \cdot \text{pr}_{\Delta \boldsymbol{r}}(\Delta \boldsymbol{r} \big| \boldsymbol{R}, z). \tag{2.13}$$

As shown in Ref. 1, the amplification PRF determines both the PDF of the displacement position $\text{pr}_{\Delta \mathbf{r}}$ and the average gain $\bar{\mathcal{K}}_n$ by

$$\mathrm{pr}_{\Delta \boldsymbol{r}}(\Delta \boldsymbol{r} \mid \boldsymbol{R}, z) = \frac{p_d(\boldsymbol{R} + \Delta \boldsymbol{r}, \boldsymbol{R} \mid z)}{\bar{\mathcal{K}}_n(\boldsymbol{R} \mid z)}, \tag{2.14}$$

$$\bar{\mathcal{K}}_n(\boldsymbol{R} \mid z) = \iint_\infty \mathrm{d}^2 r \, p_d(\boldsymbol{r}, \boldsymbol{R} \mid z), \tag{2.15}$$

where $\bar{\mathcal{K}}_n$ is the average number of secondaries per primary. We also note that the average number of secondaries per primary is independent of the particular primary, or n, and can be written as $\mathcal{K}$. Similarly, the PRF in the four-dimensional space (the 2D position and the 2D direction) is the generalized BTDF in photon number units corrected by the direction cosine:

$$\mathrm{PRF}(\boldsymbol{r}, \hat{\boldsymbol{s}}, \boldsymbol{R}, \hat{\boldsymbol{z}} \mid z) = \mathrm{BTDF}(\boldsymbol{r}, \hat{\boldsymbol{s}}, \boldsymbol{R}, \hat{\boldsymbol{z}} \mid z) \cos(\theta). \tag{2.16}$$

The joint PDF of the exit position and direction of a scattered secondary photon is thus the normalized four-dimensional PRF:

$$\mathrm{pr}(\boldsymbol{r}, \hat{\boldsymbol{s}} \mid \boldsymbol{R}, z) = \frac{\mathrm{BTDF}(\boldsymbol{r}, \hat{\boldsymbol{s}}, \boldsymbol{R}, \hat{\boldsymbol{z}} \mid z) \cos(\theta)}{\mathcal{K}(\boldsymbol{R} \mid z)}. \tag{2.17}$$

The conditional PDF of the outgoing direction can be derived from the joint PDF of Eq. (2.17). The direction is determined by the azimuth angle φ and the polar angle θ. The differential solid angle in the angular space is in fact the product of the differential azimuth angle $\mathrm{d}\varphi$ and the differential cosine of the polar angle $\mathrm{d}\cos(\theta)$. The direction PDF is thus expressed as the joint PDF of the azimuth angle φ and the cosine of the polar angle $\cos(\theta)$. The conditional PDF of the outgoing direction is expressed as

$$
\begin{aligned}
&\mathrm{pr}(\hat{\boldsymbol{s}} \mid \Delta \boldsymbol{r}, \boldsymbol{R}, z) \\
&= \frac{\mathrm{pr}(\boldsymbol{R} + \Delta \boldsymbol{r}, \hat{\boldsymbol{s}} \mid \boldsymbol{R}, z)}{\mathrm{pr}(\Delta \boldsymbol{r} \mid \boldsymbol{R}, z)} \\
&= \frac{\mathrm{BTDF}(\boldsymbol{R} + \Delta \boldsymbol{r}, \hat{\boldsymbol{s}}, \boldsymbol{R}, \hat{\boldsymbol{z}} \mid z) \cos(\theta)}{\iint_{2\pi} \mathrm{d}\Omega' \, \mathrm{BTDF}(\boldsymbol{R} + \Delta \boldsymbol{r}, \hat{\boldsymbol{s}}', \boldsymbol{R}, \hat{\boldsymbol{z}} \mid z) \cos(\theta')}.
\end{aligned}
\tag{2.18}
$$

The conditional direction PDF of Eq. (2.18) is completely determined by the generalized BTDF, but the PDF of the gain is still unknown except for the mean value. The amplification mechanism removes the independence among the different points in the resultant process because two or more secondaries can be produced by one primary. Shift invariance holds inside the boundary of well-manufactured screens, where the secondary displacement PDF [Eq. (2.14)] and the conditional direction PDF [Eq. (2.18)] are both independent of the primary position $\boldsymbol{R}$. The PRF $p_d(\mathbf{r}, \mathrm{R}|z)$ is now only a function of the displacement $\mathbf{r} - \mathbf{R}$ and the depth z, and it can be called the point-spread function. The secondaries from different primaries are independently produced since the screen should have no memory from one primary to another. The multivariate density on $\{\Delta \mathbf{r}_{nk}\}$ is therefore a product of univariate densities on each of the $\Delta \mathbf{r}_{nk}$. The product form is also valid for the multivariate density on $\{(\Delta \mathbf{r}_{nk}, \hat{\boldsymbol{s}}_{nk})\}$ and the conditional multivariate density on $\{\hat{\boldsymbol{s}}_{nk} | \Delta \mathbf{r}_{nk}\}$. Similarly, the conditional multivariate density $\mathrm{pr}(\{\mathbf{r}_{nk}\} | \mathbf{R}_n)$ is the product of the univariate densities $\mathrm{pr}(\mathbf{r}_{nk} | \mathbf{R}_n)$.

3. Point Process of Output Visible Photons—The visible light from the screen has to pass through the optical imaging lens after the screen to contribute to the image formation on the CCD camera. Because of the limited entrance pupil and the complicated vignetting effect of the lens, each visible photon has only a small chance to pass. The passing probability pass $(\mathbf{R} + \Delta \mathbf{r}, \mathbf{R}|z)$ of an output visible photon depends on its exit position $\mathbf{R} + \Delta \mathbf{r}$ and its primary

position $\mathbf{R}$. On the basis of the probability densities above, the function $\text{pass}(\mathbf{R} + \Delta\mathbf{r}, \mathbf{R}|z)$ can be obtained from the generalized BTDF function:

$$
\begin{aligned}
&\text{pass}(\boldsymbol{R} + \Delta\boldsymbol{r}, \ \boldsymbol{R}|\ z) \\
&= \iint_{\Omega(\boldsymbol{R}+\Delta\boldsymbol{r})} d\Omega \, \text{pr}(\hat{\boldsymbol{s}}|\ \Delta\boldsymbol{r}, \ \boldsymbol{R}, \ z) \\
&= \frac{\iint_{\Omega(\boldsymbol{R}+\Delta\boldsymbol{r})} d\Omega \, \text{BTDF}(\boldsymbol{R} + \Delta\boldsymbol{r}, \ \hat{\boldsymbol{s}}, \ \boldsymbol{R}, \ \hat{\boldsymbol{z}}|\ z) \cos(\theta)}{\iint_{2\pi} d\Omega \, \text{BTDF}(\boldsymbol{R} + \Delta\boldsymbol{r}, \ \hat{\boldsymbol{s}}, \ \boldsymbol{R}, \ \hat{\boldsymbol{z}}|\ z) \cos(\theta)},
\end{aligned}
\tag{2.19}
$$

where the solid angle $\Omega(\mathbf{R} + \Delta\mathbf{r})$ incorporates only the directions that a visible photon can pass through the lens from the position $\mathbf{R} + \Delta\mathbf{r}$ on the exit surface of the screen. The solid angle is determined by the lens pupil and vignetting at the field $\mathbf{R} + \Delta\mathbf{r}$. This binomial selection mechanism needs to be explicitly expressed because the input process is no longer a Poisson process. One sample function of the resultant process is

$$
y_{v,\text{out}}(\boldsymbol{r}) = \sum_{n=1}^{N} \sum_{k=1}^{K_n} \beta_{nk} \delta(\boldsymbol{r} - \boldsymbol{R}_n - \Delta\boldsymbol{r}_{nk}),
\tag{2.20}
$$

where β_{nk} is a random variable taking only two values: It is 0 when the kth visible photon produced by the nth x-ray photon does not pass through the lens or 1 when the photon reaches the CCD camera. The probability law on β_{nk} is

$$
\begin{aligned}
&\text{Pr}\,(\beta_{nk}|\ \boldsymbol{R}_n, \ \Delta\boldsymbol{r}_{nk}, \ z_n) \\
&= \begin{cases} \text{pass}\,(\boldsymbol{R}_n + \Delta\boldsymbol{r}_{nk}, \ \boldsymbol{R}_n|\ z_n), & \beta_{nk} = 1 \\ 1 - \text{pass}\,(\boldsymbol{R}_n + \Delta\boldsymbol{r}_{nk}, \ \boldsymbol{R}_n|\ z_n), & \beta_{nk} = 0 \end{cases}
\end{aligned}
\tag{2.21}
$$

Each visible photon travels independently inside the lens, so the multivariate density on $\{\beta_{nk}\}$ is the product of the univariate densities on each β_{nk}.

4. Point Process of Visible Photons in the Image—Because of aberrations, each passed visible photon is displaced from the ideal image position on the image plane. For a given photon (or ray) direction, this deviation is deterministic, but the photon directions are random. Thus aberrations translate randomness of the incident directions $\hat{\boldsymbol{s}}$ starting from the position $\mathbf{r}$ on the exit surface of the screen into randomness of the position displacements $\Delta\mathbf{r}''$ away from the ideal image positions $\mathbf{r}''$ on the CCD detector surface. Another 2D *spatial* stochastic process is produced by this translated randomness.[4] A sample function of this process is

$$
y'_g(\boldsymbol{r}'') = \sum_{n=1}^{N} \sum_{k=1}^{K_n} \beta_{nk} \delta(\boldsymbol{r}'' - \boldsymbol{r}''_{nk}),
\tag{2.22}
$$

where $\mathbf{r}''_{nk}$ is the position on the detector plane of the kth visible photon produced by the nth x-ray photon.

Because of the magnification of the lens, ideal image positions are generally not the same as the corresponding object positions. We can properly scale the sample functions on the CCD detector plane to cancel the image magnification. In particular, every position vector is divided by the magnification so that the ideal image points coincide with the corresponding object points. The resultant sample function is expressed as

$$y_g(\boldsymbol{r}') = \sum_{n=1}^{N} \sum_{k=1}^{K_n} \beta_{nk} \delta(\boldsymbol{r}' - \boldsymbol{r}'_{nk})$$

(2.23)

$$= \sum_{n=1}^{N} \sum_{k=1}^{K_n} \beta_{nk} \delta(\boldsymbol{r}' - \boldsymbol{R}_n - \Delta\boldsymbol{r}_{nk} - \Delta\boldsymbol{r}'_{nk}),$$

where $\Delta\mathbf{r}'_{nk}$ is the scaled displacement from the scaled ideal image point position that is the same as the object point $\mathbf{R}_n + \Delta\mathbf{r}_{nk}$. This scaled process again can be treated as a particular random amplification process with the nonrandom gain equaling 1. The PDF of the displacement equals its PRF $p_g(\mathbf{r}', \mathbf{r}|\mathbf{R}, z)$:

$$p_g(\boldsymbol{R} + \Delta\boldsymbol{r} + \Delta\boldsymbol{r}',\ \boldsymbol{R} + \Delta\boldsymbol{r}\ \Big|\ \boldsymbol{R},\ z) = \mathrm{pr}_{\Delta\boldsymbol{r}}'(\Delta\boldsymbol{r}'\ \Big|\ \boldsymbol{R},\ \Delta\boldsymbol{r},\ z).$$

(2.24)

We note that the PRF is not the properly scaled continuous spot diagram from optical design softwares such as CODEV or ZEMAX; the spot diagram is generated from the particular incident direction distribution density that is uniform in the direction tangent. Nor can we compute the PRF from the product of the spot diagram and the actual PDF of the direction. Analytically, we can employ the transformation law to get the PDF of $\Delta\mathbf{r}'$ from the PDF of the $\hat{\boldsymbol{s}}$ inside the passing cone $\Omega(\mathbf{R} + \Delta\mathbf{r})$. For a given lens, any displacement $\Delta\mathbf{r}'$ depends on the object point position $\mathbf{r}$ and the incident direction $\hat{\boldsymbol{s}}$ before the lens:

$$\Delta\boldsymbol{r}' = \Delta\boldsymbol{r}'(\hat{\boldsymbol{s}},\ \boldsymbol{r}) = \Delta\boldsymbol{r}'(\hat{\boldsymbol{s}},\ \boldsymbol{R} + \Delta\boldsymbol{r}).$$

