

# The Generalized NEQ and Detectability Index for Tomosynthesis and Cone-Beam CT: From Cascaded Systems Analysis to Human Observers

G. J. Gang,<sup>a</sup> J. Lee,<sup>b</sup> J. W. Stayman,<sup>c</sup> D. J. Tward,<sup>c</sup> W. Zbijewski,<sup>c</sup> J. L. Prince,<sup>b</sup>  
and J. H. Siewerdsen<sup>a,c</sup>

<sup>a</sup>Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto ON, Canada M5G2M9;

<sup>b</sup>Electrical and Computer Engineering, Johns Hopkins University, Baltimore MD, USA 21218

<sup>c</sup>Dept. of Biomedical Engineering, Johns Hopkins University, Baltimore MD, USA 21205

## ABSTRACT

**Purpose:** In the early development of new imaging modalities – such as tomosynthesis and cone-beam CT (CBCT) – an accurate predictive model for imaging performance is particularly valuable in identifying the physical factors that govern image quality and guiding system optimization. In this work, a task-based cascaded systems model for detectability index is proposed that describes not only the signal and noise propagation in the 2D (projection) and 3D (reconstruction) imaging chain but also the influence of background anatomical noise. The extent to which generalized detectability index provides a valid metric for imaging performance was assessed through direct comparison to human observer experiments.

**Methods:** Detectability index ( $d'$ ) was generalized to include anatomical background noise in the same manner as the generalized noise-equivalent quanta (NEQ) proposed by Barrett et al. (Proc. SPIE Med. Imaging, Vol. 1090, 1989). Anatomical background noise was measured from a custom phantom designed to present power-law spectral density comparable to various anatomical sites (e.g., breast and lung). Theoretical calculations of  $d'$  as a function of the source-detector orbital extent ( $\theta_{\text{tot}}$ ) was obtained from a 3D cascaded systems analysis model for tomosynthesis and cone-beam CT (CBCT). Four model observers were considered in the calculation of  $d'$ : prewhitening (PW), non-prewhitening (NPW), prewhitening with eye filter and internal noise (PWE), and non-prewhitening with eye filter and internal noise (NPWE). Human observer performance was measured from 9AFC tests for a variety of idealized imaging tasks presented within a clutter phantom. Theoretical results ( $d'$ ) were converted to area under the ROC curve ( $A_z$ ) and compared directly to human observer performance as a function of imaging task and orbital extent.

**Results:** Theoretical results demonstrated reasonable correspondence with human observer response for all tasks across the continuum in  $\theta_{\text{tot}}$  ranging from low-angle tomosynthesis ( $\theta_{\text{tot}} \sim 10^\circ$ ) to CBCT ( $\theta_{\text{tot}} \sim 180^\circ$ ). Both theoretical and experimental  $A_z$  were found to increase with acquisition angle, consistent with increased rejection of out-of-plane clutter for larger tomosynthesis angle. Of the four theoretical model observers considered, the prewhitening models tended to overestimate real observer performance, while the non-prewhitening models demonstrated reasonable agreement.

**Conclusions:** Generalized detectability index was shown to provide a meaningful metric for imaging performance, helping to bridge the gap between real observer performance and prevalent Fourier-based metrics based in first principles of spatial-frequency-dependent NEQ and imaging task.

**Keywords:** cone-beam CT, anatomical noise, anatomical clutter, cascaded systems analysis, noise-power spectrum, noise-equivalent quanta, detectability index

## 1. INTRODUCTION

Recent years have witnessed a dramatic advancement in the scope of advanced 3D imaging applications based on flat-panel detectors, such as tomosynthesis and cone-beam CT (CBCT). Each of these approaches aims to improve the conspicuity of subtle lesions through depth localization and the rejection of out-of-plane anatomical clutter, with a broad spectrum of potential applications ranging from breast and chest imaging to image-guided interventions. The knowledgeable development of such modalities stands to benefit from a general quantitative framework for performance

modeling, assessment, and system optimization. Two common approaches to image quality assessment include physical characterization of system performance [e.g., prevalent Fourier-based metrics of spatial-frequency-dependent signal and noise] and human observer-based characterization [e.g., sensitivity and specificity as assessed in a receiver operating characteristic (ROC) test]. A variety of advantages and limitations can be appreciated for each approach – for example, the former is simple and practical, yet subject to numerous assumptions of linearity and shift-invariance, while the latter directly assesses real observer response, but is time consuming, expensive, and dependent on the generation of real (or highly realistic) images. A general quantitative approach that begins with simple Fourier metrics and bridges the gap to real observer performance would be valuable in the design and optimization of such imaging systems.

Previous work has extended cascaded systems analysis of the noise-equivalent quanta (NEQ) for projection radiography<sup>1-4</sup> to advanced 2D and 3D imaging modalities<sup>5-8</sup>. Recognizing the importance of imaging task in defining meaningful image quality metrics, a task-driven approach is adopted in which NEQ is combined with the imaging task and model observers to form the detectability index ( $d'$ ). The goal of the current study is to investigate the extent to which detectability index is a valid metric for system performance and optimization through comparison to real human observers. The work below represents preliminary studies spanning initial analysis and experimentation involving a small number of simple imaging tasks, with a more complete evaluation, including a greater variety of imaging conditions and imaging tasks, to be pursued in future work.

## 2. GENERALIZED DETECTABILITY INDEX

### 2.1 Cascaded Systems Analysis (CSA) Model for Tomosynthesis and CBCT

Previous work has demonstrated the utility and accuracy of cascaded systems analysis of the performance of 2D x-ray projection imaging systems based on flat-panel detectors.<sup>3,4</sup> Such has more recently been extended to 3D FPD-based CBCT and tomosynthesis systems,<sup>5-8</sup> providing a general framework for modeling the 3D modulation transfer function (MTF), noise-power spectrum (NPS), NEQ, and detective quantum efficiency (DQE). The basic approach provides quantitative modeling of the signal and noise propagation in the 2D and 3D imaging chain as illustrated in Fig. 1. The model consists of the physical processes of 2D projection image formation followed by the mathematical processes of 3D filtered backprojection reconstruction. The former has been extensively investigated and developed over the last 15 years.<sup>1-4,9</sup> The reconstruction cascade includes<sup>7,8</sup>: log-normalization of projection data (Stage 8); application of the ramp filter (Stage 9); optional application of apodization filters (Stage 10); interpolation of the 2D projection data (Stage 11); backprojection of 2D projections, modeled as superposition of the 2D NPS along vanes in the 3D Fourier domain (Stage 12); and sampling of the 3D reconstruction (Stage 13). Such a model has demonstrated agreement with measurements of the 3D NPS and NEQ for FPD-CBCT over a broad range of experimental conditions and reconstruction techniques<sup>7,8</sup>, providing a model for imaging performance of tomosynthesis and CBCT as a continuum represented by varying extent of acquisition angle ( $\theta_{\text{tot}}$ ).

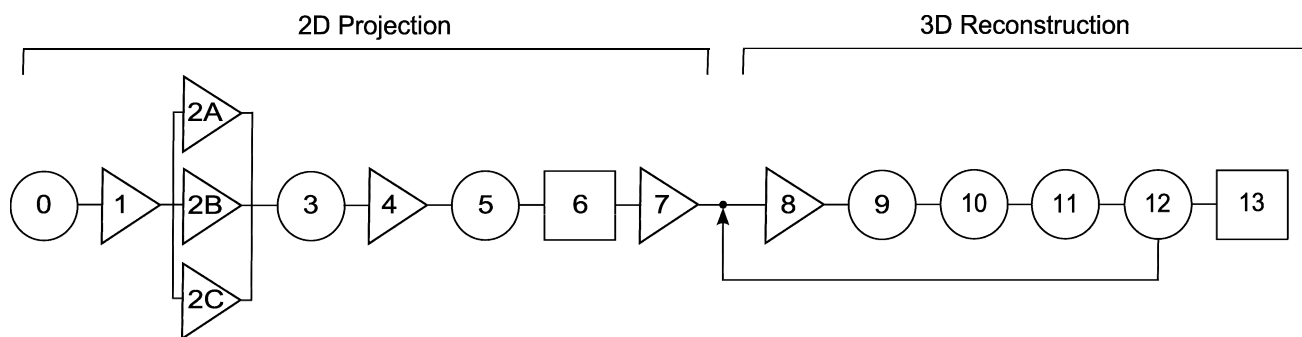


Figure. 1. Cascaded systems analysis for 3D tomosynthesis and CBCT. Stages 0-7 model signal and noise transfer through various physical processes in the detector, while Stages 8 to 13 represent mathematical processes of 3D reconstruction by filtered backprojection as detailed in previous work: 0.) incident x-ray quanta; 1.) interaction of x-rays

in detector; 2.) generation of secondary quanta; 3.) spread of secondary quanta; 4.) coupling of quanta to detector apertures; 5.) integration by detector aperture; 6.) sampling of detector pixels; 7.) readout with additive noise; 8.) log normalization; 9.) ramp filter; 10.) apodization filter; 11.) interpolation; 12.) 3D backprojection; and 13.) 3D sampling.

## 2.2 Anatomical Background: The “Generalized” NEQ and Detectability Index

Such modeling has been shown useful in characterizing the physical performance of systems under the influence of quantum and electronics noise. However, there is considerable evidence that image “noise” (generally interpreted as any random or semi-random component in an image that impedes conspicuity) is often dominated by image fluctuations associated not with the Poisson-distributed incident quanta, but rather those associated with the superposition of anatomical “clutter” in the 2D image – often referred to as “anatomical noise” or “background noise”.<sup>10-13</sup> Similarly in tomosynthesis, depending on the acquisition angle,  $\theta_{\text{tot}}$ , out-of-plane structures are superimposed to varying degrees within a slice of interest, with greater orbital extent reducing the contribution of in-plane anatomical noise. In this way, anatomical noise is expected to be a dominant source of image degradation at low tomosynthesis angles and to be progressively reduced for a larger arc. Potential tradeoffs between quantum noise and anatomical noise are immediately appreciated:

- For fixed dose per projection:
  - a smaller number of projections over a shorter arc will increase quantum noise and anatomical noise
  - a larger number of projections over a shorter arc will reduce quantum noise, reduce view aliasing noise, and increase anatomical noise
  - a smaller number of projections over a wider arc will increase quantum noise, increase view aliasing noise, and reduce anatomical noise
  - a larger number of projections over a wider arc will reduce quantum noise and increase anatomical noise
- For fixed total dose:
  - a smaller number of projections over a shorter arc will reduce quantum noise and increase anatomical noise
  - a larger number of projections over a shorter arc will reduce quantum noise, reduce view aliasing noise, and increase anatomical noise
  - a smaller number of projections over a wider arc will reduce quantum noise, increase view aliasing noise, and reduce anatomical noise
  - a larger number of projections over a wider arc will increase quantum noise and reduce anatomical noise

Clearly, a model that quantitatively describes such tradeoffs would be beneficial in understanding the factors that limits imaging performance under various system configurations. Such tradeoffs can be quantitatively appreciated through incorporation of anatomical noise (the so-called anatomical power spectrum,  $S_B$ ) in the denominator of the NEQ. Barrett et al.<sup>14</sup> termed this a form of “generalized NEQ,” written for the 3D case as:

$$GNEQ = \frac{MTF^2}{S_B + S_Q + S_E} \quad (1)$$

where  $S_B$  can take whatever form is appropriate to the anatomical “clutter” and is often empirically characterized according to a power-law form:

$$S_B(f) = \frac{\kappa}{f^\beta} \quad (2)$$

where  $\kappa$  describe the magnitude and  $\beta$  the frequency content of the anatomical clutter. Modeling such image fluctuations in this manner implies treatment of anatomical clutter as a stochastic noise source, which clearly implies numerous assumptions and limitations. Power-law modeling of anatomical noise has shown reasonable agreement with human observer performance measured with real patient background in breast imaging,<sup>15</sup> providing a convenient basis for

incorporation of a “lumpy” background in the generalized NEQ. Such approach has been applied extensively in mammography<sup>12,16-20</sup>, chest imaging<sup>21-23</sup>, and more recently in tomosynthesis and CBCT<sup>24,25</sup>.

Just as NEQ may be combined with an idealized task function as described in ICRU 54 to yield the detectability index:<sup>26</sup>

$$d'^2 = \iiint \frac{MTF^2}{S_Q + S_E} |\Delta H|^2 df_x df_y df_z \quad (3)$$

the generalized NEQ can be similarly extended to yield a generalized detectability index:

$$d'^2 = \iiint \frac{MTF^2}{MTF^2 S_B + S_Q + S_E} |\Delta H|^2 df_x df_y df_z \quad (4)$$

This idealized form for detectability in the presence of anatomical background is equivalent to the prewhitening (PW) observer model. It has been shown in previous work<sup>21,27-29</sup> to give a useful metric for imaging performance and system optimization.

Eq. (4) describes a fully 3D detectability index in which the observer can make full use of the volumetric data. Such may or may not be a reasonable model for a real observer “scrolling” slices; rather, it assumes the observer can instantly appreciate the fully 3D statistics of the volume image. The model can be reduced to describe the detectability within a single slice of 3D image, denoted  $d'_{slice}$ . Since extraction of a single slice from the 3D spatial domain corresponds to a convolution over the orthogonal direction in the Fourier domain (i.e., convolution with a sinc function corresponding to the slice thickness), a simple form of slice detectability for a prewhitening observer can be obtained by integrating the numerator and denominator of the 3D detectability index as:

$$d'_{slice}{}^2 = \iint \frac{[\int T \cdot W_{Task} df_y]^2}{\int T^2 S_B + S_Q + S_E df_y} df_x df_z \quad (5)$$

### 2.3 Model Observers

The simple PW model of Eq. (4) can be extended to slightly more sophisticated models that have been shown to better approximate real observers under various conditions. For example, the non-prewhitening (NPW) observer model is:

$$d'^2 = \frac{[\iiint [T \cdot W_{Task}]^2 df_x df_y df_z]^2}{\iiint (T^2 S_B + S_Q + S_E) \cdot [T \cdot W_{Task}]^2 df_x df_y df_z} \quad (6)$$

Similarly, each model can be extended to include a simple description of a human eye filter and internal noise, referred to respectively as the prewhitening-eye (PWE) and non-prewhitening-eye (NPWE) models:

$$d'^2 = \iiint \frac{E^2 [T \cdot W_{Task}]^2}{E^2 (T^2 S_B + S_Q + S_E) + N_i} df_x df_y df_z \quad (7)$$

$$d'^2 = \frac{[\iiint E^2 [T \cdot W_{Task}]^2 df_x df_y df_z]^2}{\iiint E^4 (T^2 S_B + S_Q + S_E) \cdot [T \cdot W_{Task}]^2 + N_i df_x df_y df_z} \quad (8)$$

The eye filter and internal noise adopted in this work are consistent previous work by Burgess<sup>30</sup>:

$$E(f) = f^n \exp(-cf^2) \quad (9)$$



where  $n = 1.3$  and  $c = 3$  gives an eye filter,  $E(f)$ , that peaks at 4 cycles/deg for a viewing distance of 50 cm. Internal noise is assumed to be uncorrelated and is set to a fraction (0.02) of the DC component of the noise-power spectrum at a viewing distance of 100 cm:

$$N_i = 0.02 \left( \frac{D}{100} \right)^2 NPS(0, 0) \quad (10)$$

where the viewing distance,  $D$ , is set to 50 cm.

## 2.4 Relating $d'$ to Human Observer Performance

Equations (4) – (8) above yield the detectability index for a given task and observer model, with the fundamental signal and noise properties described by a cascaded systems model for the imaging chain. These forms are almost purely theoretical, with the main “semi-empirical” components of the model being: the incident x-ray spectrum; the Poisson excess (Swank noise) of the x-ray converter; and the eye filter parameters (taken from Ref. 30). Relation of  $d'$  derived theoretically from cascaded systems analysis calculations can be converted to area-under the ROC curve ( $A_z$ ) by the relation:

$$A_z = \frac{1}{2} \left( 1 + \operatorname{erf} \left( \frac{d'}{2} \right) \right) \quad (11)$$

under assumption that distribution of cases under two hypotheses are Gaussian and homoscedastic<sup>31</sup>.

# 3. IMAGING TASK

## 3.1 Task Phantoms

To investigate task functions presenting stimuli in a power-law background (anatomical noise), a custom phantom was developed specifically to present power-law noise. The phantom design principles and associated background power spectra are detailed in Ref. 32. To summarize, the phantom consists of a random assortment of acrylic spheres of diameter ranging 3.18 – 15.88 mm, with proportions of each sphere selected to give a power-law background NPS with  $\beta \sim 3$ . For the small-sphere detection task, an assortment of six 3.18 mm diameter Teflon spheres was placed within the central coronal (x-z) slice of the clutter phantom as shown in Fig. 3(b). For the large sphere detection task, six 10.3 mm polypropylene spheres were used as stimuli. For the encapsulated-sphere discrimination task, each stimulus consists of a 10 mm acrylic sphere coated with an 3.18mm outer shell of paraffin wax. Example images of each task phantom are illustrated in Fig. 3(b).

Images of each task phantom were acquired over a range in  $\theta_{\text{tot}}$  as described above, and 3D images were reconstructed using the Feldkamp algorithm for 3D filtered backprojection. A central coronal slice was extracted from the volume image through the central plane of each sphere, giving the slice images used in human observer tests described below. The six stimuli give six independent trials for each task.

## 3.2 Imaging Tasks

A bridge between simple Fourier metrics of imaging performance (e.g., MTF, NPS, and NEQ) and task-dependent performance metrics can be obtained through a quantitative description of the spatial frequencies of interest in performing a given task – i.e., through a spatial-frequency-dependent “task function” or template. A simple form for such task functions is given by the Fourier transform of the difference in two hypotheses –  $H_1(x, y, z)$  and  $H_2(x, y, z)$  – where  $H_1$  and  $H_2$  represent the spatial domain representations of, for example, the “signal-present” and “noise-only” hypotheses. For example, detection of a given stimulus  $[H_1(x,y,z)]$  against a uniform background  $[H_2(x,y,z) = \text{constant}]$  is represented by the task function:

$$W_{\text{Task}} = \Delta\mu \cdot FT[H_1] \quad (12)$$

where  $FT[H_i]$  is the Fourier transform of the stimulus, and  $\Delta\mu$  are the difference in attenuation coefficients of the stimulus and background, respectively.

For detection of a given stimulus against the power-law background described previously, images presenting  $H_1$  contains a stimulus (e.g., a sphere of a different contrast) at the center of the ROI, whereas images presenting  $H_2$  have an acrylic sphere of the same size placed at the center. Thus, the two hypotheses can be written as:

$$FT[H_1] = \mu_1 FT[H_1] + clutter \quad (13a)$$

$$FT[H_2] = \mu_2 FT[H_1] + clutter \quad (13b)$$

Therefore, the task function is given by:

$$W_{Task} = \Delta\mu \cdot FT[H_1] \quad (13c)$$

Finally, discrimination of two structures – the first identified as the stimulus and the second the normal case – is given by:

$$W_{Task} = \mu_1 \cdot FT[H_1] - \mu_2 \cdot FT[H_2] \quad (14)$$

where  $H_1$  and  $H_2$  represent the two structures to be discriminated.

These three simple task functions represent different regions of the spatial frequency domain and therefore imply different susceptibility to various noise sources. High frequency tasks, such as discrimination tasks or detection of a small object, are prone to degradation by high frequency quantum and electronics noise, and are therefore more dependent on factors of detector design and reconstruction technique. Low frequency tasks, such as detection of a large, low-contrast lesion, are more dependent on low-frequency noise characteristics, such as anatomical background. Additional task functions representing higher-order tasks and representing a varied richness of spatial frequency content are under investigation in ongoing work. Example task functions are shown in Fig. 2, along with real images acquired as a function of  $\theta_{tot}$ .

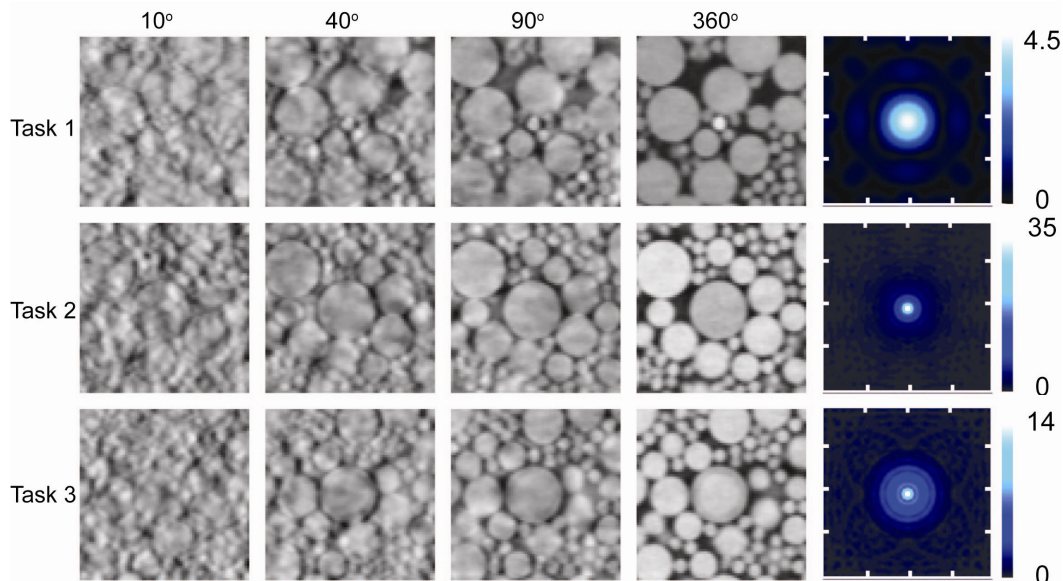


Figure 2. Hypotheses and task functions. Task 1: Detection of a small sphere in clutter. Task 2: Detection of a large sphere in clutter. Task 3. Discrimination of an encapsulated sphere vs. a uniform sphere.

### 3.3 Image Acquisition

An experimental imaging bench for tomosynthesis and cone-beam CT is illustrated in Fig. 3(a) and described in previous work<sup>7,32</sup>. The main components of the bench include: an x-ray tube (Rad 94 in a Sapphire housing; Varian Medical

Systems, Salt Lake City, UT); a flat-panel detector (RID-1640A, Perkin Elmer Optoelectronics, Santa Clara CA), and a motion control system in which a phantom placed at isocenter is rotated through an arc corresponding to  $\theta_{tot}$  (6K series translation stages, Parker Daedal, Harrison PA, and Dynaserv rotation motor, Parker Hannifin, Rohnert Park, CA).

Images were acquired over a range of  $\theta_{tot}$  spanning a continuum from low-angle tomosynthesis ( $\theta_{tot} = 10^\circ$ ) to fully 3D CBCT ( $180^\circ$  fan and  $360^\circ$ ). In the experiments reported below, the angular increment ( $\Delta\theta$ ) was held fixed at  $0.45^\circ$ , and the number of projections ( $N_{proj}$ ) was varied such that  $\theta_{tot}$  ranged  $10^\circ$  to  $360^\circ$ . In all cases, the imaging technique was 120 kVp (1.53 mm Al + 1.1 mm Cu added filtration), and 0.63 mAs per projection. The range in  $N_{proj}$  was 23 (at  $10^\circ$ ) to 800 (at  $360^\circ$ ), corresponding to a range in dose of 0.4 – 13.8 mGy.

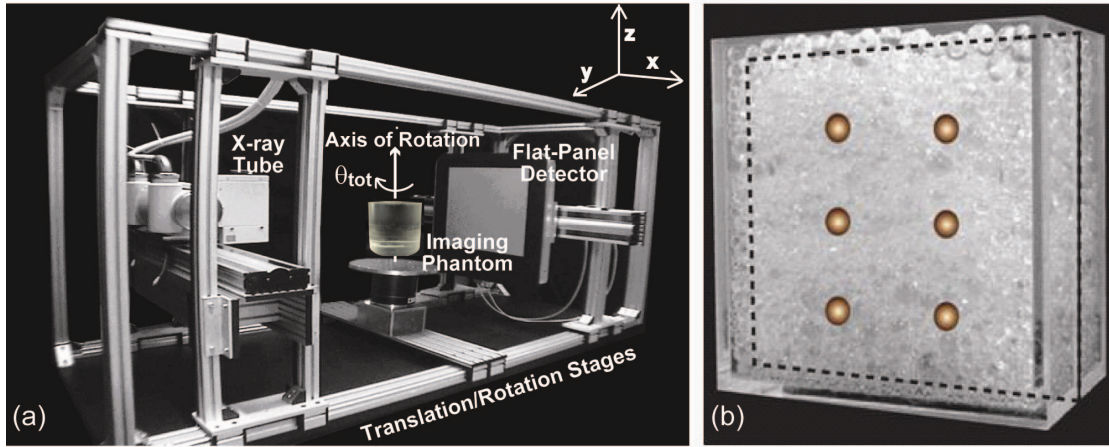


Figure 3. (a) Experimental bench showing components and coordinate systems for CBCT and tomosynthesis. (b) Experimental phantom presenting power-law background designed using principles of fractal self-similarity. Six stimuli were inserted to the central coronal slice.

#### 4. EXPERIMENTAL VALIDATION: HUMAN OBSERVER STUDY

Human observer tests were conducted to measure task performance as a function of  $\theta_{tot}$  for each of the imaging tasks described above. An efficient test well suited to simple phantom studies is offered by the multiple-alternative forced-choice (MAFC) methodology. In the studies described below, 9AFC tests were used to measure the proportion correct ( $P_{corr}$ ) which may in turn be related to  $d'$  as:

$$P_{corr}(d', M) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{(x-d')^2}{2}\right) [\Phi(x)]^{M-1} dx \quad (15)$$

where  $M$  is the number of alternatives, and  $\Phi$  is the cumulative Gaussian distribution. Note that for a 2AFC test,  $P_{corr} = A_z$  under the assumption of an ideal test in which observer response does not vary over the course of the test. For the 9AFC study,  $A_z$  corresponding to experimental value of  $P_{corr}$  is converted to  $A_z$  via a  $P_{corr} - d' - A_z$  lookup table given by Eq. (11) and (15).

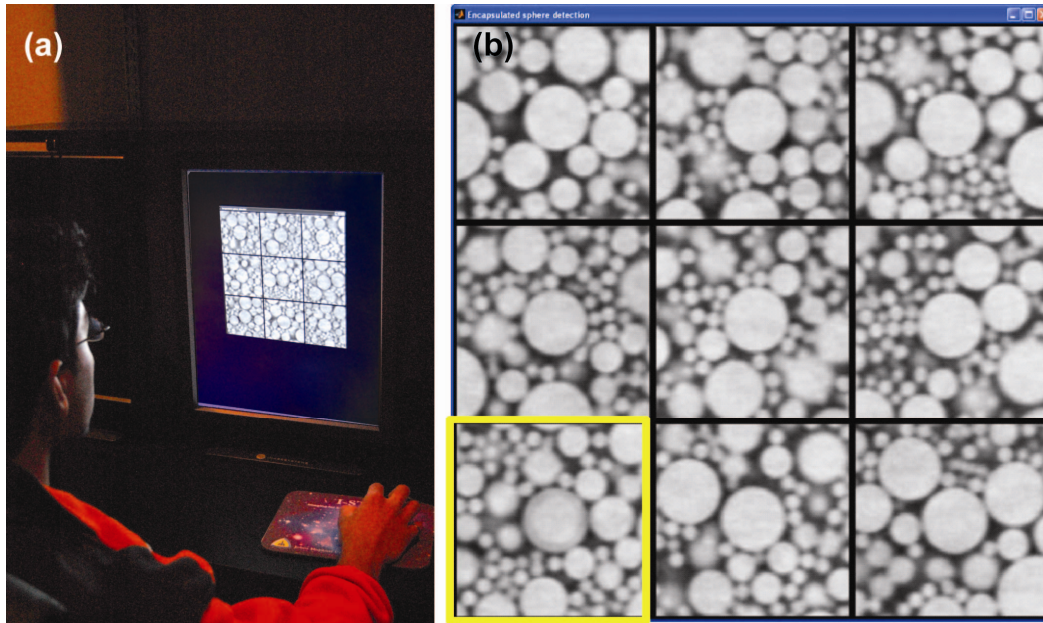


Figure 4. (a) Observer performing a 9AFC test in a darkened reading room. (b) Nine images were shown on a 3x3 grid as displayed by Matlab-based OPTEx software in a randomized order. The example shows the encapsulated sphere task, with the stimulus highlighted in the lower-left corner.

All tests were performed in a controlled image reading environment with subdued lighting and on a diagnostic quality monochrome display (Fig. 4). Reading order was randomized (using the Matlab-based OPTEx utility for human observer studies), and each test was preceded by a training set composed of images drawn from the same task phantom imaged as a function of  $\theta_{tot}$  with images distinct from those in the actual test. For these simple phantom studies, physicists / engineers were sufficiently “expert” to perform each imaging task following the training set. A total of 6 observers were used, giving 5 independent trials for each task at each setting (angle, case), and a total time of ~60 min required for each test. Responses from each observer for each setting of a task are assumed to be independent, i.e. no inter- and intra-observer variability were assumed. Mean proportion correct ( $P_{corr}$ ) was therefore calculated from the number of correct responses out of a total of 30 (=5 trials x 6 observers), and converted to  $A_z$  from the lookup table as mentioned previously. A 95% confidence interval was calculated for  $A_z$  and shown in the results below.

## 5. BRIDGING THE GAP: THEORETICAL AND EXPERIMENTAL OBSERVER PERFORMANCE

Theoretical calculations of  $A_z$  for all four observer models are plotted in Figure 5 in comparison to  $A_z$  measured from real observer studies  $P_{corr}$ . In each case, an upward trend with  $\theta_{tot}$  is observed for both theoretical and experimental results, consistent with improved rejection of clutter observed in Figure 2. Overall, the prewhitening models (PW and PWE) are seen to overestimate real observer performance in each case, whereas the non-prewhitening (NPW and NPWE) observers show reasonable agreement with experimental results. The large sphere detection (discrimination of polyethylene and acrylic spheres) task appears to be most difficult, with observer performance challenged up to  $\theta_{tot} \sim 120^\circ$ , and with the NPW and NPWE models bounding the human observer measurements. The small sphere detection and encapsulated sphere discrimination tasks are a bit more conspicuous, each challenging observers up to  $\theta_{tot} \sim 70\text{-}100^\circ$  and described reasonably well by the NPW or NPWE models. Overall, agreement of theory and measurement is fairly good, providing evidence that simple Fourier metrics such as NEQ combined with imaging task can offer meaningful descriptions of real observer performance.

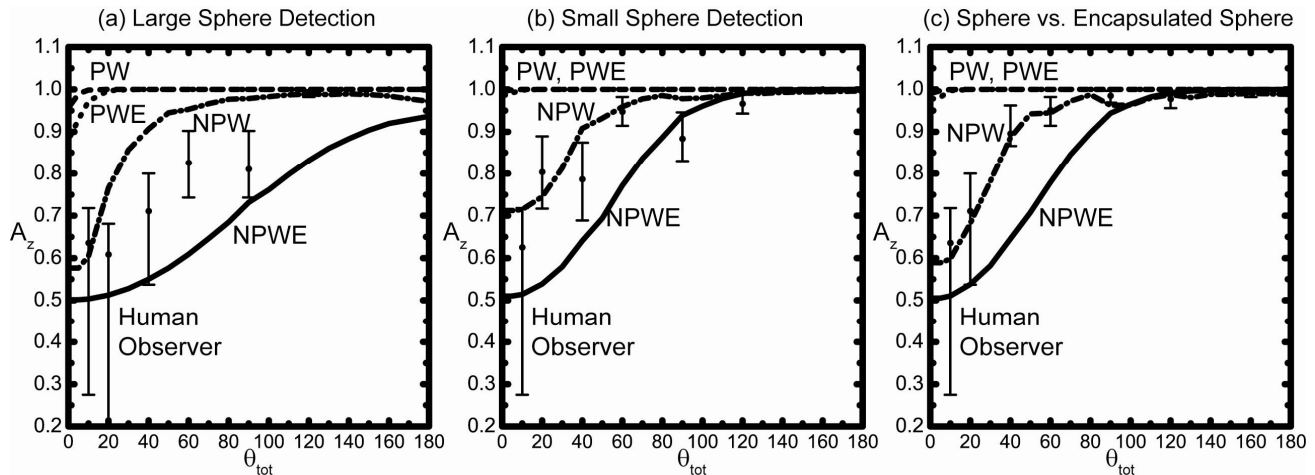


Figure 5. Theoretical and experimental imaging performance ( $A_z$ ) plotted as a function of  $\theta_{\text{tot}}$  for 3 imaging tasks. Reasonable agreement between theory and experiment is observed, with prewhitening observer models (PW and PWE) appearing to overestimate human response, while non-prewhitening observer models (NPW and NPWE) demonstrating closer agreement.

## 6. DISCUSSION AND CONCLUSION

Fourier-based physical characterization of imaging systems has been commonly applied in detector design. In designing and evaluating medical imaging systems with increasing complexity – as in tomosynthesis and CBCT – this approach can be extended to provide simple, elegant performance metrics, but correlation of theoretical predictions with human observer response has not been previously demonstrated. On the other hand, human observer-based characterization realistically describes imaging performance, but can be expensive and time-consuming. This work presented generalized detectability index ( $d'$ ) as a quantitative metric combining NEQ, anatomical background, imaging task, and model observers. Theoretical calculations of  $d'$  were validated with real human observer performance, showing  $d'$  to be a meaningful metric in modeling system performance.

It is important to acknowledge a variety of assumptions and limitations associated with in this study. Cascaded systems analysis of imaging systems assumes linearity, stationarity, and shift invariance. While tomosynthesis and CBCT are recognized to be neither strictly stationary nor shift invariant, we assume the first and second order statistics are at least locally stationary and that system response is shift-invariant to the extent that the studies involved tasks greater than the pixel / voxel size – analogous to the 2D shift-invariance described in Ref. 33. In calculating observer performance metrics  $A_z$ , the distribution of the two hypotheses are assumed to be Gaussian and homoscedestic. Finally, the evaluation involved a small number of simple, idealized imaging tasks, and the clutter phantom, though measured to present power-law background, does not truly simulate the complexity of structures in human organs.

Future work involves the investigation of a broader range of acquisition techniques, especially in the low-dose limit where detection transitions between anatomical noise-limited and quantum-limited regimes. Other reconstruction techniques will also be studied, including different reconstruction filters and sampling, which have been shown to impart significant effects on system NEQ. More complex and higher-order tasks will also be investigated as a fuller examination of the extent to which generalized detectability index demonstrates agreement with real observer performance.

## 7. ACKNOWLEDGEMENTS

The authors acknowledge numerous stimulating conversations and collaboration with Dr. Rebecca Fahrig (Stanford University), Dr. Sungwon Yoon (Stanford University), Dr. Angel Pineda (California State University – Fullerton), and Dr. Art Burgess (Harvard University, retired). This work was funded by National Institute of Health Grant No. R01-CA-112163.



## REFERENCES

1. Cunningham, I. A., Westmore, M. S., and Fenster, A. "A Spatial-Frequency Dependent Quantum Accounting Diagram and Detective Quantum Efficiency Model of Signal and Noise Propagation in Cascaded Imaging Systems," *Med Phys.* 21[3]: 417-427 (1994).
2. Cunningham, I. A., Yao, J., and Subotic, V. "Cascaded Models and the DQE of Flat-Penel Imagers: Noise Aliasing, Secondary Quantum Noise and Reabsorption," *Physics of Medical Imaging. Proc.SPIE Medical Imaging 4682*: 61-72 (2002).
3. Siewerdsen, J. H., Antonuk, L. E., el-Mohri, Y., Yorkston, J., Huang, W., Boudry, J. M., and Cunningham, I. A. "Empirical and Theoretical Investigation of the Noise Performance of Indirect Detection, Active Matrix Flat-Panel Imagers (AMFPIs) for Diagnostic Radiology," *Med Phys.* 24[1]: 71-89 (1997).
4. Siewerdsen, J. H., Antonuk, L. E., el-Mohri, Y., Yorkston, J., Huang, W., and Cunningham, I. A. "Signal, Noise Power Spectrum, and Detective Quantum Efficiency of Indirect-Detection Flat-Panel Imagers for Diagnostic Radiology," *Med Phys.* 25[5]: 614-628 (1998).
5. Siewerdsen, J. H. and Jaffray, D. A. "Cone-Beam Computed Tomography With a Flat-Panel Imager: Noise Considerations for Fully 3-D Imaging," *Proc.SPIE 3977*: 408-416 (2000).
6. Siewerdsen, J. H. and Jaffray, D. A. "Three-Dimensional NEQ Transfer Characteristics of Volume CT Using Direct and Indirect-Detection Flat-Panel Imagers," *Proc.SPIE 5030*: 92-102 (2003).
7. Tward, D. J. and Siewerdsen, J. H. "Cascaded Systems Analysis of the 3D Noise Transfer Characteristics of Flat-Panel Cone-Beam CT," *Med Phys.* 35[12]: 5510-5529 (2008).
8. Tward, D. J., Siewerdsen, J. H., Fahrig, R. A., and Pineda, A. R. "Cascaded Systems Analysis of the 3D NEQ for Cone-Beam CT and Tomosynthesis," *Proc.SPIE 6913*: 69131S-69131S-12 (2008).
9. Zhao, W. and Rowlands, J. A. "Digital Radiology Using Active Matrix Readout of Amorphous Selenium: Theoretical Analysis of Detective Quantum Efficiency," *Med Phys.* 24[12]: 1819-1833 (1997).
10. Bochud, F. O., Verdun, F. R., Valley, J. F., Hessler, C., and Moeckli, R. "The Importance of Anatomical Noise in Mammography," *Proc.SPIE 3036*: 74-80 (1997).
11. Kundel, H. L., Nodine, C. F., Thickman, D., Carmody, D., and Toto, L. "Nodule Detection With and Without a Chest Image," *Investigative radiology* 20[1]: 94-99 (1985).
12. Samei, E., Flynn, M. J., and Eyler, W. R. "Detection of Subtle Lung Nodules: Relative Influence of Quantum and Anatomic Noise on Chest Radiographs," *Radiology* 213[3]: 727-734 (1999).
13. Samei, E., Eyler, W. R., and Baron, L. "Effects of Anatomical Structure on Signal Detection," Beutel, J., Kundel, H. L., and Van Metter, R. L. *Handbook of Medical Imaging: Physics and Psychophysics*. 1[12]: 655-682 (2000). Washington, SPIE Publications.
14. Barrett, H. H., Rolland, J. P., Wagner, R. F., and Myers, K. J. "Detection and Discrimination of Known Signals in Inhomogenous, Random Backgrounds," *Proc.SPIE Medical Imaging III: Image formation 1090*: 176-182 (1989).
15. Burgess, A. E., Jacobson, F. L., and Judy, P. F. "Human Observer Detection Experiments With Mammograms and Power-Law Noise," *Med.Phys.* 28[4]: 419-437 (2001).

16. Burgess, A. E. "Mammographic Structure: Data Preparation and Spatial Statistics Analysis," