Task-based optimization of dedicated breast CT via Hotelling observer metrics

Adrian A. Sanchez,^{a)} Emil Y. Sidky,^{b)} and Xiaochuan Pan^{c)}

Department of Radiology, The University of Chicago, Chicago, Illinois 60637

(Received 16 April 2014; revised 21 July 2014; accepted for publication 5 September 2014; published 1 October 2014)

Purpose: The purpose of this work is to develop and demonstrate a set of practical metrics for CT systems optimization. These metrics, based on the Hotelling observer (HO) figure of merit, are task-based. The authors therefore take the specific example of optimizing a dedicated breast CT system, including the reconstruction algorithm, for two relevant tasks, signal detection and Rayleigh discrimination.

Methods: A dedicated breast CT system is simulated using specifications in the literature from an existing prototype. The authors optimize configuration and image reconstruction algorithm parameters for two tasks: the detection of simulated microcalcifications and the discrimination of two adjacent, high-contrast signals, known as the Rayleigh discrimination task. The effects on task performance of breast diameter, signal location, image grid size, projection view number, and reconstruction filter were all investigated. Two HO metrics were evaluated: the percentage of correct decisions in a two-alternative forced choice experiment (equivalent to area under the ROC curve or AUC), and the HO efficiency, defined as the squared ratio of HO signal-to-noise ratio (SNR) in the reconstructed image to HO SNR in the projection data.

Results: The ease and efficiency of the HO metric computation allows a rapid high-resolution survey of many system parameters. Optimization of a range of system parameters using the HO results in images that subjectively appear optimal for the tasks investigated. Further, the results of assessment through the HO reproduce closely many existing results in the literature regarding the impact of parameter selection on image quality.

Conclusions: This study demonstrates the utility of a task-based approach to system design, evaluation, and optimization. The methodology presented is equally applicable to determining the impact of a wide range of factors, including patient parameters, system and acquisition design, and the reconstruction algorithm. The results demonstrate the versatility of the proposed HO formalism by not only generating a set of parameters that are optimal for a given task but also by qualitatively reproducing many existing results from the breast CT literature. Meanwhile, the implementation of the proposed methodology is straightforward and entirely simulation-based. This is an attractive feature for many system optimization problems, where the goal is to analyze the individual system components such as the image reconstruction algorithm. Final assessment of the system as a whole should be based also on real data studies. © 2014 American Association of Physicists in Medicine. [<http://dx.doi.org/10.1118/1.4896099>]

Key words: Hotelling observer, image quality, breast CT, image reconstruction

1. INTRODUCTION

Quantification of image quality in x-ray computed tomography (CT) is useful not only in evaluating a completed imaging system but also in optimizing decisions in the system design and implementation. A wide array of parameters exists in the CT imaging chain, including various physical parameters such as patient size, acquisition parameters such as the number of projection views, as well as reconstruction algorithm parameters such as image pixel size. Each of these parameters has an effect on the noise properties of the CT images, as well as the resolution in the image, however, the impact of these factors on the utility of the images produced is often unclear. Further, for parameters which are controllable, such as acquisition and reconstruction parameters, the optimal selection of implementation choices is rarely obvious. We adopt a task-based

approach to system design and image quality evaluation via the Hotelling observer (HO).¹⁻⁴ In order to demonstrate the optimization of a CT system via the HO, a specific system must be investigated, along with relevant tasks. We have chosen to optimize a dedicated breast CT system, and demonstrate the use of the HO to quantify breast CT's image quality for the relevant task of microcalcification detection. Our approach differs from current standard methods in that the resulting metrics are nonstochastic, and we impose no restrictions on the structure of the signal or on the nature of the noise correlations. Instead, we only assume that the relevant noise correlations and signal extent can be restricted to a region-of-interest (ROI) in the image.

The particular HO implementation we propose is exact in that it does not rely on assumptions such as shift-invariance or noise stationarity. These assumptions are commonly invoked,

at least locally, because they enable the construction of image quality metrics based on discrete Fourier transform domain image properties, resulting in familiar metrics such as the noise power spectrum (NPS)⁵⁻⁸ and modulation transfer function (MTF)^{5,7,9,10} or local NPSs (Refs. 11 and 12) and MTFs.¹³ Instead, the approach we propose is not limited by assumptions regarding stationarity, but instead by the assumption that relevant information for HO assessment is restricted to a ROI in the image, and that the reconstruction operation is linear. Meanwhile, the propagation of the signal and noise into the image domain is fully and exactly accounted for through a linear systems theory approach. While the assessment performed here makes an additional assumption of a fixed (nonrandom) background, the framework described is sufficiently general that additional terms could be added to account for anatomical noise covariance when the task of interest is limited by anatomical noise. Likewise, performance-degrading psychophysical factors could be modeled if direct estimation of human performance were of interest, however, in this work, we are concerned instead with determining ideal observer performance.

In the present study, we focus on optimization of dedicated breast CT, which is currently being investigated as a potential screening and diagnostic tool for breast cancer at several institutions.¹⁴⁻¹⁸ Breast CT is an attractive technology, largely because it has the potential to provide improved visualization of breast lesions in cases where conventional mammography lacks sensitivity.^{15,19,20} Examples include young women (less than 50 yr old) and those with dense breast tissue, wherein the overlay of anatomic breast structures in a mammographic projection image can obscure lesions and impede diagnostic accuracy.²¹

While breast CT can potentially improve the utility of breast screening relative to mammography, the individual projection views of the CT acquisition are limited in terms of x-ray fluence due to dose concerns, resulting in considerable quantum noise in the projection data. Further, dedicated breast CT cannot currently match the “in plane” resolution of full-field digital mammography,^{14,22} leading to potential shortcomings in terms of microcalcification detection or visualization of tumor margins, two tasks which can help to indicate malignancy of breast lesions.

In this work, we demonstrate the impact of several system and object parameters on image quality. First, we investigate the effects of patient breast diameter for both fixed total exposure and fixed mean glandular dose. Since patient breast diameter cannot be controlled, the impact of breast size on task performance is important to ascertain. We therefore demonstrate that our proposed formalism is capable of qualitatively reproducing previous findings regarding the impact of breast size on image quality. Next, we investigate the impact of signal location within the breast, which is important since it leads to the conclusion that accounting for location variability can meaningfully impact the outcome of the image quality assessment. In order to demonstrate the use of HO metrics to address a question of system parameters, we next evaluate the effects of varying the number of projection views while keeping the total radiation exposure fixed. The trade-off between noise in the individual projection views and adequate angular

sampling provides a meaningful evaluation of the HO metrics’ ability to optimize acquisition parameters. Finally, since the reconstruction algorithm itself connotes some potential loss in information, we demonstrate optimization of two of the most ubiquitous parameters in the conventional FBP algorithm: the image pixel size and the reconstruction filter width. Image quality evaluation in each case was performed with respect to two tasks: the detection of a simulated microcalcification and the resolution of two adjacent, high-contrast objects. We choose these tasks because the trade-off between image noise and resolution is quite sensitive to CT system parameters.

The noise/resolution trade-off also depends strongly on the noise model covariance and the form of object being viewed in the detection/discrimination task. The HO combines this complex image quality information into a single meaningful metric for a given task. In order to illustrate this advantage of the HO, we present several realizations of reconstructed noisy images for subjective assessment. We include several examples of images created with reconstruction filters and view numbers deemed optimal by the HO. Also shown are images obtained from view numbers or reconstruction filter widths deemed suboptimal by the HO. The numerical phantoms used to create these images do not correspond exactly to the tasks for which these parameters were optimized, but instead are either anthropomorphic breast phantoms or phantoms intended to mimic clinical quality assurance phantoms. This comparison is important because it illustrates the extent to which these two approaches to image quality assessment agree, demonstrates how such images provide a useful check for HO optimization, and highlights specific limitations of each approach.

The outline of the paper is as follows: Sec. 2 provides background in image quality evaluation in dedicated breast CT, along with motivation for the evaluation formalism we propose. In Secs. 3.A and 3.B, we describe the simulated breast CT system and numerical phantoms. In Secs. 3.C and 3.D, we describe the data and noise models in the projection data and image domains, while in Secs. 3.E and 3.F, we outline the task-based assessment formalism. Next, we describe the application of the formalism to ascertaining the impact of a range of imaging parameters, some of which are inherent to the breast being imaged, Sec. 3.F.1, and others which are controllable through system design or reconstruction algorithm implementation, Secs. 3.F.2 and 3.F.3. Results of the application of the methodology are shown in Sec. 4, with corresponding discussion in Sec. 5. Examples of reconstructed images generated using the determined optimal system parameters are provided in Sec. 6 and discussed in Sec. 7. Finally, we provide conclusions in Sec. 8.

2. BACKGROUND

Since in this work we have decided to focus on dedicated breast CT, we shall focus our discussion of CT image quality metrics to those which have most commonly been applied to that modality. A variety of approaches have been adopted to aid in image quality evaluation for dedicated breast CT, some of which have been applied to the optimization of the

physical system, technique factors, or algorithm implementation. Boone *et al.*¹⁴ performed image quality evaluation by computing contrast-to-noise (CNR) and corresponding signal-to-noise ratio (SNR) values for lesions in breast cadavers. Comparison of these values to the Rose criterion²³ was then performed to establish feasibility of low-dose dedicated breast CT. A similar approach was also used more recently by Shen *et al.*²⁴ in order to ascertain the impact of projection view number with fixed total exposure.

Meanwhile, Refs. 25–27 explored the impact of an extensive range of imaging parameters on image resolution (Refs. 25 and 26) and noise (Ref. 27) by computing Fourier-based metrics. Specifically, MTFs and NPSs were measured experimentally by repeated imaging of phantoms with varied imaging parameters. In the case of resolution measurement, a high-contrast wire was used to mimic a point source, so that empirical point-spread functions could be seen in the reconstructed images. Meanwhile, uniform cylindrical polyethylene phantoms were used to generate mean-subtracted noise images, whose discrete Fourier transform yielded empirical NPSs. These metrics have the benefit of ubiquity across many medical imaging applications.

A task-based approach to breast CT image quality evaluation was pursued by Lindfors *et al.* and Lai *et al.* in Refs. 22 and 28, respectively. The former study used human observers to subjectively classify a variety of lesion types as visible or not visible, for the purpose of comparing lesion detectability in dedicated breast CT to screen-film mammography. The authors allude to performing human observer ROC studies, which indeed would provide a very meaningful metric for comparison of these two modalities, but would be too costly and time-consuming to perform investigation of the impact of many system parameters in isolation. Meanwhile, Lai *et al.*²⁸ also performed a subjective image quality evaluation by having human observers count the number of microcalcifications in various microcalcification clusters which were clearly visible. The clusters were in known locations and had fixed, known locations for individual microcalcifications. The purpose of this study was to investigate the effects of x-ray tube voltage and radiation dose.

Moving toward methods which are more robust against the subjectivity of human observer responses, the results of Refs. 17 and 19 are particularly significant in that they help to establish the capability of dedicated breast CT in microcalcification detection through assessment of human performance in the form of full human ROC studies.²⁹ Meanwhile, the methods proposed in Ref. 30 overcome some of the limitations of using humans by computing ideal observer performance to study the impact of x-ray spectral shape. The authors' approach in that work is similar to that which we propose, however the assumptions upon which each formalism is based are different. Glick *et al.*³⁰ apply a parallel cascade approach to studying the signal and noise properties of dedicated breast CT by modeling it as a linear shift-invariant (LSIV) system with stationary Gaussian noise (both anatomical and quantum). They then analyze lesion detectability in the projection domain, rather than optimizing with respect to the reconstructed image.

As an alternative to simulation of the breast CT system, Packard *et al.* used sample images to obtain mean signal profiles and breast power spectra, in turn constructing a mathematical model observer, the prewhitening matched filter, based on these quantities derived from samples.³¹ The goal of that work was to ascertain the effect of slice thickness on lesion detectability. More recently, Chen *et al.*³² have used similar methods to investigate an association between lesion detectability in breast imaging and spectral components of anatomic noise.

We propose an assessment methodology based on the HO. By considering the HO's performance on a relevant task within a small ROI, summary metrics can be computed quickly enough to sweep over a range of imaging parameters and evaluate their impact on image quality. This approach has the benefit of being task-based and nonstochastic, meaning that it does not rely on a collection of noisy images or on the assessment of human performance. Further, the methodology we propose does not assume a LSIV system or stationary noise. Instead of making assumptions regarding the noise and signal properties in the reconstructed images, our approach is based on assumptions of noise and signal properties in the original projection data, which, in the absence of object variability, are simpler to model accurately. We then exactly account for the action of the reconstruction operation on HO task performance, including the noise correlations and signal bias which reconstruction introduces. The details of the noise model assumed in this work are given in Sec. 3.D, and the resulting scalar image quality metrics are described in Sec. 3.F.

3. METHODS

3.A. Imaging system geometry

The simulated CT system in this work was modeled based on that from Refs. 25 and 27. Those studies investigated noise and resolution properties, respectively, of the same dedicated breast CT system. This system utilizes a pendant geometry and a flat-panel CsI detector to acquire cone-beam CT data of the breast. We chose to model that system here in order to facilitate a meaningful comparison of our task-based assessment approach to more standard approaches, namely, performing measurements of the MTF and NPS, pursued by the authors of those works.

The chosen detection/discrimination tasks are performed viewing a single slice of the reconstructed volume, and in order to simplify the HO presentation, we restrict the observer model computation to the midplane of the circular cone-beam scanning configuration. As a result, we consider only the 2D circular fan-beam configuration for the midplane. The axial position of the midplane for the simulated detector is approximately 2 cm below the top edge of the detector face. The simulated detector had a width of approximately 40 cm and consisted of 2048 detector pixels, each 0.194 mm in width. The distance from the x-ray source to the axis of rotation is 458 mm, and the distance from the x-ray source to the detector is 878 mm. The resulting geometric magnification for an object at the axis of rotation is therefore 1.92. The system's field-of-view (FOV) is restricted to 165.9 cm in diameter, and this circular FOV

is inscribed in a 1024×1024 image pixel grid, with a corresponding image pixel size of 0.162×0.162 mm. Meanwhile, the detector pixel size, back-projected to the axis of rotation, is 0.101 mm. Finally, a 0.4 mm focal spot is modeled, as described in Sec. 3.C.

3.B. Breast phantom

The digital breast phantom used in this work is circular with a diameter of 14 cm. The phantom composition is modeled as being 50% adipose tissue and 50% fibroglandular tissue, uniformly distributed throughout the breast (see Fig. 2 for ROIs from the phantom). The corresponding attenuation values are taken from Ref. 33. Simulated microcalcifications are modeled as being circularly symmetric Gaussians of peak x-ray attenuation equal to that of calcium. However, the attenuation value is decreased in order to simulate the first-order partial volume averaging in a finite slice thickness of 0.5 mm containing a microcalcification.

Whenever the simulated microcalcification has a diameter less than the slice thickness of 0.5 mm, the modified x-ray attenuation for the microcalcification is computed as

$$\mu_{\text{calc}} = \frac{d_{\text{calc}}}{0.5 \text{ mm}} (\mu_{\text{Ca}} - \mu_{\text{br}}) + \mu_{\text{br}}, \quad (1)$$

where d_{calc} is the diameter of the calcification, defined as the full width at half maximum of the Gaussian signal, μ_{Ca} is the x-ray attenuation of calcium, and μ_{br} is the attenuation of the 50/50 breast phantom.

3.C. Projection data model

We model the CT data after the point where the transmitted x-ray intensity is processed by the negative logarithm, using the standard line integral formulation as the mean

$$g_i = \int_L f(s_i + l\hat{\theta}_i) dl, \quad (2)$$

where g_i is the sinogram data corresponding to the i th detector element, and f is the numerical phantom value at a point which is l distance from the source position s_i in the direction $\hat{\theta}_i$. The line integral is performed along all values within the compact object support in the direction $\hat{\theta}$, and projection through the discrete numerical phantoms is performed by considering the path-length of a given ray through a square pixel. The pixel size of the numerical phantoms is always set to be at least an order of magnitude smaller than the pixel size of the final reconstructed image.

The simulated x-ray spectrum corresponds to a 80 kV(peak) setting with added Be and Al filtration of 0.8 and 2.5 mm, respectively. Methods from Refs. 34–36 were used in simulation of the x-ray spectrum. In order to model the finite size of the detector pixels, we perform 16-fold subsampling, so that for each detector pixel, 16 equally spaced rays within that pixel are simulated. We do not perform subsampling along the trajectory of the x-ray source, modeling a step-and-shoot acquisition. A 0.4 mm focal spot is modeled by convolution of a rect function with the subsampled line integral data. The width of the convolution kernel (rect function) includes geometric magnification

and is computed as

$$d = d_{\text{fspot}} \frac{\text{SDD} - r}{r}, \quad (3)$$

where d_{fspot} is the width of the focal spot (0.4 mm here), SDD is the source-to-detector distance (878 mm), and r is the angle-dependent distance from the focal spot to the center of a ROI containing the signal of interest. For the present model, we do not include other physical factors such as x-ray scatter and spread of optical photons in the CsI detector. Although inclusion of these physical factors may alter the presented results, they do not alter the proposed assessment methodology outlined in Sec. 3.F.

3.D. Noise model

We view images and sinograms as 1D vectors where the index scans through the respective data structure in lexicographical order. We then consider additive Gaussian noise applied to the line integral data model of Eq. (2)

$$\mathbf{g}_i = \int_L f(s_i + l\hat{\theta}_i) dl + \mathbf{n}_i, \quad (4)$$

where bold font indicates a random variable. In particular, we consider the noise term $\mathbf{n}_i$ to be zero-mean and have variance determined by

$$\text{Var}\{\mathbf{n}_i\} = \frac{\sigma_e^2 + \text{Var}\{\mathbf{I}_i\}}{\bar{I}_i^2}, \quad (5)$$

where $\mathbf{I}_i$ is the number of optical quanta incident on the imaging array at location i and converted to electrons, $\bar{I}_i$ is the mean of $\mathbf{I}_i$, and σ_e^2 is the variance due to electronic noise. This noise model is a generalization of that put forward by Barrett and Swindell,⁵ with the addition of the electronic noise term. Following Gong *et al.*,¹⁹ we take $\sigma_e = 2200$, the optical collection efficiency (OCE) to be 0.64, the detector pixel fill factor to be 0.57, and the x-ray-to-optical conversion of the scintillator to be 58/keV. The OCE, fill factor, and optical conversion factor can all be combined to a single gain factor $\Omega = 0.64 \times 0.57 \times 58$, so that the x-ray spectrum incident on the detector can be related to $\mathbf{I}_i$ via

$$\mathbf{I}_i = \Omega \sum_{k=1}^{N_E} E_k \mathbf{I}_{i,k}^{\text{x-ray}}, \quad (6)$$

where E_k is the energy in keV of the k th energy bin, and $\mathbf{I}_{i,k}^{\text{x-ray}}$ is the incident number of x-rays at energy E_k in location i . We take each $\mathbf{I}_{i,k}^{\text{x-ray}}$ to be Poisson distributed, with $\text{Var}\{\mathbf{I}_{i,k}^{\text{x-ray}}\} = \bar{I}_{i,k}^{\text{x-ray}}$, where the mean x-ray count data $\bar{I}_{i,k}^{\text{x-ray}}$ can be obtained through noiseless simulation. The resulting model for the line integral data variance is then given by

$$(K_g)_{i,j} = \begin{cases} \frac{\sigma_e^2 + \Omega^2 \sum_k E_k^2 \bar{I}_{i,k}^{\text{x-ray}}}{\left(\Omega \sum_k E_k \bar{I}_{i,k}^{\text{x-ray}} \right)^2} & : i = j \\ 0 & : \text{else,} \end{cases} \quad (7)$$

where $(K_g)_{i,j} = \text{Cov}\{g_i, g_j\}$. In order to simulate a fluence which would result in the same mean glandular dose as two-view mammography, we approximate the total incident number of x-rays based on the results for the photon fluence at the axis of rotation given in Ref. 37, which range from approximately 3×10^7 to 4×10^8 photons/mm² for breast diameters between 10 and 16 cm. This photon fluence is equally distributed among each projection view.

While Eq. (7) describes the modeled additive noise in the projection data domain, the noise in the image domain is correlated by the reconstruction algorithm. In our case, we employ the FBP algorithm, whose action can be fully described by multiplication with a matrix A . The image covariance matrix is then given by (see Eq. 8.50 in Ref. 38)

$$K_y = AK_gA^T, \quad (8)$$

where A^T denotes the matrix transpose of the reconstruction operator (not the discretization of the continuous adjoint operator). Recall that the image and sinogram can be represented by 1D vectors. In this way, the image covariance and reconstruction operator can be coded as two-index matrices.

For large systems, where A cannot directly be computed and stored, A represents the action of all linear steps in the FBP algorithm and a separate implementation for A^T is needed. For the present studies, the ROI is small enough that A can be directly computed column-by-column and stored in computer memory. Depending on the ROI size used, the dimensions of A in this study range roughly from 16×400 to $7000 \times 80\,000$, with the number of columns varying because we consider only the shadow of the ROI projected onto the detector. With the actual matrix A available, A^T is trivial to compute and the covariance K_y is obtained directly by matrix multiplication. In general, K_y is rank-deficient and we use its Moore–Penrose inverse, $K_y^\dagger$. Computation of $K_y^\dagger$ can be obtained by singular value decomposition (SVD). With the SVD, the covariance is written as

$$K_y = U\Sigma V^T, \quad (9)$$

where U and V are orthogonal matrices, and for the present case, where K_y is always symmetric and positive semidefinite, $U = V$. The matrix Σ is diagonal containing the singular values of K_y . The Moore–Penrose pseudoinverse becomes

$$K_y^\dagger = V\Sigma^\dagger U^T, \quad (10)$$

where $\Sigma^\dagger$ is diagonal, containing the multiplicative inverse of the nonzero singular values and zero elements of Σ remain zero in $\Sigma^\dagger$.

For some configurations investigated, rank-deficiency enters into K_y through A , which can be rank-deficient when image pixel size is mismatched to the size of the detector pixels or through a filtering operation on the data. An example of the singular value spectrum for K_y is shown in Fig. 1 for a 5×5 mm² ROI located 5.7 cm from the center of the field of view. The plateau of small singular values is within numerical precision of 0.

FIG. 1. The singular value spectrum for the matrix K_y is shown for one combination of parameters investigated. The smallest singular values shown are within numerical precision of 0, and K_y is, therefore, rank-deficient.

3.E. Classification tasks

We focus on optimizing parameters with respect to two relevant tasks: the detection of microcalcifications and the resolution of two high-contrast signals in close proximity, known as the Rayleigh discrimination task.^{39–41} Both of these tasks address different forms of an image noise-resolution trade-off, and they involve single slice viewing. This latter point facilitates studies relating human observers to the HO, because observer studies can be conducted with single or pairs of images and there is no need to consider slice scrolling or volume rendering, which would be needed to present volume data in a human consumable format. Each of these tasks can be seen as a two-class classification task, where one attempts to classify an image as corresponding to one of two hypotheses. In the case of microcalcification detection, the two hypotheses correspond to “signal-absent” and “signal-present.” In the case of the Rayleigh task, the two hypotheses are “one elongated object”, namely a blurred line, or “two distinct objects.” The blurred points and line used in the Rayleigh discrimination task are assigned a peak attenuation value equal to that of calcium. The sizes of the Rayleigh task signals are made proportional to those described in Ref. 41. The numerical phantoms corresponding to each hypothesis are shown in Fig. 2 for each of the two tasks. The difference in scale between the two tasks is a result of the fact that the Rayleigh task must span several pixels so that the “trough” between the high-contrast signals is visible. Meanwhile, a detectable microcalcification can be much smaller.

3.F. Hotelling observer metrics

The HO is the optimal linear model observer in the limit of Gaussian noise, and it has been used for prediction of human observer task performance.^{42,43} Even in cases where the HO does not accurately predict human performance, it often still constitutes a useful upper bound on human task performance.^{3,4,44} In a classification task, the HO operates by computing a scalar value which is a linear combination of the image pixel values. It then uses this value to classify an image as belonging to one of two hypotheses, and it performs this classification in a way which is optimal with respect to all

FIG. 2. Top Row: The numerical phantoms corresponding to the two hypotheses for the signal detection task, namely, signal-absent and signal-present. Bottom Row: The two hypotheses for the Rayleigh task, “one object” and “two objects.” The width of each image is approximately 5 mm. Note that the simulated microcalcification is much smaller than the Rayleigh signals because the Rayleigh task must span several pixels in order to be feasible, whereas the microcalcification can be detected by the HO when it spans only one or two pixels. The display window is $[0, 0.122] \text{ mm}^{-1}$.

other linear observers. The scalar value used for classification is termed the Hotelling test statistic and is computed as

$$t = w_y^T \mathbf{y}, \quad (11)$$

where w_y is the optimal set of weights for the image pixels, known as the Hotelling template. The Hotelling template is in turn defined as

$$\bar{K}_y w_y = \Delta \bar{y}, \quad (12)$$

where $\bar{K}_y = A \bar{K}_g A^T$ is the average of the image covariance matrix under the two hypotheses, and $\Delta \bar{y}$ is the mean difference between images from the two hypotheses. The quantity $\Delta \bar{y}$ can be obtained by generating noiseless reconstructed images of

each of the two hypotheses. Since the detector noise is additive and zero-mean and the reconstruction operator is linear, these noiseless images correspond to the mean image under each hypothesis. Eq. (12) is solved via the aforementioned Moore–Penrose pseudoinverse for cases where K_y is rank-deficient. Otherwise, for large system models, Eq. (12) can be solved by LU decomposition followed by backsubstitution, and for very large system models, iterative methods such as conjugate gradients can be applied.

Because the signal sizes involved are small, we compute t by using a small ROI. Since we will consider a range of projection view numbers, some of which lead to angular undersampling artifacts far from the signal of interest, we restrict the ROI size based on the radius from the signal at which these artifacts begin to appear. We define the signal energy in the reconstructed image within an annulus of radius r and width $2\Delta r$, by

$$E_y(r) = \int_0^{2\pi} d\theta \int_{r-\Delta r}^{r+\Delta r} (\Delta \bar{y}(r', \theta))^2 r' dr'. \quad (13)$$

An example of the integrand of Eq. (13) with a corresponding plot of $E_y(r)$ is shown in Fig. 3 for a microcalcification signal. We determine the ROI size used for HO assessment by approximating the above integral at evenly spaced values of r and setting the ROI radius to the value of r which minimizes $E_y(r)$.

The HO defined within a ROI is the optimal linear observer if we assume that only the pixels within the ROI are known. Equivalence between the HO in a ROI and the HO for the full image depends on whether the ROI contains the full signal extent as well as relevant noise correlations. Regarding the former point, while undersampling artifacts are present in $\Delta \bar{y}$, we intentionally exclude this region of the image in the task model so that the HO is not performing the task based on artifacts. As for the latter concern, we have consistently found that the ROIs determined using the proposed method are larger than the typical pixel correlation length.

In the following, we consider the number of projections to be at most 500 views. Therefore, for the ROI sizes inves-

FIG. 3. Left: The integrand of Eq. (13) for the microcalcification detection task. Right: The reconstructed microcalcification signal energy as a function of radius from the signal, $E_y(r)$.

Hanning filter is defined as

$$H(\nu) = 0.5 + 0.5\cos(\nu/\nu_C) \quad 0 \leq |\nu| \leq \nu_C, \\ = 0 \quad \text{otherwise,} \quad (19)$$

where ν_C is the cutoff frequency defining the filter's width.

The impact of breast diameter on image quality is investigated by computing HO P_C for a range of microcalcification sizes and Rayleigh separations for breasts between 10 and 16 cm in diameter. The results are generated first for a fixed exposure level, and subsequently for a fixed average glandular dose, where x-ray fluence for a fixed average glandular dose is again obtained from Ref. 37. The values of HO SNR² are computed at the three different signal locations mentioned above and averaged, so that the final P_C is computed from the average HO SNR² for the three locations.

3.F.2. Impact of view number

In order to study the impact of the number of projection views (at a fixed total dose), the size of the microcalcification is set to 100 μm , and the Rayleigh task separation is set to 0.3 mm. These signal sizes are chosen since they were typically shown to result in a HO P_C value of near 95%, corresponding to a possible but not trivial task. 50, 100, 200, and 500 view cases are investigated, and for each of these cases, a range of reconstruction filter widths are considered (see below).

3.F.3. Reconstruction parameters

The Hotelling efficiency ε is used to determine the image quality impact of two reconstruction parameters: image pixel size and Hanning window width. HO efficiency is computed for a range of signal sizes for two image pixel widths, 0.162 and 0.384 mm. For this study and the view number study, we consider the performance of the HO averaged at three locations within the breast. The HO efficiency is then computed for a Hanning cutoff ranging between $0.125\nu_N$ and $2.0\nu_N$, where ν_N is the Nyquist frequency. The Fourier domain representation of the corresponding apodized ramp kernels is shown in Fig. 5.

FIG. 5. Examples of the impact of various Hanning filter cutoff frequencies on the apodized ramp filter (frequency domain).

3.F.4. Comparison to subjective assessment

After using the HO methodology to determine an optimal set of acquisition and reconstruction parameters, we reconstruct simulated images of several numerical phantoms: a simulated breast phantom containing microcalcifications, a breast phantom with a Rayleigh signal, and simulated high- and low-contrast resolution quality assurance phantoms, based roughly on the *Catphan* phantom (The Phantom Laboratory, Salem, NY). Example reconstructions are also computed for reconstruction filter widths and view numbers which are nonoptimal, so that the scale of refinement achieved through HO optimization can be seen qualitatively.

4. RESULTS—PATIENT AND SYSTEM PARAMETERS

4.A. Breast diameter study

Figure 6 illustrates the impact of breast size on HO performance for the two tasks considered in this work, microcalcification detection (Left) and Rayleigh discrimination (Right). Exposure for these results is fixed so that the mean glandular dose to the 14 cm diameter breast is roughly equal to that of two-view planar mammography. Clearly, since exposure is fixed in these results, larger breasts lead to poorer task performance, as an increasing amount of noise is inherent in the projection data. This finding is in keeping with the results of Yang *et al.*²⁷ For a 14 cm breast diameter, these results indicate that 90 μm microcalcifications are the smallest microcalcifications which can be detected with 95% accuracy, while microcalcifications larger than 120 μm can likely be detected with near perfect accuracy. Gong *et al.* found that using a similar system in simulation resulted in a human observer performance of greater than 95% correct in detection of microcalcifications with diameter equal to or greater than 175 μm .¹⁷ Given that the HO can be expected to establish an upper bound on human observer performance, a 95% accuracy in detection of 90 μm microcalcifications is reasonable, particularly in light of the differences in system simulation and observer experiment methodology between the current work and the study by Gong *et al.* Specifically, one would expect the inclusion of optical light spread, which was not considered in our work, to lessen the gap between the two results.

The results of the Rayleigh discrimination experiment illustrate that for a 14 cm diameter breast, two high-contrast objects can be well resolved by the HO (with 95% accuracy) when their centers are separated by roughly 0.45 mm. This separation is significantly larger than the smallest detectable microcalcification, simply because the Rayleigh task is limited by pixelization in the image, whereas microcalcification detection is still possible even when the image pixel size is somewhat larger than the microcalcification to be detected. The image pixel size for these results is 0.162 mm, and the HO can only consistently perform the Rayleigh discrimination task once the separation between objects spans several pixels.

Figure 7 illustrates the impact of breast size on task performance when the radiation dose is held constant, equal to a

FIG. 6. Left: The HO P_C as a function of microcalcification diameter for a range of breast sizes and a constant x-ray exposure. The quantity d in the legend is the breast diameter simulated. In a 2AFC trial, the 95% correct performance level of the HO would range from microcalcifications of roughly $60\ \mu\text{m}$ for a 10 cm breast versus roughly $100\ \mu\text{m}$ for a 16 cm breast. Right: The HO P_C as a function of the separation between the blobs in the Rayleigh task. As expected, for reasonable task performance, the scale of the Rayleigh signal is substantially larger than that of a detectable microcalcification, as reflected in the left-hand figure. For the breast diameters studied here, the separation at 95% correct ranges from roughly 0.35 to 0.5 mm.

two-view mammogram. Again, these results correspond closely to those from Yang *et al.*,²⁷ who demonstrated an increase in image noise power with decreasing breast size, which would lead to a decrease in task performance. That work demonstrated a noise level for smaller breasts which was nearly an order of magnitude higher for a 10.4 cm breast compared to a 18.4 cm breast. While our results do reproduce the same trend observed in that work, the noise penalty for smaller breast size does not appear to be as dramatic in our simulations.

4.B. Signal location study

Figure 8 illustrates the effect of radial distance from the axis of rotation on the HO task performance for microcalcification detection (Left) and Rayleigh discrimination (Right). A slight increase in performance is seen for signals located away from the axis of rotation for the microcalcification and Rayleigh tasks, and this effect is seen for the full range of signal sizes investigated here. Kwan *et al.*²⁵ demonstrated a significant reduction in resolution measure via the MTF with increasing radial distance from the axis of rotation. For microcalcification

detection, our result therefore demonstrates that the location-dependent image noise properties have a stronger impact on task performance than image resolution. Note that a bowtie filter was not simulated in our results, so that image noise is lower in the periphery of the image.

4.C. Image pixel size

Figure 9 illustrates the impact of image pixel size on HO efficiency for a range of detection and Rayleigh tasks. Clearly, finer pixelization in the image grid results in a substantial benefit in HO efficiency for the full range of investigated microcalcification sizes (Left). For the Rayleigh discrimination task, there is also a pronounced benefit for smaller pixels, which is to be expected since the separation of the Rayleigh signals is less than 2 pixels when 0.324 mm pixels are used. The jump in efficiency for object separations below 0.15 mm on 0.162 mm pixels is an artifact of the nature of the Rayleigh discrimination task changing once the two objects are in adjacent pixels. While the algorithm regains some efficiency in this signal regime, this is immaterial since the HO SNR which

FIG. 7. Left: The HO P_C as a function of microcalcification diameter for a range of breast sizes and constant breast AGD. These results are inverted relative to the corresponding results for constant exposure. This is likely due to the increased noise power for smaller breasts at constant dose which, as described in Ref. 25, is a consequence of the physical relationship between photon flux and breast diameter at a constant AGD. Right: Shown is HO P_C as a function of the separation between the blobs in the Rayleigh task for various breast diameters with AGD held constant. These results mirror those shown in the left-hand figure.

FIG. 8. Left: Plots of HO P_C as a function of calcification size for a range of signal locations. The radial position r denotes the distance of the signal from the axis of rotation. Note that a bowtie filter was not simulated, and the noise at the periphery of the phantom is therefore reduced relative to the center of the FOV. Right: Corresponding results to the left-hand figure, but for Rayleigh discrimination.

is being preserved from the projection domain is low to begin with. In other words, although the algorithm becomes slightly more efficient in this domain, the task is essentially impossible regardless. For an illustration of this point, see Fig. 10, which shows the HO percent of correct decisions for the same image pixel sizes and tasks as are shown in Fig. 9.

These results stand in contrast to those from Kwan *et al.*,²⁵ who evaluated the impact of image pixel size using MTF measurements. That study found that selection of a 512×512 image pixel grid (0.324 mm pixels) versus a 1024×1024 grid (0.162 mm pixels) had minimal impact on image resolution. Modeling of additional physical factors in our simulated signal, such as spread of optical photons in the CsI flat-panel detector, could account for the discrepancy in these findings. It is also possible that directly modeling the nickel–chromium wire used by Kwan *et al.* instead of a microcalcification could lead to somewhat different results. Finally, the presence of correlated noise in the images could impact task performance for the HO, facilitating improved performance with smaller pixels. This effect would not be captured by MTF measurements.

4.D. View number study

Figure 11 shows the impact of the number of projection views on the HO efficiency. The size of the microcalcification signal is held at $100\ \mu\text{m}$, and the separation of the Rayleigh signal is set to 0.3 mm. The total exposure for each acquisition is held fixed, so that the exposure per projection view varies. Kwan *et al.* predict improved resolution properties for larger numbers of views,²⁵ while Yang *et al.* predict decreased image noise for fewer projection views,²⁷ so that the impact of view number on task performance in this case is unclear without a task-based approach to image quality evaluation. We find that when the signal of interest is at the center of the FOV, the noise benefits of restricting view number and the resolution benefits of increasing view number roughly cancel and lead to fairly uniform task performance for a wide range of projection view numbers. By contrast, as the signal moves toward the periphery of the FOV, a slight benefit is seen for lower view numbers. After averaging performance over the three signal locations, this effect is only noticeable for the Rayleigh task using 50 views and a wide Hanning filter. Since view number appears to only negligibly impact task performance in this case, the

FIG. 9. Left: The HO efficiency ϵ is shown as a function of microcalcification diameter for two image pixel sizes: 0.324 and 0.162 mm. In general, smaller pixels result in greater algorithm efficiency in preserving detectability. Right: Shown are the HO efficiency for two image pixel grid sizes for the Rayleigh discrimination task. See text for discussion.

FIG. 10. Left: The HO percent correct is shown for the same microcalcification sizes and pixel sizes as in Fig. 9. Right: Same as left-hand figure, but for the Rayleigh discrimination task.

remaining results in this work correspond to 100 projection views in order to facilitate computational efficiency.

It is worth noting that our simulations neglect x-ray source motion, so that the undersampling artifacts of an x-ray source operating in continuous mode are not captured in these results. In a study not presented here, we find that such blurring from source motion becomes the dominant factor in task performance for signals near the FOV periphery. In a real system, this effect could be controlled somewhat by either using a step-and-shoot acquisition or a pulsed x-ray source.

4.E. Hanning filter width study

Figure 11 also illustrates the effect of a Hanning filter on the performance of the two tasks. Moderate smoothing via the Hanning filter results in a modest improvement in HO efficiency for each task. Meanwhile, over-filtering impairs HO performance for the Rayleigh task more significantly than for the detection task, however the optimal filter widths for the two tasks seem to correlate closely.

FIG. 11. Left: Detection task HO efficiency ϵ as a function of Hanning window width for a range of projection views from 50 to 500 and 0.162 mm pixels image pixels. All plots between 50 views and 500 essentially are overlaid. The total image dose is held constant. Right: Rayleigh task HO efficiency versus the width of the Hanning window for 0.162 mm image pixels. As before, there is little discernible difference for view numbers ranging from 50 to 500.

5. DISCUSSION—PATIENT PARAMETERS AND SYSTEM OPTIMIZATION

The results illustrated in Figs. 6 and 7 illustrate the impact of patient breast diameter on two tasks: detectability of microcalcifications and the Rayleigh task. The performance of each of these tasks is limited not only by the spatial frequency content in the images but also by the presence of noise, and the latter effect is clearly what is modified by a change in breast diameter. The study by Yang *et al.* which investigated the impact of breast diameter on image noise illustrated a degradation in image quality for decreased breast diameter when the mean radiation dose to the breast is kept constant. The same conclusion is reproduced here, but through a task-based approach that also quantifies an upper bound for task performance for a range of signal sizes and breast diameters.

The effect of signal location within the breast CT scanner's field of view appeared to impact task performance somewhat less than previous studies would suggest, however the basic trend that the detection task is somewhat more difficult closer to the axis of rotation in the absence of a bowtie filter is shown in Fig. 8. Use of the HO efficiency metric demonstrated that

for all investigated signal sizes, performance was improved through use of smaller image pixels (see Fig. 9). While the magnitude of this effect depends somewhat on the signal size, the improvement was relatively constant for signal sizes near the threshold of reliable task performance, roughly $100\ \mu\text{m}$ for microcalcification size, and $0.3\ \text{mm}$ for Rayleigh blob separation. Figure 11 also shows the impact of reconstruction implementation by illustrating the impact of width of a Hanning window in reconstruction. For each task considered, a slow improvement in performance was seen with increasing filtering until a shallow maximum followed by a sharp drop in performance for over-filtering.

Figure 11 also illustrates the impact of view number with fixed total radiation dose. The number of projection views was shown to have a minimal impact on task performance, with a slight improvement for few-view acquisitions (50 projection views). The lack of impact from view number is likely due to a combination of the small size of the ROIs investigated and the step-and-shoot model for the x-ray source motion. The small ROIs tended to remain free from angular undersampling artifacts. This highlights the relevance of task modeling, in that since we were only interested in improvement of classification for small signals, global changes to the image quality did not affect the metrics used. Clearly, if a fluoroscopic x-ray mode were modeled, wherein the x-ray source is constantly emitting x rays, there would also be a degradation in task performance due to blurring in the projection data from the moving x-ray source, and 50 projection views would be inadequate. Pulsed x-ray sources present an intermediate regime between step-and-shoot and fluoroscopic acquisitions, and investigation of the impact of view number with these sources requires further study.

6. RESULTS—OPTIMIZED RECONSTRUCTIONS

The foregoing image quality studies demonstrated that a Hanning window of width between $0.75\nu_N$ and $1.0\nu_N$, along with a 1024×1024 image pixel grid ($0.162\ \text{mm}$ pixel width) yields optimal HO performance for detection of $100\ \mu\text{m}$ microcalcifications for the imaging system model used here. As validation of this result, noisy reconstructed images with artificially amplified microcalcifications are shown in Fig. 12. The top center reconstruction was generated with the optimized parameters, $1.0\nu_N$ Hanning width and $0.162\ \text{mm}$ pixels. The left column used a Hanning window width of $0.5\nu_N$, while the right column was generated with a Hanning window width of $1.5\nu_N$. The bottom row images were all reconstructed using a 512×512 image pixel grid ($0.324\ \text{mm}$ pixels). Each image was reconstructed from 50 equally spaced projection views.

From inspection of these noisy reconstructed images, one sees that the HO metrics lead to a set of reconstruction parameters that produce a reasonable image, with enough smoothing from the Hanning filter to limit the background noise, but not so much smoothing that the signal itself became substantially attenuated. With the artificially amplified microcalcifications, it is not immediately clear that the larger pixels would result in decreased task performance. This illustrates that the HO

FIG. 12. Noisy reconstructed images are shown with microcalcification signals. The microcalcifications are $100\ \mu\text{m}$ in diameter and artificially amplified. The top center image was reconstructed using a Hanning window width of $1.0\nu_N$, where ν_N is the Nyquist frequency. For all images, 50 projection views were used. The top row images correspond to a 1024×1024 image pixel grid ($0.162\ \text{mm}$ pixels). The left column of images was reconstructed with heavier smoothing using a $0.5\nu_N$ Hanning window, and the right column used a $1.5\nu_N$ Hanning window. The bottom row demonstrates reconstruction using a 512×512 image pixel grid ($0.324\ \text{mm}$ pixels). The diameter of the circular ROI is approximately $5\ \text{mm}$, and the display window used is $[0, 0.07]\ \text{mm}^{-1}$.

enables the determination of image quality information which cannot be reliably obtained through subjective inspection of a single simulated image by a human.

By the nature of the metrics used in this work, there is a clear dependence of the outcome of system parameter optimization on the task being modeled, i.e., on the objects being imaged and noise correlations. This dependence is important to understand for two reasons. First, any task which is modeled for assessing model observer performance will necessarily be a simplification of the more realistic task which one hopes to ultimately perform. In order to help ensure that this simplification does not lead to unreasonable conclusions for system optimization, inspection of simulated images of more realistic tasks provides a useful check for the derived optimal system parameters. In this way, far from solely relying on a noisy simulated image for system design, simulations constitute a useful supplement to the more fundamental objective assessment methods advocated in this work. Figures 13 and 14 address this point by showing noisy reconstructed images of a numerical breast phantom with simulated microcalcifications. A reconstructed image with determined optimal system parameters is shown, along with images corresponding to two suboptimal Hanning filter widths. Since the HO appeared insensitive to the number of projection views with fixed total dose, an image corresponding to 500 projection views is shown in order to ensure that using 100 views indeed does not negatively impact the microcalcification detection task. The microcalcifications are not artificially enhanced in this case, but instead are made large enough to be visualized with relative ease.

Another sense in which the form of the object being imaged is important is that stylized phantoms commonly used for quality assurance in CT may not be appropriate for task-specific

FIG. 13. Example noisy reconstructions from a simulated breast phantom with microcalcification cluster. The ROIs outlined in white are shown in the images on the right. (a) A reconstruction from 100 views with reconstruction parameters determined to be optimal for microcalcification detection by the HO. (b) The same phantom reconstructed from 500 views. (c) A reconstruction with a wider-than-optimal ($2.0\nu_N$) Hanning filter. 100 views were acquired for this image. (d) A 100-view reconstruction with a narrower-than-optimal ($0.375\nu_N$) Hanning filter. The display window is $[0, 0.05] \text{ mm}^{-1}$.

system optimization. Figures 15 and 16 address this point by showing noisy reconstructions of a phantom based on the *Catphan* phantom for the same system parameters as Figs. 13 and 14. We outline the specifics of each of the four phantoms used in this study.

Figure 13 shows noisy reconstructions from a simulated breast phantom with a microcalcification cluster. Five microcalcifications of sizes 0.125, 0.15, 0.175, 0.2, and 0.225 mm were included, with the smallest microcalcification located to the right, and the largest in the center. The image labeled Fig. 13(a) was reconstructed from 100 views with parameters found to be optimal by the HO. The image labeled Fig. 13(b) was reconstructed from 500 views, but the reconstruction parameters were unchanged. Figures 13(c) and 13(d) were reconstructed from 100 views with Hanning filters that were $2.0\nu_N$

and $0.375\nu_N$ wide, respectively. These widths were chosen because they each corresponds to HO efficiency of approximately 75%–80%, compared to a maximal HO efficiency of 90%–95%.

Figure 14 shows noisy reconstructed images from a similar numerical breast phantom, but with the addition of a pair of simulated microcalcifications in close proximity to simulate the Rayleigh task. The top-left image was reconstructed from 100 views using parameters deemed optimal by the HO for the Rayleigh task. Figures 14(b)–14(d) correspond to a 500-view reconstruction, a $2.0\nu_N$ -wide Hanning filter, and a $0.375\nu_N$ -wide Hanning filter, respectively.

Figure 15 shows reconstructions from a bar phantom based on the *Catphan* phantom. The ROI shown to the right contains bars with separations of 0.173, 0.186, 0.20, and 0.22 mm. The

FIG. 14. Noisy reconstructed images from a breast phantom including two barely resolvable microcalcifications to emulate the Rayleigh task. (a) Reconstruction from 100 views and optimal reconstruction parameters. (b) Reconstruction from 500 views. (c) Reconstruction with a wide ($2.0\nu_N$) Hanning filter. (d) Reconstruction with a narrow ($0.375\nu_N$) Hanning filter. The display window is $[0, 0.1] \text{ mm}^{-1}$.

FIG. 15. Reconstructed noisy images from a bar phantom. The bars shown in the ROI image are of separations ranging from 0.173 to 0.22 mm. (a) Reconstruction from 100 views with parameters optimal for Rayleigh discrimination. (b) Reconstruction from 500 views. (c) Reconstruction with wider Hanning filter ($2.0\nu_N$). (d) Reconstruction with narrower Hanning filter ($0.375\nu_N$). The 0.22 mm separation bars are resolvable in all images except (d). The display window is $[0, 0.0544] \text{ mm}^{-1}$.

image labeled Fig. 15(a) is reconstructed from 100 views with parameters optimal for performing Rayleigh discrimination. The other subfigures are, as before, either from 500 views, reconstructed with a wide Hanning filter, or reconstructed with a narrow Hanning filter. The 0.22 mm bars appear distinct in all but image Fig. 15(d), while the 0.20 mm bars are only questionably resolved in each image. Resolution of smaller bars seems impossible based on these images, regardless of the parameter choices.

Finally, Fig. 16 demonstrates noisy reconstructions of a low-contrast resolution phantom, roughly based on the *Catphan* phantom. The optimal image, Fig. 16(a), was reconstructed using parameters which were optimal for microcalcification detection. The other images were generated as in the preceding figures. Faint undersampling artifacts are visible in the images reconstructed from 100 views [Figs. 16(a), 16(c),

and 16(d)]. Low-contrast visibility is consistently improved with heavier filtering, while the higher-contrast inserts are arguably equally visible in each image.

7. DISCUSSION OF OPTIMIZED RECONSTRUCTED IMAGES

Figures 13–16 show examples of noisy reconstructed images for a variety of numerical phantoms. In each case, an optimal Hanning filter setting and view number are used, along with two suboptimal filters and a larger number of views. Based on these few image realizations, in each case, increasing the view number from 100 to 500 views appears to have minimal impact on the resulting image within the selected ROIs. For the microcalcification detection phantom (Fig. 13), under-

FIG. 16. Noisy reconstructed images from a low-contrast resolution numerical phantom. (a) Reconstruction with optimal parameters from 100 views. (b) Reconstruction from 500 views. (c) $2.0\nu_N$ Hanning width. (d) $0.375\nu_N$ Hanning width. The display window is $[0, 0.05] \text{ mm}^{-1}$.

smoothing seems to result in the smallest microcalcification being overwhelmed by noise, while over-filtering increases the noise correlation distance to approximately the size of the microcalcification. Therefore, the smallest microcalcification seems marginally easier to visualize for the optimized Hanning filter width. Meanwhile, the Rayleigh signal shown in Fig. 14 appears approximately equivalent among all except the over-smoothed image. This and the preceding figure demonstrate that subjective assessment is not necessarily as sensitive to parameter choices or acquisition settings as the objective HO assessment. However, subjective assessment of these particular simulated images tends to corroborate the HO optimization by demonstrating that the optimal parameters do, in fact, produce reasonable images.

Figures 15 and 16 highlight the importance of task-based assessment in that while each of these two phantoms is intended to summarize image quality, neither provides results which agree with the more relevant HO optimization or with the subjective assessment in Figs. 13 and 14. In the case of Fig. 15, only the over-smoothed reconstruction resulted in a noticeable difference in resolution of the bar patterns. Meanwhile, the results in Fig. 16 seem to contradict the conclusion of the HO and of Fig. 13, with the over-smoothed image more clearly showing the small disks.

8. CONCLUSIONS

We have investigated the use of a model observer based approach to image quality evaluation for the purpose of guiding design and use of a dedicated breast CT system. The use of the proposed methodology, based on the Hotelling observer, was demonstrated through investigation of performance in two tasks: microcalcification detection and Rayleigh discrimination. Breast diameter, signal location, image grid size, reconstruction filter, and projection view number were all considered, and the impact of each parameter on the HO metrics was computed. The proposed approach to system optimization and evaluation is applicable for any linear reconstruction algorithm, and the results of this study indicate that the HO and its associated metrics are a versatile and easily implemented tool for characterizing breast CT image quality and guiding decisions in system design and optimization. The value of this approach is further increased by the fact that it does not rely on human observers, repeated experimental scans, or assumptions of shift-invariance or stationarity in the reconstructed images. Further work is needed to include additional realism in the simulations, such as background variability and realistic anatomic background structure.

ACKNOWLEDGMENTS

This work was supported in part by NIH R01 Grant Nos. EB000225, CA158446, CA182264, EB018102 and NIH Grant No. F31EB018742. The contents of this paper are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

^{a)}Electronic mail: adrianalberto@uchicago.edu

^{b)}Electronic mail: sidky@uchicago.edu

^{c)}Also at Department of Cellular and Radiation Oncology, The University of Chicago, Chicago, Illinois 60637.

¹K. J. Myers, H. H. Barrett, M. C. Borgstrom, D. D. Patton, and G. W. Seeley, "Effect of noise correlation on detectability of disk signals in medical imaging," *JOSA A* **2**(10), 1752–1759 (1985).

²R. D. Fiete, H. H. Barrett, W. E. Smith, and K. J. Myers, "Hotelling trace criterion and its correlation with human-observer performance," *JOSA A* **4**(5), 945–953 (1987).

³H. H. Barrett, J. Yao, J. P. Rolland, and K. J. Myers, "Model observers for assessment of image quality," *Proc. Natl Acad. Sci.* **90**(21), 9758–9765 (1993).

⁴J. P. Rolland and H. H. Barrett, "Effect of random background inhomogeneity on observer detection performance," *JOSA A* **9**(5), 649–658 (1992).

⁵H. H. Barrett and W. Swindell, *Radiological Imaging: The Theory of Image Formation, Detection, And Processing* (Academic, New York, NY, 1996).

⁶G. M. Jenkins and D. G. Watts, *Spectral analysis*, 1968.

⁷J. C. Dainty and R. Shaw, *Imaging Science. Principles, Analysis and Evaluation of Photographic-type Imaging Processes* (Academic, London, 1974).

⁸R. B. Blackman, J. W. Tukey, J. W. Tukey, and J. W. Tukey, *The Measurement of Power Spectra: From the Point of View of Communications Engineering* (Dover, New York, NY, 1959), Vol. 190.

⁹K. Rossmann, "Measurement of the modulation transfer function of radiographic systems containing fluorescent screens," *Phys. Med. Biol.* **9**(4), 551–557 (1964).

¹⁰K. Rossmann, "The spatial frequency spectrum: A means for studying the quality of radiographic imaging systems," *Radiology* **90**(1), 1–13 (1968).

¹¹J. Baek and N. J. Pelc, "Local and global 3d noise power spectrum in cone-beam CT system with FDK reconstruction," *Med. Phys.* **38**(4), 2122–2131 (2011).

¹²A. R. Pineda, D. J. Tward, A. Gonzalez, and J. H. Siewerdsen, "Beyond noise power in 3d computed tomography: The local NPS and off-diagonal elements of the fourier domain covariance matrix," *Med. Phys.* **39**(6), 3240–3252 (2012).

¹³M. Deffrise, A. Geissbuhler, and D. W. Townsend, "A performance study of 3d reconstruction algorithms for positron emission tomography," *Phys. Med. Biol.* **39**(3), 305–320 (1994).

¹⁴J. M. Boone, T. R. Nelson, K. K. Lindfors, and J. A. Seibert, "Dedicated breast CT: Radiation dose and image quality evaluation I," *Radiology* **221**(3), 657–667 (2001).

¹⁵B. Chen and R. Ning, "Cone-beam volume CT breast imaging: Feasibility study," *Med. Phys.* **29**, 755–770 (2002).

¹⁶W. T. Yang, S. Carkaci, L. Chen, C.-J. Lai, A. Sahin, G. J. Whitman, and C. C. Shaw, "Dedicated cone-beam breast CT: Feasibility study with surgical mastectomy specimens," *Am. J. Roentgenol.* **189**(6), 1312–1315 (2007).

¹⁷X. Gong, A. A. Vedula, and S. J. Glick, "Microcalcification detection using cone-beam CT mammography with a flat-panel imager," *Phys. Med. Biol.* **49**(11), 2183–2195 (2004).

¹⁸R. L. McKinley, M. P. Tornai, L. A. Tuttle, D. Steed, and C. M. Kuzmiak, "Development and initial demonstration of a low-dose dedicated fully 3d breast CT system," *Breast Imaging* (Springer, New York, NY, 2012), pp. 442–449.

¹⁹X. Gong, S. J. Glick, B. Liu, A. A. Vedula, and S. Thacker, "A computer simulation study comparing lesion detection accuracy with digital mammography, breast tomosynthesis, and cone-beam CT breast imaging," *Med. Phys.* **33**, 1041–1052 (2006).

²⁰S. J. Glick, "Breast CT," *Annu. Rev. Biomed. Eng.* **9**, 501–526 (2007).

²¹E. D. Pisano, C. Gatsonis, E. Hendrick, M. Yaffe, J. K. Baum, S. Acharyya, E. F. Conant, L. L. Fajardo, L. Bassett, and C. D'Orsi, "Diagnostic performance of digital versus film mammography for breast-cancer screening," *N. Engl. J. Med.* **353**(17), 1773–1783 (2005).

²²K. K. Lindfors, J. M. Boone, T. R. Nelson, K. Yang, A. L. C. Kwan, and D. F. Miller, "Dedicated breast CT: Initial clinical experience I," *Radiology* **246**(3), 725–733 (2008).

²³A. Rose, "Human vision," *Vision* (Springer, New York, NY, 1974), pp. 29–53.

²⁴Y. Shen, Y. Zhong, C. Lai, T. Wang, and C. Shaw, "Effects of projection view sampling on CT numbers and noise level in cone beam breast CT," *Med. Phys.* **40**(6), 125 (2013).

²⁵Alexander L. C. Kwan, J. M. Boone, K. Yang, and S.-Y. Huang, "Evaluation of the spatial resolution characteristics of a cone-beam breast CT scanner," *Med. Phys.* **34**(1), 275–281 (2006).

- ²⁶K. Yang, A. L. C. Kwan, and J. M. Boone, "Computer modeling of the spatial resolution properties of a dedicated breast CT system," *Med. Phys.* **34**(6), 2059–2069 (2007).
- ²⁷K. Yang, A. L. C. Kwan, S.-Y. Huang, N. J. Packard, and J. M. Boone, "Noise power properties of a cone-beam CT system for breast cancer detection," *Med. Phys.* **35**(12), 5317–5327 (2008).
- ²⁸C.-J. Lai, C. C. Shaw, L. Chen, M. C. Altunbas, X. Liu, T. Han, T. Wang, W. T. Yang, G. J. Whitman, and S.-J. Tu, "Visibility of microcalcification in cone beam breast CT: Effects of x-ray tube voltage and radiation dose," *Med. Phys.* **34**(7), 2995–3004 (2007).
- ²⁹C. E. Metz, "Basic principles of ROC analysis," *Semin. Nucl. Med.* **8**, 283–298 (1978).
- ³⁰S. J. Glick, S. Thacker, X. Gong, and B. Liu, "Evaluating the impact of x-ray spectral shape on image quality in flat-panel CT breast imaging," *Med. Phys.* **34**(1), 5–24 (2007).
- ³¹N. J. Packard, C. K. Abbey, K. Yang, and J. M. Boone, "Effect of slice thickness on detectability in breast CT using a prewhitened matched filter and simulated mass lesions," *Med. Phys.* **39**(4), 1818–1830 (2012).
- ³²L. Chen, C. K. Abbey, and J. M. Boone, "Association between power law coefficients of the anatomical noise power spectrum and lesion detectability in breast imaging modalities," *Phys. Med. Biol.* **58**(6), 1663–1681 (2013).
- ³³P. C. Johns and M. J. Yaffe, "X-ray characterisation of normal and neoplastic breast tissues," *Phys. Med. Biol.* **32**(6), 675–695 (1987).
- ³⁴J. M. Boone and J. A. Seibert, "An accurate method for computer-generating tungsten anode x-ray spectra from 30 to 140 kV," *Med. Phys.* **24**(11), 1661–1670 (1997).
- ³⁵J. M. Boone, T. R. Fewell, and R. J. Jennings, "Molybdenum, rhodium, and tungsten anode spectral models using interpolating polynomials with application to mammography," *Med. Phys.* **24**(12), 1863–1874 (1997).
- ³⁶J. M. Boone, "Spectral modeling and compilation of quantum fluence in radiography and mammography," *Proc. SPIE* **3336**, 592–601 (1998).
- ³⁷J. M. Boone, A. L. C. Kwan, J. A. Seibert, N. Shah, K. K. Lindfors, and T. R. Nelson, "Technique factors and their relationship to radiation dose in pendant geometry breast CT," *Med. Phys.* **32**, 3767–3776 (2005).
- ³⁸H. H. Barrett and K. J. Myers, *Foundations of Image Science* (John Wiley & Sons, Inc., Hoboken, New Jersey, 2004).
- ³⁹R. F. Wagner, D. G. Brown, and C. E. Metz, "On the multiplex advantage of coded source/aperture photon imaging," *Proc. SPIE* **314**, 72–76 (1981).
- ⁴⁰K. Hanson and K. J. Myers, "Rayleigh task performance as a method to evaluate image reconstruction algorithms," in *Maximum Entropy and Bayesian Methods*, The Fundamental Theories of Physics (Springer, New York, NY, 1991), pp. 303–312.
- ⁴¹K. J. Myers, R. F. Wagner, and K. M. Hanson, "Rayleigh task performance in tomographic reconstructions: Comparison of human and machine performance," *Proc. SPIE* **1898**, 628–637 (1993).
- ⁴²K. J. Myers, Visual perception in correlated noise (models) (1985).
- ⁴³R. D. Fiete, H. H. Barrett, E. B. Cargill, K. J. Myers, and W. E. Smith, "Psychophysical validation of the Hotelling trace criterion as a metric for system performance," *Proc. SPIE* **767**, 298–305 (1987).
- ⁴⁴J. Yao and H. H. Barrett, "Predicting human performance by a channelized Hotelling observer model," *Proc. SPIE* **1768**, 161–168 (1992).