(2.25)

We should be aware that the conditional direction PDF inside the passing cone is the original PDF normalized by the probability to be inside the cone:

$$\mathrm{pr}_{\hat{\boldsymbol{s}}}(\hat{\boldsymbol{s}}\ \Big|\ \boldsymbol{R},\ \Delta\boldsymbol{r},\ z)$$

$$= \frac{\mathrm{BTDF}(\boldsymbol{R} + \Delta\boldsymbol{r},\ \hat{\boldsymbol{s}},\ \boldsymbol{R},\ \hat{\boldsymbol{z}}|\ z)\cos(\theta)}{\iint_{\Omega(\boldsymbol{R}+\Delta\boldsymbol{r})} d\Omega'\ \mathrm{BTDF}(\boldsymbol{R} + \Delta\boldsymbol{r},\ \hat{\boldsymbol{s}},\ \boldsymbol{R},\ \hat{\boldsymbol{z}}|\ z)\cos(\theta')}.$$

(2.26)

When finding the mean value of some function of the displacement $u(\Delta\mathbf{r}')$, we obtain the integral relation between the point-spread function p_g and the conditional direction PDF as

$$\iint_{\infty} d^2\Delta r'\ u(\Delta\boldsymbol{r}')\mathrm{pr}_{\Delta\boldsymbol{r}}'(\Delta\boldsymbol{r}'\ \Big|\ \boldsymbol{R},\ \Delta\boldsymbol{r},\ z) = \iint_{\Omega(\boldsymbol{R}+\Delta\boldsymbol{r})} d\Omega\, u[\Delta\boldsymbol{r}'(\hat{\boldsymbol{s}},\ \boldsymbol{r})]\mathrm{pr}(\hat{\boldsymbol{s}}\ \Big|\ \boldsymbol{R},\ \Delta\boldsymbol{r},\ z)$$

$$= \frac{\iint_{\Omega(\boldsymbol{R}+\Delta\boldsymbol{r})} d\Omega\, u[\Delta\boldsymbol{r}'(\hat{\boldsymbol{s}},\ \boldsymbol{r})]\mathrm{BTDF}(\boldsymbol{R} + \Delta\boldsymbol{r},\ \hat{\boldsymbol{s}},\ \boldsymbol{R},\ \hat{\boldsymbol{z}}|\ z)\cos(\theta)}{\iint_{\Omega(\boldsymbol{R}+\Delta\boldsymbol{r})} d\Omega'\ \mathrm{BTDF}(\boldsymbol{R} + \Delta\boldsymbol{r},\ \hat{\boldsymbol{s}}',\ \boldsymbol{R},\ \hat{\boldsymbol{z}}|\ z)\cos(\theta')},$$

(2.27)

where the direction $\hat{\boldsymbol{s}}$ is chosen inside the passing cone Ω $(\mathbf{R} + \Delta\mathbf{r})$. Whenever the angular space of passed $\hat{\boldsymbol{s}}$ is mapped onto the $\Delta\mathbf{r}'$ space, this integral relation is always true. We do not usually know the analytical expression of $\Delta\mathbf{r}'(\hat{\boldsymbol{s}},\mathbf{r})$ or $\mathbf{r}'(\hat{\boldsymbol{s}},\mathbf{r})$, but we can find precisely $\mathbf{r}'$ $(\hat{\boldsymbol{s}},\mathbf{r})$ with Snell's law and geometric optics. Given a group of randomly emanating photons corresponding to proper direction density, computers can trace them through the lens onto the CCD detector plane and automatically produce the requested sample of the above formidable PDF of the scaled final displacement on the CCD.

5. Random Image Vector—The last step is to map this 2D continuous image function $y_g(\mathbf{r}')$ into a 2D discrete image array $\mathbf{g}$. Remember that we ignore the electronic noise from the CCD for now, so the final image array is

$$g = \{g_{\boldsymbol{m}}, \ \boldsymbol{m} = \{1, \ 1\}, \ \{1, \ 2\}, \ \dots, \ \{M, \ M\}\},$$

$$g_{\boldsymbol{m}} = \iint_{\mathrm{supp}} \mathrm{d}^2 r' \, g(\boldsymbol{r}')h_{\boldsymbol{m}}(\boldsymbol{r}'), \tag{2.28}$$

where $h_{\mathbf{m}}(\mathbf{r}')$ is the response function of the **m**th pixel. This mapping from a continuous function to a discrete vector can be denoted by a discretization operator *D*:

$$g = D y_g, \tag{2.29}$$

$$[D y_g]_{\boldsymbol{m}} = \iint_{\mathrm{supp}} \mathrm{d}^2 r' \, y_g(\boldsymbol{r}')h_{\boldsymbol{m}}(\boldsymbol{r}'). \tag{2.30}$$

Under the ideal CCD model, the response function is the 2D rectangle function centered at the corresponding pixel with the size of the pixel. Every response function has the same size, and adjacent response functions have no gap in between.

3. MERIT FUNCTION

Improving image quality is the central goal for imaging system optimization such as lens design. A task-based assessment of image quality is rooted in statistical decision theory, where the observer performance in the specific task is used to measure the image quality. The major type of task of a mammographic imaging system is to help diagnose breast tumors, which is a classification task. Therefore the measure of the observer performance in the classification task can be used as the merit function for optimizing lens design.

A. Observer Performance

The observer is the strategy by which the task gets done. Both human and computer algorithms can perform classification tasks, and the algorithms are called mathematical observers. Mathematical observers calculate the outcome from some operation on the data, and the outcome is called the test statistic. In addition to the test statistic, mathematical observers also have a rule of dividing the data space into several regions without ambiguity based on the test statistic. Each region corresponds to a class or hypothesis.

For a general *L*-class task, the Bayesian observer has complete knowledge of the statistical information regarding the task and uses that knowledge to make the best decision. Thus the performance of the Bayesian observer provides an upper bound for all the other observer performances. It is also called the ideal observer. The test statistic of the ideal observer (the likelihood ratio) is usually nonlinear in the data and difficult to implement. Another type of observer has the linear test statistic $\mathbf{t} = \mathbf{W}^t \mathbf{g}$ in the data $\mathbf{g}$, and the template $\mathbf{W}$ is a matrix. The Hotelling observer is the optimal linear observer in the sense that it maximizes the Hotelling trace. The Hotelling trace is a useful performance measurement that calculates the separability of the classes. The expression of the Hotelling trace in a two-class case is given in this section.

The first task of mammographic imaging systems is to make the decision between two hypotheses: tumor-present H_1 and tumor-absent H_0. For this binary or two-class task, every mathematical observer uses a scalar test statistic t and compares it with a threshold t_c. The observer makes decision D_1 when the statistic is larger than the threshold; otherwise, the observer makes decision D_0. The conditional probability $\Pr(D_1|H_1)$ is called the true positive fraction, and $\Pr(D_1|H_0)$ is called the false positive fraction; they are both functions of the threshold. We can plot the true positive fraction versus the false positive fraction with the threshold as the changing parameter. The plot is referred to as the receiver operating characteristic (ROC) curve. The area under the ROC curve (AUC) can also be used as a

performance measure. The Hotelling observer for the binary detection task has the template vector **w** as follows:

$$w = K_g^{-1} \Delta \bar{g}, \tag{3.1}$$

$$K_g = \frac{1}{2}(K_1 + K_0), \tag{3.2}$$

$$\Delta \bar{g} = \bar{g}_1 - \bar{g}_0. \tag{3.3}$$

where $\mathbf{K}_1$, $\mathbf{K}_0$, $\bar{g}_1$, and $\bar{g}_0$ are the covariance matrices and the mean image vector under each hypothesis, respectively. The Hotelling trace is now $\mathcal{T}_{\text{Hot}} = 4\Delta\bar{g}^t K_g^{-1} \Delta\bar{g}$. We define a new performance measure called the Hotelling SNR as a quarter of the Hotelling trace:

$$\text{SNR}_{\text{Hot}}^2 = \Delta\bar{g}^t K_g^{-1} \Delta\bar{g}. \tag{3.4}$$

Only the first and second moments of the data appear in the SNR definition, which relaxes the requirement for the full conditional probabilities needed in the AUC calculation. When the test statistic is Gaussian distributed under both hypotheses, the AUC can be monotonically mapped to the SNR. The mapping is through the error function

$$\text{SNR}_{\text{Hot}} = 2\,\text{erf}^{-1}(2 \times \text{AUC} - 1). \tag{3.5}$$

For non-Gaussian test statistics, the result of this mapping is called the detectability index d_A.

Observers for the tumor detection task are usually radiologists who look at the images formed by the mammographic system. The human observer's AUC under this task can be measured from psychophysical studies. This experiment is limited by the resources and is very time-consuming. We need a mathematical observer to model the human observer to facilitate the performance measurement. The ROC curve and the AUC are used to compare the models. The channelized Hotelling observer is a computationally feasible model whose ROC curve correlates well with that of the human observer. The channelized Hotelling observer was first introduced to medical image quality assessment by Myers and Barrett.[5] On the basis of previous work in our group,[6] the channelized Hotelling observer can satisfactorily predict the human observer performance in a wide range of situations when the tumor profile is deterministic. The notion of channels in the human visual system has been studied intensively in vision science for many years.[7] Gabor functions are widely used as one type of channel set. Several other kinds of functions have been used instead of Gabor functions. Difference of Gaussians (sometimes called DOGs),[8,9] a Gaussian-derivative model,[10] the Laplacian operator,[11] and Cauchy functions[12] have all been used in the research. On the basis of these studies in vision research, the introduction of the channels into the model of the human observer has been successful in a variety of circumstances.[13–15]

We will use the channelized Hotelling observer in our study here. The observer applies the Hotelling observer's strategy on the channel output of the image. The channel output is defined by the transformation $\mathbf{u} = \mathbf{T}^t\mathbf{g}$, where each column of the matrix **T** represents the spatial profile of a channel. The observer template is given by

$$w_{\text{ch}} = T(T^t K T)^{-1} T^t \Delta\bar{g}, \tag{3.6}$$

where $\mathbf{K}$ and $\Delta \bar{\boldsymbol{g}}$ are defined in Eqs. (3.2) and (3.3). The AUC of the channelized Hotelling observer is difficult to get directly. By the central limit theorem, the test statistic $\mathbf{w}^t \mathbf{g}$ is approximately normally distributed since it is a linear combination of all the image vector elements. The monotonic relationship between the SNR and the AUC in Eq. (3.5) is approximately true. Because a monotonic transformation preserves the performance order, we use the SNR as the performance measure. The SNR on the channel outputs is calculated as

$$\mathrm{SNR}_{\mathrm{ch}}^2 = \Delta \bar{\boldsymbol{g}}^t T (T^t K T)^{-1} T \Delta \bar{\boldsymbol{g}}. \tag{3.7}$$

The SNR of the channelized Hotelling observer is the task-based merit function to be used in this paper.

B. Merit Function Calculation

To compute the merit function, we need the first and second moments of the image array $\bar{\boldsymbol{g}}_j$ and $\mathbf{K}_j$ under both hypotheses $j = 0, 1$. According to the relations among all the random variables, we should average from the bottom up. Similar analyses can be found in Refs. 1–3 and [16]. The procedure is as follows:

1. Average over $\{\Delta \mathbf{r}'_{nk}\}$ for fixed $\mathbf{R}_n$, z_n, and $\{\Delta \mathbf{r}_{nk}\}$ with density [Eq. (2.27)],

2. average over $\{\beta_{nk}\}$ for fixed $\mathbf{R}_n$, z_n, and $\{\Delta \mathbf{r}_{nk}\}$ with density [Eq. (2.21)],

3. average over $\{\Delta \mathbf{r}_{nk}\}$ for fixed $\mathbf{R}_n$ and z_n with density [Eq. (2.14)],

4. average over K_n for fixed $\mathbf{R}_n$ and z_n,

5. average over z_n for fixed f with density [Eq. (2.4)],

6. average over $\mathbf{R}_n$ for fixed f with density [Eq. (2.3)],

7. average over N for fixed f with the Poisson density law, and

8. average over f.

We will follow this procedure in the calculations of both the mean image vector and the image covariance matrix.

The mean image vectors and the covariance matrices are given by the following (see Appendix A or Chap. 11 in Ref. 1 for the derivation):

$$\bar{\boldsymbol{g}} = D \overline{\mathcal{J}_g}, \tag{3.8}$$

$$\overline{\mathcal{J}_g}(\boldsymbol{r}') = [H_1 \overline{f}](\boldsymbol{r}'), \tag{3.9}$$

$$\boldsymbol{K}_{\boldsymbol{g}} = D K_g D^\dagger, \tag{3.10}$$

$$\begin{aligned}
K_g(\boldsymbol{r}'_1, \boldsymbol{r}'_2) &= [H_1 \overline{f}](\boldsymbol{r}'_2) \delta(\boldsymbol{r}'_1 - \boldsymbol{r}'_2) \\
&+ [H_2 \overline{f}](\boldsymbol{r}'_1, \boldsymbol{r}'_2) \\
&+ [H_1 K_f H_1^\dagger](\boldsymbol{r}'_1, \boldsymbol{r}'_2),
\end{aligned} \tag{3.11}$$

where D is the discretization operator in Eq. (2.30). The linear operators H_1 and H_2 are defined as

$$[H_1 \bar{f}](\mathbf{r}') = \iint_{\text{supp}} \mathrm{d}^2 R\, \bar{f}(\mathbf{R}) p_{\text{tot}}(\mathbf{r}', \mathbf{R}), \tag{3.12}$$

$$p_{\text{tot}}(\mathbf{r}', \mathbf{R}) = \frac{1}{d_x} \int_0^d \mathrm{d}z\, \exp(-d/d_x) p_{\text{tot}}(\mathbf{r}', \mathbf{R}|z), \tag{3.13}$$

$$p_{\text{tot}}(\mathbf{r}', \mathbf{R}|z) = \iint_\infty \mathrm{d}^2 r\, p_d(\mathbf{r}, \mathbf{R}|z)\text{pass}$$
$$\times (\mathbf{r}, \mathbf{R}|z) p_g(\mathbf{r}', \mathbf{r}|\mathbf{R}, z), \tag{3.14}$$

$$[H_2 \bar{f}](\mathbf{r}_1', \mathbf{r}_2') = \frac{1}{d_x} \iint_{\text{supp}} \mathrm{d}^2 R\, \bar{f}(\mathbf{R}) \int_0^d \mathrm{d}z$$
$$\times \exp(-z/d_x) Q(\mathbf{R}, z) p_{\text{tot}}$$
$$\times (\mathbf{r}_1', \mathbf{R}|z) p_{\text{tot}}(\mathbf{r}_2', \mathbf{R}|z), \tag{3.15}$$

$$Q(\mathbf{R}, z) = \frac{\overline{m^2} - m}{m^2}, \tag{3.16}$$

$$[H_1 K_f H_1^\dagger(\mathbf{r}_1', \mathbf{r}_2') = \iint_{\text{supp}} \iint_{\text{supp}} \mathrm{d}^2 R_1 \mathrm{d}^2 R_2\, p_{\text{tot}}$$
$$\times (\mathbf{r}_1', \mathbf{R}_1) K_f(\mathbf{R}_1, \mathbf{R}_2) p_{\text{tot}}(\mathbf{r}_2', \mathbf{R}_2), \tag{3.17}$$

where f is the x-ray fluence; m is the number of visible photons produced by an x-ray photon absorbed at position $\mathbf{R}$ and depth z; and p_d, pass, and p_g are given in Eqs. (2.13), (2.19), and (2.24), respectively. The conditional total PRF $p_{\text{tot}}(\mathbf{r}', \mathbf{R}|z)$ at depth z and the overall total PRF $p_{\text{tot}}(\mathbf{r}', \mathbf{R})$ describe the total blur effect on the final image by the whole imaging system. The mean number of visible photons produced by an absorbed x-ray photon m is not necessarily the same as the mean number of secondaries per primary $\bar{K}(\mathbf{R}, z)$ in the incident visible photons stage because there may be visible light loss inside the screen due to the visible light absorption and light leakage through the edges. The visible light loss mechanism can depend on the position and depth, but m is usually well approximated as independent of them. The Q factor can thus be a constant as well. The Q factor is related to the Swank factor S used to describe the fluorescent screen in x-ray imaging.[1,17] The Swank factor is defined as the ratio of the squared mean gain and the second moment of the gain, $S = m^2 / \overline{m^2}$. The Q factor can be found by the following formula:

$$Q = \frac{1}{S} - \frac{1}{m}. \tag{3.18}$$

If we denote the x-ray fluence after normal breast tissue as the background $b(\mathbf{r})$, the tumor can be the additive x-ray fluence with the profile $s(\mathbf{r})$, so the x-ray fluence after tumor-present tissue is denoted as $b + s$. In the statistical decision theory, the tumor is the signal we want to detect. In the signal-known-exactly case, the signal profile s is nonrandom, whereas the background b can be random. The ensemble-average object function and the covariance function are $\bar{b}$ and K_b under the signal-absent hypothesis H_0; they are $\bar{b} + s$ and K_b under the signal-present hypothesis H_1. The SNR of the channelized Hotelling observer is readily derived from the formulas above. Note that all the operators involved are linear, so additions can interchange with these linear operators and work directly on the fluences. Specifically,

$$\text{SNR}_{\text{ch}}^2 = \Delta \bar{g}^t T(T^t K T)^{-1} T \Delta \bar{g}, \tag{3.19}$$

$$\Delta \bar{g} = D\Delta \overline{\mathcal{J}}_g = DH_1 \Delta f, \tag{3.20}$$

$$\Delta f = s, \tag{3.21}$$

$$K = D\left[\frac{1}{2}(K_1 + K_0)\right]D^{\dagger}, \tag{3.22}$$

$$\frac{1}{2}(K_1 + K_0) = [H_1(b + \frac{1}{2}s)]\delta(r_1' - r_2')$$
$$+ [H_2(b + \frac{1}{2}s)] + H_1 K_b H_1^{\dagger}. \tag{3.23}$$

Expanding the compact expressions, we are able to see the details underneath. The *i*th element of the difference of the mean image vector $\Delta \bar{g}$ is

$$[\Delta \bar{g}]_i = \iint d^2 r' \, [H_1 s](r') h_i(r')$$
$$= \iint d^2 R s(R) \iint d^2 r' \, P_{\text{tot}}(r', R) h_i(r'). \tag{3.24}$$

We write the *ij*th element of the matrix **K** in three parts, as follows:

$$K_{ij} = \iint d^2 r' \left[H_1\left(b + \frac{1}{2}s\right)\right](r') h_i(r') h_j(r')$$
$$+ \iiiint d^2 r_1' d^2 r_2' \left[H_2\left(b + \frac{1}{2}s\right)\right]$$
$$\times (r_1', r_2') h_i(r_1') h_j(r_2')$$
$$+ \iiiint d^2 r_1' d^2 r_2' (H_1 K_b H_1^{\dagger})$$
$$\times (r_1', r_2') h_i(r_1') h_j(r_2')$$
$$= \iint d^2 R \left[b(R) + \frac{1}{2}s(R)\right]$$
$$\times \iint d^2 r' \, P_{\text{tot}}(r', R) h_i(r') h_j(r')$$
$$+ \iint d^2 R \left[b(R) + \frac{1}{2}s(R)\right]\frac{1}{d_X}$$
$$\times \int_0^d dz \exp\left(-z/d_X\right) Q(R, z)$$
$$\times \iint d^2 r_1' P_{\text{tot}}(r_1', R \mid z) h_i(r_1')$$
$$\times \iint d^2 r_2' P_{\text{tot}}(r_2', R \mid z) h_j(r_2')$$
$$+ \iiiint d^2 R_1 d^2 R_2 K_b(R_1, R_2)$$
$$\times \iint d^2 r_1' P_{\text{tot}}(r_1', R_1) h_i(r_1')$$
$$\times \iint d^2 r_2' P_{\text{tot}}(r_2', R_2) h_j(r_2'). \tag{3.25}$$

The two formulas in Eqs. (3.24) and (3.25) are the basis for the following numerical computation.

4. NUMERICAL CALCULATION

We should not get lost in the messy expressions above. Instead, we can see that all the system's information is encapsulated into $p_{tot}(\mathbf{r}', \mathbf{R}|z)$ and $Q(\mathbf{R}, z)$. The object randomness is separated out in the mean b and the covariance function K_b. Thus when we design the system, all we need to consider are the two functions that describe the system. Moreover, the definition of $Q(\mathbf{R}, z)$ implies its constant nature since the conversion number m is independent of the absorption depth z and the position $\mathbf{R}$. There is only one function left, the total PRF $p_{tot}(\mathbf{r}', \mathbf{R}|z)$. Since the pixels in a CCD do not overlap, the first part of the covariance matrix is a diagonal matrix with the elements as $DH_1 f$. In fact, we need only the numerical simulation of the integral $\iint d^2 r' p_{tot}(\mathbf{r}', \mathbf{R}|z) h_i(\mathbf{r}')$.

The total PRF $p_{tot}(\mathbf{r}', \mathbf{R}|z)$ contains three parts: the spatial point-spread function of the screen $p_d(\mathbf{r}, \mathbf{R}|z)$, the passing probability of a visible photon pass$(\mathbf{r}, \mathbf{R}|z)$ at $\mathbf{r}$ with its parent x-ray photon at $\mathbf{R}$, and the PRF of the optical imaging system $p_g(\mathbf{r}', \mathbf{r}|\mathbf{R}, z)$. By the PDF of the displacement given in Eq. (2.27) on the CCD detector plane, the integral can be written as

$$
\begin{aligned}
&\iint d^2 r' \, p_{tot}(r', R| z) h_i(r') \\
&= \iint d^2 r' \iint d^2 r \, p_d(r, R| z) \\
&\quad \times \text{pass}(r, R| z) p_g(r', r| R, z) \\
&= \iint d^2 r \, p_d(r, R| z) \, \text{pass}(r, R| z) \\
&\quad \times \iint d^2 \Delta r' \, \text{pr}_{\Delta r}(\Delta r' | r, R| z) h_i(r + \Delta r') \\
&= \iint d^2 r \, p_d(r, R| z) \text{pass}(r, R| z) \\
&\quad \times \iint_{\Omega(r)} d\Omega \, \text{pr}_{\hat{s}}(\hat{s}| r, R, z) h_i(r + \Delta r'(r + R, \hat{s})).
\end{aligned}
\tag{4.1}
$$

When referred back to the expressions of p_d, pass, and $pr_{\hat{s}}$ in Eqs. (2.13), (2.19), and (2.18), the final integral form used in the Monte Carlo calculation is

$$
\begin{aligned}
&\iint d^2 r' \, p_{tot}(r', R| z) h_i(r') \\
&= \mathcal{K}(R, z) \iint \\
&\quad \times d^2 r \frac{\iint_{2\pi} d\Omega' \, \text{BTDF}(r, \hat{s}', R, \hat{z}| z)}{\iint_\infty d^2 r' \iint_{2\pi} d\Omega' \, \text{BTDF}(r', \hat{s}', R, \hat{z}| z)} \\
&\quad \times \iint_{\Omega(r)} d\Omega \frac{\text{BTDF}(r, \hat{s}, R, \hat{z}| z)}{\iint_{2\pi} d\Omega' \, \text{BTDF}(r, \hat{s}', R, \hat{z}| z)} \\
&\quad \times h_i(r + \Delta r'(\hat{s}, r)).
\end{aligned}
\tag{4.2}
$$

The first fraction is the spatial marginal of the BTDF and the second one is the angular part for the given spatial position. The optical system manifests itself in the integration area $\Omega(\mathbf{r})$ and the transformation $\Delta r'(\hat{s}, r)$. The mean gain of the screen $\mathcal{K}$ can also be independent of the x-ray absorption position $\mathbf{R}$ and the depth z to the first approximation. The inner and outer integrals can be computed with a Monte Carlo method, where the position vectors $\mathbf{r}$ and the direction vector $\hat{s}$ of a visible photon can be sampled from the corresponding densities. To find the displacement $\Delta \mathbf{r}'$ on the CCD, we can employ a geometrical-optical ray-tracing routine to

trace a ray starting at **r** and propagating into $\hat{\boldsymbol{s}}$ from the screen. Since we need only the $\mathcal{K}$ and Q, the entire PDF of the gain is not necessary.

The above integral contains the entire system and is critical for the optimization of the system. The rest of the SNR calculation is about the objects to be observed. In the statistical description of objects, the entire density is again not necessary; rather, only the first two moments for the Hotelling observer are necessary. In fact, only the mean and the covariance function of the background are needed since the signal is deterministic.

The clustered lumpy background is a way of generating textures that simulate real mammograms while still being analytically trackable.[18] It is generated when we first select a random number of cluster centers in the object area. Then each cluster center is replaced by a random number of blob centers according to a certain distribution density about the center. Finally, there is a blob centered at each blob position. The shapes of the blob profiles have the same functional forms but different parameters. The shape is elliptical with random orientation. This random background is a wide-sense stationary background with constant mean function and shift-invariant covariance function. A sample function is shown in Fig. 2.

A. Positive Semidefinite Functions

To evaluate the performance of the channelized Hotelling observer, we can compute the squared SNR directly. We are not calculating sample covariance matrices but are using the numerical integral to calculate directly the elements of the covariance matrix. The natural guarantee of positivity of the estimated SNR square is lost in our method. The positive semidefiniteness of the average covariance function ensures the nonnegativity of the squared SNR. The direct Monte Carlo method cannot guarantee the positive definiteness because small deviations in the elements of the matrix usually break it. Taking a look at the covariance function expression again, we find it is important to ensure the positive semidefiniteness of the background covariance function in the numerical computation. The theory of the positive semidefinite functions and their construction is beyond the scope of this paper. A detailed discussion on this issue can be found in Refs. 1 and [19–21].

The covariance function of the clustered lumpy background $K_b(\mathbf{R}_1, \mathbf{R}_2)$ is shift invariant since the background is stationary. In fact, we can find a real-valued symmetric function $\Phi(\mathbf{R})$ to construct the covariance function in the following form:

$$K_b(\boldsymbol{R}_1 - \boldsymbol{R}_2) = \iint \mathrm{d}^2 R \, \Phi(\boldsymbol{R}_1 - \boldsymbol{R}) \Phi(\boldsymbol{R} - \boldsymbol{R}_2). \tag{4.3}$$

The exact expression for the function $\Phi(\mathbf{r}')$ is derived in Appendix C. With this function, the third term of the covariance matrix of the final image is given as

$$
\begin{aligned}
\boldsymbol{K}_{ij}^{(3)} &\iint \mathrm{d}^2 R \iiiint \mathrm{d}^2 r_1' \mathrm{d}^2 R_1 \Phi \\
&\times (\boldsymbol{R}_1 - \boldsymbol{R}) P_{\mathrm{tot}}(\boldsymbol{r}_1', \, \boldsymbol{R}_1) h_i(\boldsymbol{r}_1') \\
&\times \iiiint \mathrm{d}^2 r_2' \mathrm{d}^2 R_2 \Phi(\boldsymbol{R}_2 \\
&- \boldsymbol{R}) P_{\mathrm{tot}}(\boldsymbol{r}_2', \, \boldsymbol{R}_2) h_j(\boldsymbol{r}_2'),
\end{aligned}
\tag{4.4}
$$

where $\mathbf{K}^{(3)}{}_{i,j}$ is the ijth element of the third term of the covariance matrix. The set $\{\iiiint \mathrm{d}^2 r_1' \mathrm{d}^2 R_1 \Phi(\mathbf{R}_1 - \mathbf{R}) \times p_{\mathrm{tot}}(\mathbf{r}_1', \mathbf{R}_1) h_i(\mathbf{r}_1')\}$ composes $n^2 \times 1$ column vector functions. The matrix function given by the product of the vector function and its transpose is thus a positive-definite matrix. The integral of this positive-definite matrix over its argument is the third term of the

image covariance matrix and therefore is positive definite. We then perform the Monte Carlo integration over the vector function directly to get the sample function. After the above procedure, the resultant estimate of the third term is always positive definite.

B. Channel Profiles

We use the radially symmetric difference-of-Gaussian channels in this study because the signal profiles in the image are close to radially symmetric. The channel profiles are described in terms of the Fourier-transform space of the original 2D space:

$$C_j(\rho) = \exp\left[-\frac{1}{2}\left(\frac{\rho}{A\sigma_j}\right)^2 \right] - \exp\left[-\frac{1}{2}\left(\frac{\rho}{\sigma_j}\right)^2 \right], \qquad (4.5)$$

where ρ is the radial spatial-frequency variable. The standard deviation of each channel is defined by $\sigma_j = \sigma_0 \alpha^j$ from an initial value σ_0. The multiplicative factor $A > 1$ defines the bandwidth of the channel. We then implement the channels in the discrete image space by using the inverse Fourier transform followed by a digitization operator D defined above. To compare the response of each channel, the channel profile is normalized so that $2\pi\int \mathrm{d}r' r' t_j(r')^2 = 1$, where $t_j(r')$ is the spatial sensitivity function calculated as the inverse Fourier transform of $C_j(\rho)$.

5. RESULTS AND DISCUSSIONS

The clustered lumpy background has the same parameters throughout the simulation. The characteristic lengths of each blob before any random rotation are $Lx = 1.5$ mm and $L_y = 0.6$ mm. The mean number of blobs per cluster is $\overline{K}_{blob} = 20$, and the mean number of the clusters per square centimeter is $\overline{N}_{cluster} = 10.17$ cm^{-2}. The average number of blobs $\overline{K}_{blob}$ should not be confused with the mean number of secondaries per primary $\overline{K}$ in the previous discussion. The adjustable coefficients in the blob's exponent are $\alpha = 3.6373$ mm$^{1/2}$ and $\beta = 0.5$. The standard deviation of the distribution density of the blob centers inside a cluster is $\sigma_\Phi = 3.6$ mm. The signal we used is a Gaussian-profile blob with the standard deviation of $\sigma = 1.0$ mm. The maximum signal level at its center is 10% of the mean background level. A unit level in the background and signal corresponds to 58.1 x-ray photons so that the exposure on the screen is comparable to the typical screen-film mammographic system with a grid that is ~9.5 mR.[22]

The parameters of the fluorescent screen are primarily from the Kodak company's mammographic screens, which use gadolinium oxysulfide activated with terbium as the phosphor. The size of the screen is 24 cm × 30 cm, and the thickness is 84 μm. The x-ray absorption coefficient is $\alpha_x = 1.3$ cm^{-1}. On average, each absorbed x-ray photon in the mammographic range produces 1000 visible photons that can come out of the exit surface of the screen regardless of the absorption depth. To get the Q factor in the screen model, we need the Swank factor S.[23] The value S for this screen is 0.8.[24] To a first approximation, the generalized BTDF can be the product between the separate spatial part and the angular part. The angular part is independent of the spatial information. We approximate the spatial distribution density of the visible photons as Gaussian with the standard deviation equal to the absorption depth. The angular distributions on all spatial positions are assumed to follow the Lambertian approximation.

The imaging lens is the part containing changing parameters. We will study the effects of these parameters on the image quality defined by the SNR of the channelized Hotelling observers. There is a folding mirror between the screen and the lens to prevent the x-ray photons from striking onto the CCD camera directly. In the simulation, the mirror is unfolded and not included in the ray-tracing routine.

The CCD camera has 128×128 pixels with a pixel size of 50 μm $\times$ 50 μm. The quantum efficiency is unity, and there are no gaps among adjacent pixels. This simple model of an ideal CCD camera is adequate for our simulation.

The channels used in the observer model are the dense difference-of-Gaussian channels mentioned in Abbey and Barrett's study.[6] This model uses ten channels. The radial frequency profile of the jth channel is the difference of the two Gaussians. The initial standard deviation in frequency is $\sigma_0 = 0.005$ pixel^{-1}, and the ratio of adjacent standard deviations is $\alpha = 1.4$. The multiplicative factor defining the bandwidth of the channel is $A = 1.67$.

The signal position changes from an on-axis spot up to the edge of the screen. The signal position is known to the observer so that the center of the channel set is on the image of the signal center each time. When the magnification is not large enough to cover the full screen in a single CCD shot, we will move the CCD accordingly until the edge is imaged. Each time the CCD camera is shifted, the signal to be imaged will be guaranteed to be within the CCD detector and not too close to the boundary.

The first experiment is to study the effect of the aperture size of the lens. The plots of SNR versus signal position at different aperture sizes are shown in Fig. 3, and the corresponding spot diagrams at four different fields are shown in Fig. 4. The CCD camera is at the minimum rms point size position, which is commonly called the best focus position. An iris is used on the stop to control the numerical aperture of the lens. With the parameters unchanged, except for the iris diameter, the aperture size of the imaging system can be controlled. The CCD camera remains at the minimum rms spot size position as the aperture size is changed throughout this experiment. When the iris is opened up, the light throughput becomes larger so that the exposure during the same time interval becomes larger. The relatively large demagnification of the lens cuts off a large portion of the light by the relatively small numerical aperture in object space. The visible light that finally gets onto the CCD detector is so dim that it makes the Poisson noise from the x-ray photons in-significant. Instead, the lens becomes the quantum sink. Larger exposure generally improves the discrimination ability of observers because it makes the lens less of a quantum sink. Meanwhile, more aberrations are introduced into the image and make the images more blurred. The signal is gradually smeared into the background by more and more blur. Blurring effects are usually detrimental to the observer's ability to pick a signal out of the background. The trade-off between the exposure and the blur, i.e., between the numerical aperture and the aberration, can be defined quantitatively by a certain aperture size, which makes the observer have the highest SNR.

In this experiment, we vary the object-space numerical aperture from 0.005 up to 0.055, which corresponds to f-numbers from 10 down to 0.91. The SNR curves versus the signal position are compared in all f-number cases. From Fig. 3, the SNR increases with decreasing f-number at all fields in the f-number range from 10 to 2. The enhancement in the SNR demonstrates that the flux increase improves the observer performance at the relatively small apertures. From f-number 2 to under 1, the SNR has only limited increase at first and then is maximized at $\sim f/$ 1. Although we cannot open the aperture further, we can already observe the saturation of the SNR at around $f/1$. In this case, decreasing the f-number further will not help in detecting signals. The maximum SNR curve numerically defines the best f-number for this lens design. This f-number quantitatively indicates the balancing point between the flux and the aberrations in the task-based manner.

At the same f-number, the aberration changes with changing field, although the exposure is almost the same within a 10° field of view in this Lambertian condition. Therefore the aberration differences are the main reason for the SNR differences at the different signal positions given any single f-number. However, from the largest f-number at 10 down to 2, each

SNR curve is almost flat. From the spot diagram plots, the spot size is limited to approximately 2×2 pixels throughout the full range of the field of view. Therefore the observer performance is insensitive to the aberration changes when the spot width is within ~2 pixels. From *f*-number 2 to under 1, the SNR is not flat but has a maximum at ~ 0.7 field. From the spot diagrams, the spot size generally increases with the decreasing *f*-number and has a minimum at ~ 0.7 field. Most of the spots are wider than 2 pixels, which means that the observer performance is sensitive to the aberration when the spot size is beyond the 2-pixel limit. Looking at the 0.7 field in the spot diagrams, we find that the spot sizes are still around 2 pixels at different *f*-numbers. The trend of the SNR increasing with decreasing *f*-numbers is again observable at that field. Away from the 0.7 field, both aberration and exposure have visible effects on observer performance. Even though the spot size is less in the far off-axis field than in the near-axis field, the SNR is higher in the near-axis field than that far from the axis. This is because vignetting begins to reduce the exposure in the off-axis field at a low *f*-number, and the SNR enhancement from the smaller spot size is overcompensated by the SNR decrease from less exposure.

The second experiment is to study the effect of the working distance change of the same lens. The SNR plots at different working distances are shown in Fig. 5, and the corresponding spot diagrams are shown in Fig. 6. We choose the *f*/0.91 lens design from the above experiment since it yields the best observer performance. Each time we change the working distance, the lens is adjusted to the best focus position for that working distance. The closer the object is to the lens, the farther the CCD camera is from the lens and the higher the absolute magnification value. The magnification of the imaging system varies monotonically with the working distance, so we use the magnification to denote the working distance. At the designed working distance, the magnification is −0.10. This experiment examined the SNR changes at the different magnifications from −0.07 to −0.13. Because of the large demagnification used, the change in the working distance is much smaller than the distance between the screen and the lens. The observer's performance varies considerably in spite of the small changes in the working distance.

The flux becomes larger when the absolute magnification is increased because the solid angle viewed from the object plane becomes larger with the increased working distance. At the close-to-axis field range, the SNR curves decrease with the increasing absolute magnifications. In this field range, the spot size quickly increases when the absolute magnification becomes larger. The SNR drop shows that the small improvement in observer performance by the flux increase cannot overcome the negative effects of blurring. At a lower absolute magnification portion, the curves are largely overlapped with an insignificant decreasing trend with increasing absolute magnifications. From the spot diagrams, the spot sizes are within the 2-pixel width limit in this portion. Therefore the observer performs similarly since the flux difference is small. When we extend to the far off-axis field region, the changing trend of the SNR curves versus the magnifications reverts, with the SNR becoming higher at the higher absolute magnification. For magnifications lower than 0.1, the SNR increases quickly when the absolute magnification becomes larger. For the cases when the absolute magnifications are higher than 0.1, the SNR curves are very close. The spot size now has a minimum at the magnification −0.1 and increases toward both ends. Therefore the rapid improvement in observer performance below the magnification − 0.1 indicates the help from both the decreasing spot size and the increasing flux. When the absolute magnification is larger than 0.1, the detrimental effects from the increased spot size balance the positive effects from the flux, and observer performance remains the same. At ~0.7 field, the spot sizes vary little among different working distances, which makes the SNR almost unchanged.

When the absolute magnification is small, the SNR curves are relatively flat at the close-to-axis field range and generally go downward at the far off-axis field region. This is because the

observer can tolerate the aberrations within 2-pixel-wide spots, which is true in the near-axis field. At the higher absolute magnification area, the effects from the aberrations become visible, and the observer performs best at ~0.7 field since the spot sizes are minimum there. Although the spot sizes are obviously different between the full field and the zero field, the observer performs similarly. This indicates that the effects of aberration should be more than just the effects of the spot size.

In this experiment, both the flux and the aberration clearly affects the observer performance. The spot size is only a coarse summary on the effect of the aberrations. More detailed effects by the aberration should not be overlooked. We can select the working distance based on the region where the signal will show up. Above all, there is only a small improvement in the near-axis fields and a larger drop in the far off-axis fields at the small absolute magnifications, and there is nearly no improvement in all fields at large absolute magnifications. We should consider the original designed working distance as the best among all choices.

The last experiment is on the focus shift of the lens. The SNR plots at different image plane positions are shown in Fig. 7, and the corresponding spot diagrams are shown in Fig. 8. On the basis of the first two experiments, we wanted to study the effect of the aberration alone. We designed the third experiment since the defocus is the most convenient way to control the blur size of the imaging lens. Also, we can check whether the best focal plane defined by the smallest rms spot size is the best in the task-based sense. In this experiment, we again use the $f/0.91$ lens from the first experiment. The defocus is in terms of a CCD detector plane shift from the best focal plane. The shift ranges from -200 to $400\ \mu m$, where the negative shift means moving the CCD toward the lens, and the positive shift means moving it away. Now the fluxes in all the cases are the same. At the near-axis field region, the SNR becomes smaller when the CCD camera is moved closer to the lens. When the focus shift is between 100 and 300 μm, the SNR stays close and almost always remains at the maximum. When going out to the far off-axis field region, the SNR increases when the CCD camera is moved closer to the lens. When the shift is positive, the SNR curves overlap and remain at the upper bound. The complicated changes in SNR are related to the complex blurs by the different amounts of aberrations. From the spot diagrams, the SNR change can be explained approximately by the rms spot size, but we can also see the effect from the further details of aberrations. For the spots at the far off-axis field region and the negative focus shift, the rms spot size increases when moving closer to the lens, while the SNR is almost the same. This indicates that the tail in the spot does not affect the observer performance, even though it contributes to the rms spot size. A similar situation happens in the near-axis field region and the moderate positive focus shift; the tail around a much darker center has little effect on the SNR. On the basis of these observations, the rms spot size is not a very pertinent measure of the position of the best focus. In fact, we can see that the best focal plane should be moved a little farther away from the lens, within 100 μm from the rms best focal plane, as can be seen from the SNR curve of the 100-μm focus shift in Fig. 7.

6. CONCLUSIONS

From our simulation, the task-based lens design merit function in terms of the SNR of the channelized Hotelling observers can quantitatively take into account not only aberrations but also the flux and noise. The designers' subjective opinion on the relationship between the aberration and other facts of the design is eliminated. The connection is instead numerically established based on statistical signal detection theory and the mathematical human vision model through the channelized Hotelling observer. We therefore call this SNR an objective or task-based design criterion. To our knowledge, this is the first time the observer's performance has been introduced into the lens design.

We should note that the conclusions in this work, such as the best *f*-number at ~1.0 and the observer's tolerance at 2-pixel-wide spots, are pertinent only to the mammographic imaging condition. Objective performance measures are critically determined by the relevant flux level and noise characteristics together with the base lens type.

References

1. H. H. Barrett and K. J. Myers, Foundations of Image Science (Wiley, New York, 2004).

2. Rabbani M, Shaw R, Van Metter R. "Detective quantum efficiency of imaging systems with amplification and scattering mechanisms,". J Opt Soc Am A 1987;4:895–901. [PubMed: 3598742]

3. Cunningham IA, Westmore MS, Fenster A. "A spacial-frequency dependent quantum accounting diagram and detective quantum efficiency model of signal and noise propagation in cascaded imaging systems,". Med Phys 1994;21:417–427. [PubMed: 8208217]

4. B. R. Frieden, Probability, Statistical Optics, and Data Testing, 2nd ed. (Springer-Verlag, New York, 1991).

5. Myers KJ, Barrett HH. "Addition of a channel mechanism to the ideal-observer model,". J Opt Soc Am A 1987;4:2447–2457. [PubMed: 3430229]

6. 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.

7. N. Graham, "Complex channels, early nonlinearities, and normalization in texture segmentation," in Computational Models of Visual Processing, M. S. Landy and J. A. Movshon, eds. (MIT Press, Cambridge, Mass., 1990), pp. 273–290.

8. Frishman LJ, Freeman AW, Troy JB, Schweitzer-Tong DE, Enroth-Cugell C. "Spatiotemporal frequency responses of cat retinal ganglion cells,". J Gen Physiol 1987;89:599–628. [PubMed: 3585279]

9. Hawken MJ, Parker AJ. "Spatial properties of neurons on the monkey striate cortex,". Proc R Soc London, Ser B 1987;231:251–288. [PubMed: 2889214]

10. Webster MA, De Valois RL. "Relationship between spatial-frequency and orientation tuning of striate-cortex cells,". J Opt Soc Am A 1985;2:1124–1132. [PubMed: 4020509]

11. Marr D, Hildreth E. "Theory of edge detection,". Proc R Soc London, Ser B 1980;207:187–217. [PubMed: 6102765]

12. Klein SA, Levi DM. "Hyperacuity thresholds of 1 sec: theoretical predictions and empirical validation,". J Opt Soc Am A 1985;2:1170–1190. [PubMed: 4020514]

13. Fiete RD, Barrett HH, Smith WE, Myers KJ. "The Hotelling trace criterion and its correlation with human observer performance,". J Opt Soc Am A 1987;4:945–953. [PubMed: 3598746]

14. E. B. Cargill, "A mathematical liver model and its application to system optimization and texture analysis," Ph.D. dissertation (University of Arizona, Tucson, Ariz., 1989).

15. Eckstein MP, Whiting JS. "Lesion detection in structured noise,". Acad Radiol 1995;2:249–253. [PubMed: 9419557]

16. H. H. Barrett, R. F. Wagner, and K. J. Myers, "Correlated point processes in radiological imaging," in Physics of Medical Imaging, R. L. Van Metter and J. Beutel, eds., Proc. SPIE 3032, 110–124 (1998).

17. Swank RK. "Absorption and noise in x-ray phosphors,". J Appl Phys 1973;44:4199–4203.

18. Bochud FO, Abbey CK, Eckstein MP. "Statistical texture synthesis of mammographic images with clustered lumpy backgrounds,". Opt Exp 1999;4:33–43.http://www.opticsexpress.org

19. C. Berg, J. P. Reus Christensen, and P. Russel, Harmonic Analysis on Semigroups: Theory of Positive Definite and Related Functions (Springer-Verlag, New York, 1984).

20. M. H. Stone, Linear Transformations in Hilbert Space and Their Applications to Analysis, Vol. 15 of the American Mathematical Society Colloquium (American Mathematical Society, New York, 1932).

21. Aronszajn N. "Theory of reproducing kernels,". Trans Am Math Soc 1950;68:337–404.

22. Roehrig H, Yu T, Krupinski E. "Image quality control for digital mammographic systems: initial experience and outlook,". J Digital Imaging 1995;8:52–66.

23. Blevis IM, Hunt DC, Rowlands JA. "X-ray imaging using amorphous selenium: detection of swank factor by pulse height spectroscopy,". Med Phys 1998;25:638–641. [PubMed: 9608472]

24. D. P. Trauernicht and R. Van Metter, "The measurement of conversion noise in x-ray intensifying screens," in Medical Imaging II: Image Formation, Detection, Processing, and Interpretation, R. H. Schneider and S. J. Dwyer, eds., Proc. SPIE 914, 100–116 (1988).

APPENDIX A: MEAN IMAGE VECTOR

The discrete mean image array $\bar{\boldsymbol{g}}$ is given when the operator D is applied to the continuous mean image function $\langle y_g(\mathbf{r}')\rangle$:

$$
\begin{aligned}
\bar{\boldsymbol{g}} &= \left\langle \iint_{\text{supp}} y_g(\boldsymbol{r}')\boldsymbol{h}(\boldsymbol{r}')\mathrm{d}^2 r' \right\rangle \\
&= \iint_{\text{supp}} \left\langle y_g(\boldsymbol{r}') \right\rangle \boldsymbol{h}(\boldsymbol{r}')\mathrm{d}^2 r' = D(\left\langle y_g(\boldsymbol{r}') \right\rangle),
\end{aligned}
\tag{A1}
$$

where $\mathbf{h}(\mathbf{r}')$ is the vector whose elements are the pixel response functions $\{h_i(\mathbf{r}')\}$.

The calculation of $\langle y_g(\mathbf{r}')\rangle$ is performed as follows:

1. Average over the displacement on the CCD detector plane $\{\Delta\mathbf{r}'_{nk}\}$:

$$
\begin{aligned}
\left\langle y_g(\boldsymbol{r}') \right\rangle_{\{\Delta\boldsymbol{r}'_{nk}\}} &= \sum_n \sum_k \iint_\infty \mathrm{d}^2 r'_{nk}\beta_{nk}\delta \\
&\times (\boldsymbol{r}' - \boldsymbol{R}_n - \Delta\boldsymbol{r}_{nk} - \Delta\boldsymbol{r}'_{nk}) \\
&\times p_g(\boldsymbol{R}_n + \Delta\boldsymbol{r}_{nk} + \Delta\boldsymbol{r}'_{nk}, \boldsymbol{R}_n \\
&+ \Delta\boldsymbol{r}_{nk}\big| \boldsymbol{R}_n, z_n) \\
&= \sum_n \sum_k \beta_{nk} p_g(\boldsymbol{r}', \boldsymbol{R}_n + \Delta\boldsymbol{r}_{nk} \mid \boldsymbol{R}_n, z_n).
\end{aligned}
\tag{A2}
$$

1. Average over $\{\beta_{nk}\}$:

$$
\begin{aligned}
\left\langle y_g(\boldsymbol{r}') \right\rangle_{(\beta_{nk})} &= \sum_n \sum_k \left\langle \beta_{nk} \right\rangle p_g(\boldsymbol{r}', \boldsymbol{R}_n + \Delta\boldsymbol{r}_{nk} \big| \boldsymbol{R}_n, z_n) \\
&= \sum_n \sum_k \text{pass}(\boldsymbol{R}_n + \Delta\boldsymbol{r}_{nk}, \boldsymbol{R}_n \big| z_n) p_g(\boldsymbol{r}', \boldsymbol{R}_n \\
&+ \Delta\boldsymbol{r}_{nk}\big| \boldsymbol{R}_n, z_n).
\end{aligned}
\tag{A3}
$$

1. Average over the displacements on the exit surface of the screen $\{\Delta\mathbf{r}_{nk}\}$:

$$\langle y_g(\mathbf{r}') \rangle_{(\Delta \mathbf{r}_{nk})}$$

$$= \sum_n \sum_k \iint_\infty \mathrm{d}^2 r_{nk} \, \mathrm{pass}(\mathbf{R}_n + \Delta \mathbf{r}_{nk}, \ \mathbf{R}_n \big| \ z_n)$$

$$\times p_g(\mathbf{r}', \ \mathbf{R}_n + \Delta \mathbf{r}_{nk} \big| \ \mathbf{R}_n, \ z_n)$$

$$\times \overline{K}_n(\mathbf{R}_n, \ z_n)^{-1} p_d(\mathbf{R}_n + \Delta \mathbf{r}_{nk}, \ \mathbf{R}_n \big| \ z_n)$$

$$= \sum_n \overline{K}_n(\mathbf{R}_n, \ z_n)^{-1} \sum_k \iint_\infty \mathrm{d}^2 r \, p_d(\mathbf{r}, \ \mathbf{R}_n \big| \ z_n)$$

$$\times \mathrm{pass}(\mathbf{r}, \ \mathbf{R}_n \big| \ z_n) p_g(\mathbf{r}', \ \mathbf{r} \big| \ \mathbf{R}_n, \ z_n)$$

$$= \sum_n \overline{K}_n(\mathbf{R}_n, \ z_n)^{-1} K_n(\mathbf{R}_n, \ z_n) \iint_\infty \mathrm{d}^2 r \, p_d(\mathbf{r}, \ \mathbf{R}_n \big| \ z_n)$$

$$\times \mathrm{pass}(\mathbf{r}, \ \mathbf{R}_n \big| \ z_n) p_g(\mathbf{r}', \ \mathbf{r} \big| \ \mathbf{R}_n, \ z_n).$$

(A4)

1. Average over the gains $\{K_n\}$:

$$\langle y_g(\mathbf{r}') \rangle_{\{K_n\}} = \sum_n \iint_\infty \mathrm{d}^2 r \, p_d(\mathbf{r}, \ \mathbf{R}_n \big| \ z_n)$$

$$\times \mathrm{pass}(\mathbf{r}, \ \mathbf{R}_n \big| \ z_n) p_g(\mathbf{r}', \ \mathbf{r} \big| \ \mathbf{R}_n, \ z_n).$$

(A5)

1. Average over the absorption depth $\{z_n\}$:

$$\langle y_g(\mathbf{r}') \rangle_{(z_n)} = \sum_n \int_0^d \mathrm{d}z_n \frac{1}{d_x} \cdot \frac{\exp(-z_n / d_x)}{1 - \exp(-d / d_x)}$$

$$\times \iint_\infty \mathrm{d}^2 r \, p_d(\mathbf{r}, \ \mathbf{R}_n \big| \ z_n)$$

$$\times \mathrm{pass}(\mathbf{r}, \ \mathbf{R}_n \big| \ z_n) p_g(\mathbf{r}', \ \mathbf{r} \big| \ \mathbf{R}_n, \ z_n).$$

(A6)

1. Average over the absorption positions of x-ray photons $\{\mathbf{R}_n\}$:

$$\langle y_g(\mathbf{r}') \rangle_{(\mathbf{R}_n)}$$

$$= \sum_n \iint_{\mathrm{supp}} \mathrm{d}^2 R_n \frac{f(\mathbf{R}_n)}{\iint f(\mathbf{R}) \mathrm{d}^2 R}$$

$$\times \int_0^d \mathrm{d}z_n \frac{1}{d_x} \cdot \frac{\exp(-z_n / d_x)}{1 - \exp(-d / d_x)}$$

$$\times \iint_\infty \mathrm{d}^2 r \, p_d(\mathbf{r}, \ \mathbf{R}_n \big| \ z_n)$$

$$\times \mathrm{pass}(\mathbf{r}, \ \mathbf{R}_n \big| \ z_n) p_g(\mathbf{r}', \ \mathbf{r} \big| \ \mathbf{R}_n, \ z_n)$$

$$= \frac{1}{d_x \overline{N}(f)} \sum_n \iint_{\mathrm{supp}} \mathrm{d}^2 R_n \, f(\mathbf{R}_n) \int_0^d \mathrm{d}z_n$$

$$\times \exp(-z_n / d_x)$$

$$\times \iint_\infty \mathrm{d}^2 r \, p_d(\mathbf{r}, \ \mathbf{R}_n \big| \ z_n)$$

$$\times \mathrm{pass}(\mathbf{r}, \ \mathbf{R}_n \big| \ z_n) p_g(\mathbf{r}', \ \mathbf{r} \big| \ \mathbf{R}, \ z_n)$$

$$= \frac{N(f)}{d_x \overline{N}(f)} \iint_{\mathrm{supp}} \mathrm{d}^2 R f(\mathbf{R}) \int_0^d \mathrm{d}z \, \exp(-z / d_x)$$

$$\times \iint_\infty \mathrm{d}^2 r \, p_d(\mathbf{r}, \ \mathbf{R} \big| \ z) \mathrm{pass}(\mathbf{r}, \ \mathbf{R} \big| \ z) p_g(\mathbf{r}', \ \mathbf{r} \big| \ \mathbf{R}, \ z).$$

(A7)

1. Average over the number of absorbed x-ray photons N:

$$
\begin{aligned}
\left\langle y_g(\mathbf{r}^{'})\right\rangle_N = {}& \iint_{\text{supp}} d^2 R f(\mathbf{R}) \int_0^d dz \frac{1}{d_x} \exp\left(-z \middle/ d_x\right) \\
& \times \iint_\infty d^2 r\, p_d(\mathbf{r},\ \mathbf{R}\middle|\ z) \\
& \times \text{pass}(\mathbf{r},\ \mathbf{R}\middle|\ z) p_g(\mathbf{r}^{'},\ \mathbf{r}\middle|\ \mathbf{R},\ z).
\end{aligned}
\tag{A8}
$$

1. Average over the x-ray fluence before the entrance surface of the screen f:

$$
\begin{aligned}
\left\langle y_g(\mathbf{r}^{'})\right\rangle_N = {}& \frac{1}{d_x} \iint_{\text{supp}} d^2 R\, \overline{f}(\mathbf{R}) \int_0^d dz\, \exp\left(-z \middle/ d_x\right) \\
& \times \iint_\infty d^2 r\, p_d(\mathbf{r},\ \mathbf{R}_n\middle|\ z) \\
& \times \text{pass}(\mathbf{r},\ \mathbf{R}_n\middle|\ z) p_g(\mathbf{r}^{'},\ \mathbf{r}\middle|\ \mathbf{R},\ z).
\end{aligned}
\tag{A9}
$$

We define a conditional total PRF $p_{\text{tot}}(\mathbf{r}', \mathbf{R}|z)$, an overall total PRF $p_{\text{tot}}(\mathbf{r}', \mathbf{R})$, and a linear operator H_1:

$$
\begin{aligned}
p_{\text{tot}} = (\mathbf{r}^{'},\ \mathbf{R}\middle|\ z) = {}& \iint_\infty d^2 r\, p_d(\mathbf{r},\ \mathbf{R}\middle|\ z) \\
& \times \text{pass}(\mathbf{r},\ \mathbf{R}\middle|\ z) p_g(\mathbf{r}^{'},\ \mathbf{r}\middle|\ \mathbf{R},\ z),
\end{aligned}
\tag{A10}
$$

$$
p_{\text{tot}}(\mathbf{r}^{'},\ \mathbf{R}) = \frac{1}{d_x} \int_0^d dz\, \exp\left(-d \middle/ d_x\right) p_{\text{tot}}(\mathbf{r}^{'},\ \mathbf{R}),
\tag{A11}
$$

$$
[H_1\, f](\mathbf{r}^{'}) = \iint_{\text{supp}} d^2 R\, \overline{f}(\mathbf{R}) p_{\text{tot}}(\mathbf{r}^{'},\ \mathbf{R}).
\tag{A12}
$$

The mean image array is

$$
\bar{g} = [D H_1\, \overline{f}].
\tag{A13}
$$

APPENDIX B: COVARIANCE MATRIX

The covariance matrix $\mathbf{K_g}$ can be deduced from the autocorrelation matrix $\mathbf{R_g}$ and the mean image vector $\bar{g}$:

$$
\mathbf{K_g} = \mathbf{R_g} - \langle g \rangle \langle g \rangle^t.
\tag{B1}
$$

With the help of the operator D, we can also obtain the autocorrelation matrix $\mathbf{R_g}$ out of the autocorrelation function $R_g(\mathbf{r}'_1, \mathbf{r}'_2)$:

$$
\begin{aligned}
\mathbf{R_g} = {}& \left\langle g g^t \right\rangle \\
= {}& \left\langle \iint d^2 r_1^{'} y_g(\mathbf{r_1}) h(\mathbf{r_1}) \iint d^2 r_2^{'} y_g(\mathbf{r_2}) h^t(\mathbf{r_2}) \right\rangle \\
= {}& \iiiint d^2 r_1^{'} d^2 r_2^{'} h(\mathbf{r_1}) h^t(\mathbf{r_2}) \left\langle y_g(\mathbf{r_1}) y_g(\mathbf{r_2}) \right\rangle \\
= {}& [D R_g(\mathbf{r_1^{'}}, \mathbf{r_2^{'}}) D^\dagger].
\end{aligned}
\tag{B2}
$$

The exact expression for the autocorrelation function is

$$
R_g(r'_1, r'_2) = \left\langle \sum_{n_1} \sum_{k_1} \beta_{n_1 k_1} \delta(r'_1 - R_{n_1} - \Delta r_{n_1 k_1} \right.
$$

$$
- \Delta r'_{n_1 k_1}) \sum_{n_2} \sum_{k_2} \beta_{n_2 k_2} \delta \tag{B3}
$$

$$
\left. \times (r'_2 - R_{n_2} - \Delta r_{n_2 k_2} - \Delta r'_{n_2 k_2}) \right\rangle.
$$

In the double sum over n_1 and n_2, there are N terms with $n_1 = n_2$ and $N^2 - N$ terms with $n_1 = n_2$. For $n_1 = n_2$, there are also K_n terms with $k_1 = k_2$ and $K_n^2 - K_n$ terms with $k_1 \neq k_2$. For $n_1 \neq n_2$, it is irrelevant whether $k_1 = k_2$. Hence there are three cases to consider:

Case 1: $n_1 = n_2$ and $k_1 = k_2$. The calculation parallels the derivation of the mean image array. The contribution from this part is

$$
[R_g(r'_1, r'_2)]_1 = [H_1 b](r'_2) \delta(r'_1 - r'_2). \tag{B4}
$$

Case 2: $n_1 = n_2$ and $k_1 \neq k_2$. In this case, $\Delta r'_{nk_1}$ and $\Delta r'_{nk_2}$ are independent, as are Δr_{nk_1} and Δr_{nk_2}. The result from step (3) in Appendix A is

$$
[R_g(r'_1, r'_2)]_2 = \sum_n \frac{K_n^2 - K_n}{\langle K_n \rangle^2} p_{tot} \tag{B5}
$$

$$
\times (r'_1, R_n | z_n) p_{tot}(r'_2, R_n | z_n).
$$

Averaging over K_n in the next step requires $\langle K_n(R_n, z_n) \rangle$ and $\langle K_n^2(R_n, z_n) \rangle$. Although we do not know the PDF of K_n, we can still go a little further because K_n is from the binomial selection of $m(R_n, z_n)$, which is the total number of visible photons generated by the nth x-ray photon. The probability of a visible photon to come out of the screen is denoted by $A(R_n, z_n)$:

$$
\Pr_{K_n}(K_n) = \begin{bmatrix} m \\ K_n \end{bmatrix} A^{K_n} (1 - A)^{m - K_n}, \tag{B6}
$$

$$
\overline{K_n} = mA, \tag{B7}
$$

$$
\langle K_n^2 | m \rangle = mA(1 - A) + m^2 A^2, \tag{B8}
$$

$$
\langle K_n^2 \rangle = mA(1 - A) + \overline{m^2} A^2, \tag{B9}
$$

$$
Q(R_n, z_n) \equiv \frac{\overline{K_n^2} - \overline{K_n}}{\overline{K_n}^2},
$$

$$
= \frac{mA(1 - A) + \overline{m^2} A^2 - mA}{m^2 A^2}, \tag{B10}
$$

$$
= \frac{\overline{m^2} - m}{m^2}.
$$

With the function $Q(R_n, z_n)$ introduced above, the result from step (4) in Appendix A is

$$[R_g(r_1', r_2')]_2 = \sum_n \frac{\overline{K_n^2} - \overline{K_n}}{\overline{K_n}^2} P_{tot}$$

$$\times (r_1', R_n | z_n) P_{tot}(r_2', R_n | z_n)$$

$$= \sum_n Q(R_n, z_n) P_{tot}(r_1', R_n | z_n) P_{tot}$$

$$\times (r_2', R_n | z_n). \tag{B11}$$

The remaining steps now yield

$$[R_g(r_1', r_2')]_2 \iint_{supp} d^2 R \, \mathcal{F}(R) \int_0^d dz \frac{1}{d_x}$$

$$\times \exp(-z / d_x) Q(R, z)$$

$$\times P_{tot}(r_1', R | z) P_{tot}(r_2', R | z). \tag{B12}$$

We can express this term more compactly by defining a linear operator H_2 :

$$[H_2 \mathcal{F}](r_1', r_2') = \iint_{supp} d^2 R \, \mathcal{F}(R) \int_0^d dz \frac{1}{d_x}$$

$$\times \exp(-z / d_x) Q(R, z) P_{tot}$$

$$\times (r_1', R | z) P_{tot}(r_2', R | z). \tag{B13}$$

This part of the autocorrelation function is therefore expressed as

$$[R_g(r_1', r_2')]_2 = [H_2 \mathcal{F}](r_1', r_2'). \tag{B14}$$

Case 3: $n_1 \neq n_2$. If $n_1 \neq n_2$, and considering the independence between $\Delta r'_{n1,k1}$ and $\Delta r'_{n2,k2}$, there is no correlation between R_{n1} and R_{n2} except possible randomness induced by $f(R)$. For any fixed f, step (6) in Appendix A gives

$$[R_g(r_1', r_2')]_3 \frac{N^2(f) - N(f)}{N^2(f)} [H_1 f](r_1')[H_1 f](r_2'). \tag{B15}$$

Since N is a Poisson random variable at the fixed fluence f, we know that

$$\left\langle N^2 - N | b \right\rangle = \mathrm{Var}(N | f) + N^2(f) - N(f) = N^2(f). \tag{B16}$$

From step (7) of Appendix A, we have

$$[R_g(r_1', r_2')]_3 = [H_1 f](r_1')[H_1 f](r_2'). \tag{B17}$$

After step (8) in Appendix A, this case yields

$$[R_g(r_1', r_2')]_3 = [H_1 R_f H_1^\dagger](r_1', r_2'). \tag{B18}$$

By adding the Eqs. (B4), (B12), and (B18) together, we find that the autocorrelation function is

$$R_g(r_1', r_2') = [H_1 \bar{f}](r_2') \delta(r_1' - r_2') + [H_2 \bar{f}](r_1', r_2')$$
$$+ [H_1 R_f H_1^\dagger](r_1', r_2').$$

(B18)

Because the operator H_1 is linear, the covariance function is given as

$$K_g(r_1', r_2') = R_g(r_1', r_2') - \langle y_g(r_1') \rangle \langle y_g(r_2') \rangle$$
$$= [H_1 \bar{f}](r_2') \delta(r_1' - r_2') + [H_2 \bar{f}](r_1', r_2')$$
$$+ [H_1 K_f H_1^\dagger](r_1', r_2').$$

(B20)

APPENDIX C: COVARIANCE FUNCTION DECOMPOSITION

If the stationary covariance function can be decomposed into $K(\mathbf{R}) = \iint d^2 r f(\mathbf{R} + \mathbf{r}) f(\mathbf{r})^*$, the Fourier transform is a positive function defined as $[F_2 K](\rho) = \|F(\rho)\|$, where F_2 is the 2D Fourier-transform operator and $F(\rho)$ is the Fourier transform of $f(\mathbf{R})$. The amplitude of the Fourier transform of the decomposition function is the square root of the Fourier transform of the covariance function.

The covariance function is[18]

$$K(\mathbf{R}_1 - \mathbf{R}_2)$$
$$= \frac{\bar{N}\bar{K}}{A}[R_b(\mathbf{R}_1 - \mathbf{R}_2) + \bar{K} R_s(\mathbf{R} - 1 - \mathbf{R}_2)],$$

(C1)

$$b(\mathbf{R}, R_\theta) = b(R, \varphi, R_\theta)$$
$$= \exp\left\{ - aR^\beta \right.$$
$$\times \left. \frac{[L_y^2 \cos^2(\theta + \varphi) + L_x^2 \sin^2(\theta + \varphi)]^{1/2}}{L_x L_y} \right\},$$

(C2)

where b is the blob function, R_b is the autocorrelation function of the blob averaged over all the orientations,

$$R_b(\mathbf{R}_1 - \mathbf{R}_2) = \frac{1}{2\pi} \int_0^{2\pi} d\theta \iint_A d^2 R b(\mathbf{R}_1 - \mathbf{R}, R_\theta) b$$
$$\times (\mathbf{R}_2 - \mathbf{R}, R_\theta),$$

(C3)

and R_s is the autocorrelation function of S_θ averaged over all the orientations, given by

$$R_s = \frac{1}{2\pi} \int_0^{2\pi} d\theta \iint_A d^2 R S_\theta(\mathbf{R}_1 - \mathbf{R}, R_\theta) S_\theta$$
$$\times (\mathbf{R}_2 - \mathbf{R}, R_\theta),$$

(C4)

$$S_\theta(\mathbf{R}) = \iint d^2 R' \Phi(\mathbf{R}') b(\mathbf{R} - \mathbf{R}', R_\theta),$$

(C5)

where S_θ is the convolution of the cluster PDF Φ and the blob function, where the cluster PDF is the Gaussian with the standard deviation σ_φ. The constant $\bar{N}$ is the mean number of clusters within the field of view, $\bar{K}$ is the mean number of blobs per cluster, and A is the area of the field of view.

The Fourier transform of R_b results in the absolute squared Fourier transform of the blob function averaged over all the orientations:

$$[F_2 R_b](\rho) = \frac{1}{2\pi} \int_0^{2\pi} d\theta \left| \left| [F_2 b](\rho, \theta) \right| \right|^2 . \tag{C6}$$

The Fourier transform of R_s is similarly given as

$$[F_2 R_s](\rho) = \frac{1}{2\pi} \int_0^{2\pi} d\theta \left| \left| [F_2 S_\theta](\rho, \theta) \right| \right|^2 , \tag{C7}$$

and the Fourier transform of S_θ is the product of the Fourier transform of the blob function and the Gaussian φ:

$$[F_2 S_\theta](\rho, \theta) = \exp(-2\pi^2 \sigma_\varphi^2 \rho^2)[F_2 b](\rho, \theta). \tag{C8}$$

The Fourier transform of the covariance function is given as

$$[F_2 K_b](\rho) = \frac{\bar{N} \bar{K}}{A}[1 + \bar{K}$$
$$\times \ \exp(-4\pi^2 \sigma_\varphi^2 \rho^2)]\frac{1}{2\pi} \int_0^{2\pi} d\theta \left| \left| F_2 b \right| (\rho, \theta) \right| \right|^2 . \tag{C9}$$

The amplitude of the Fourier transform of the decomposition function is the square root of that of the covariance function:

$$\left| F(\rho) \right| = \left\{ \frac{\bar{N} \bar{K}}{A}[1 + \bar{K} \ \exp(-4\pi^2 \sigma_\varphi^2 \rho^2)] \right\}^{1/2}$$
$$\times \left[\left\langle \left| \left| F_2 b(\rho, \theta) \right| \right|^2 \right\rangle_\theta \right]^{1/2}, \tag{C10}$$

where $\left\langle \| F_2 b(\rho, \theta) \|^2 \right\rangle_\theta$ is the averaged absolute square of the Fourier transform of the blob function.

The Fourier transform of the blob function can be simplified as follows:

$$[F_2 b](\rho, \theta)$$
$$= [F_2 b](\rho, \varphi', \theta)$$
$$= \int_0^\infty \int_0^{2\pi} dr d\varphi r b(r, \varphi, \theta)$$
$$\times \ \exp[-j 2\pi \rho r \cos(\varphi' - \varphi)]$$
$$= \int_0^\infty \int_0^{2\pi} dr d\varphi r$$
$$\times \ \exp\left\{ -\alpha r \beta \frac{[L_y^2 \cos^2(\theta + \varphi) + L_x^2 \sin^2(\theta + \varphi)]^{1/2}}{L_x L_y} \right\}$$
$$\times \ \exp[-j 2\pi \rho r \cos(\varphi' - \varphi)]$$
$$= \int_0^\infty \int_0^{2\pi + \theta} dr d\varphi r$$
$$\times \ \exp\left\{ -\alpha r \beta \frac{[L_y^2 \cos^2(\varphi) + L_x^2 \sin^2(\varphi)]^{1/2}}{L_x L_y} \right\}$$
$$\times \ \exp[-j 2\pi \rho r \cos(\varphi' + \theta - \varphi)]. \tag{C11}$$

Because the spatial function to be transformed and the Fourier-transform kernel are both periodic functions of the variable φ with 2π as one period, the integral over one period is independent of the starting point of the integration interval. The interval in Eq. (C11) can be changed into $(0, 2\pi)$:

$$
\begin{aligned}
&[F_2 b](\rho, \ \varphi', \ \theta) \\
&= \int_0^\infty \int_0^{2\pi} \mathrm{d}r\,\mathrm{d}\varphi\, r \\
&\times \ \exp\left\{-ar\beta \frac{[L_y^2 \cos^2(\varphi) + L_x^2 \sin^2(\varphi)]^{1/2}}{L_x L_y}\right\} \\
&\times \ \exp[-j2\pi\rho r \cos(\varphi' + \theta - \varphi)] \\
&= [F_2 b](\rho, \ \varphi + \theta, \ 0).
\end{aligned}
$$

(C12)

The Fourier transform of the blob function is still a periodic function of θ with a period of 2π; thus

$$
\begin{aligned}
\left\langle |\,|\,F_2 b(\rho, \ \theta)\,|\,|^2 \right\rangle_\theta &= \left\langle |\,|\,F_2 b(\rho, \ \varphi + \theta, \ 0)\,|\,|^2 \right\rangle_\theta \\
&= \left\langle |\,|\,F_2 b(\rho, \ \theta, \ 0)\,|\,|^2 \right\rangle_\theta.
\end{aligned}
$$

(C13)

The absolute square of the transform averaged over all the orientations is a radially symmetric function at last. For a radially symmetric function, the inverse Fourier transform is actually the Hankel transform, and the corresponding spatial function is a 2D real function. On the basis of this fact, the phase of the Fourier transform of the decomposition function is uniformly 0. The function can be calculated by the inverse Fourier transform of the amplitude given above:

$$
f(\boldsymbol{R}) = F_2^{-1}\left(\left\{\frac{NK}{A}[1 + K \exp(-4\pi^2 \sigma_\varphi^2 \rho^2)]\right\}^{1/2} \left[\left\langle |\,|\,F_2 b(\rho, \ \theta, \ 0)\,|\,|^2 \right\rangle_\theta\right]^{1/2}\right).
$$

(C14)

The Fourier and inverse Fourier transform were performed for this study by the fast-Fourier-transform algorithm on a computer.
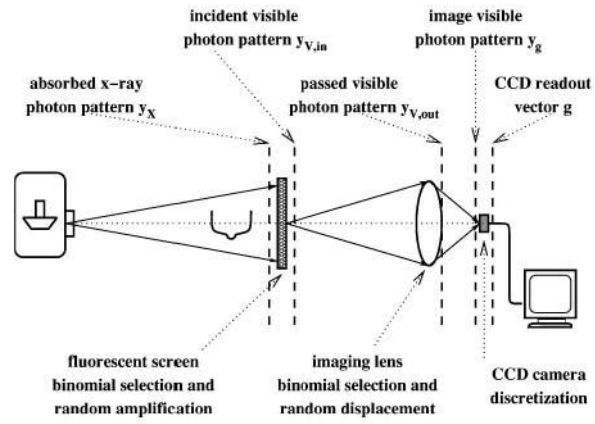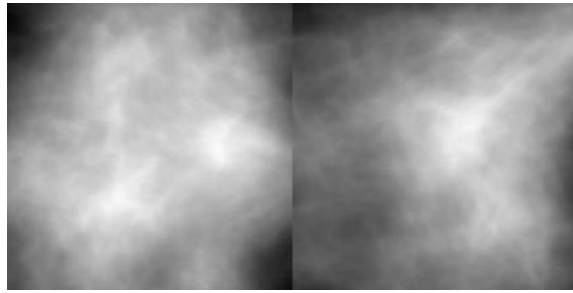
Fig. 1.
Layout of the model system.

Fig. 2.
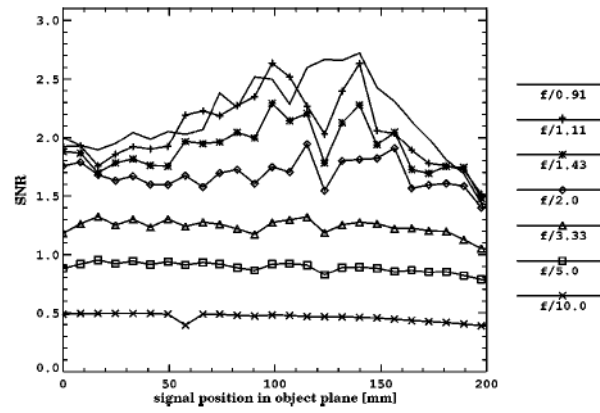Two samples of the clustered lumpy background.

Fig. 3.
SNR plots at different aperture sizes. The signal position is measured from the optical axis on the entrance surface of the screen. The image plane position is unchanged in all cases. The aperture size is in terms of the numerical aperture in object space.
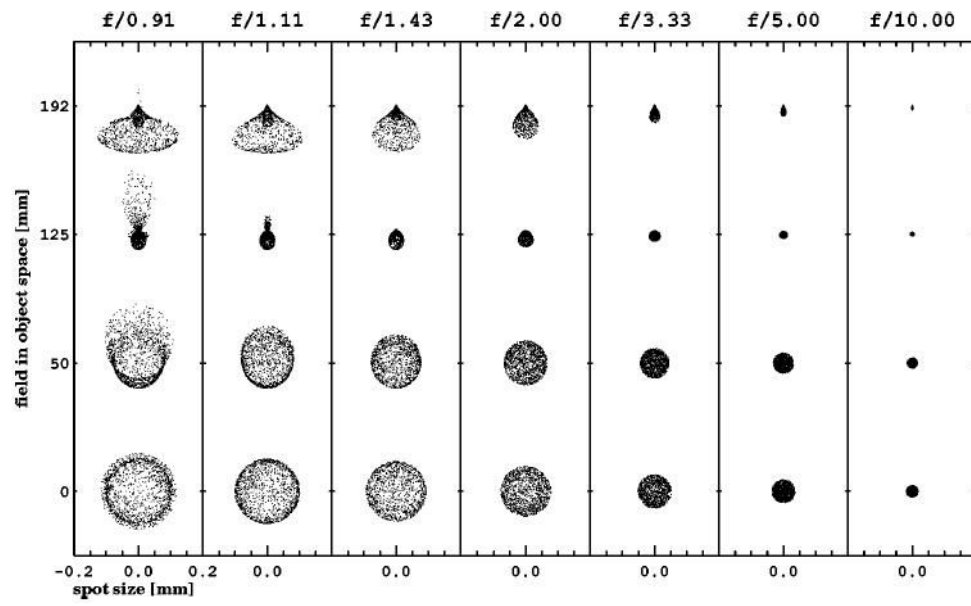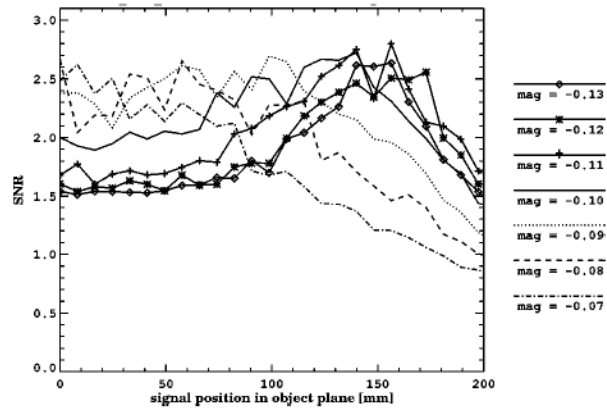
Fig. 4.
Spot diagrams at four fields of different aperture sizes. The field position is in terms of the radial distance from the optical axis on the image plane.

Fig. 5.
SNR plots at different working distances. The signal position is measured from the optical axis on the entrance surface of the screen. The aperture size is unchanged in all cases, with a numerical aperture of 0.055. The working distance is in terms of the magnification.
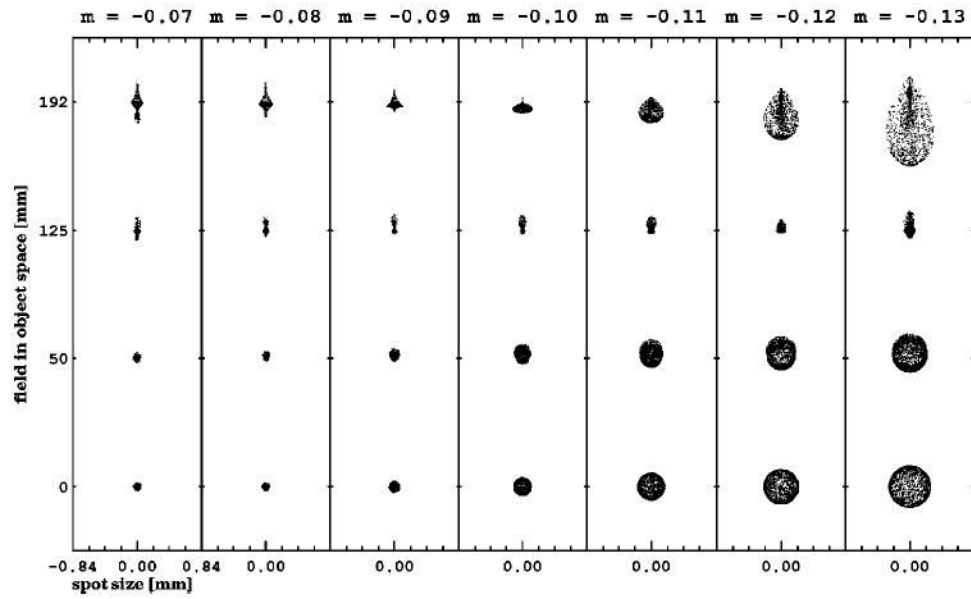
Fig. 6.
Spot diagrams at four fields of different working distances. The field position is in terms of the radial distance from the optical axis on the image plane.
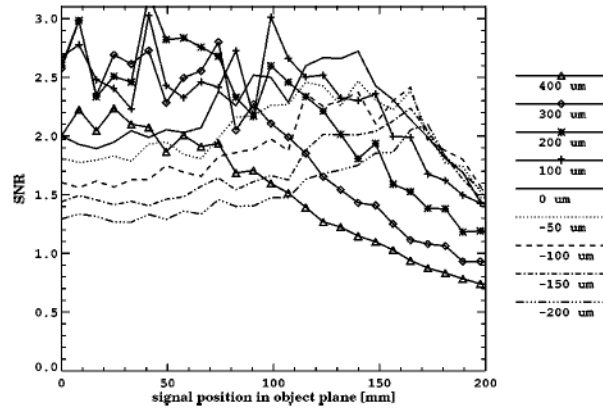
Fig. 7.
SNR plots at different image plane positions. The signal position is measured from the optical axis on the entrance surface of the screen. The aperture size is unchanged in all cases. The defocus is in terms of the distance from the ideal image plane in image space.
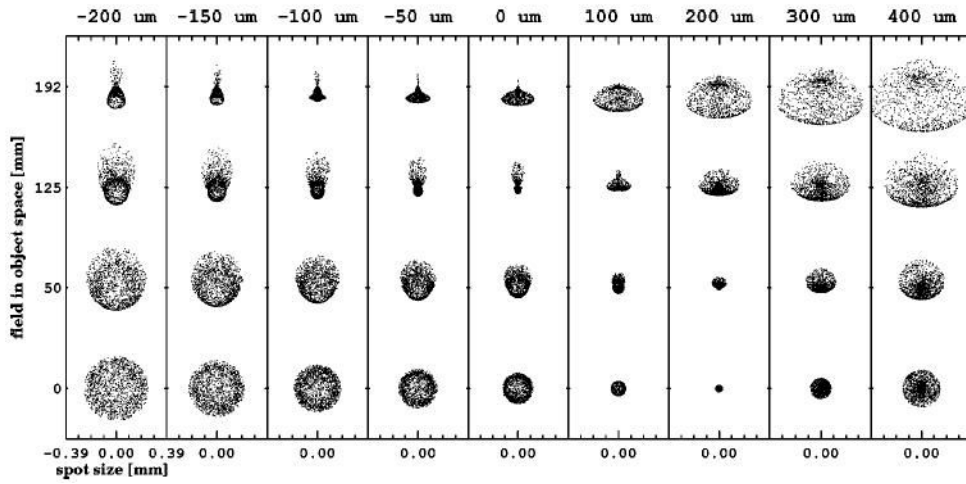
Fig. 8.
Spot diagrams at four fields of different image plane positions. The field position is in terms of the radial distance from the optical axis on the image plane.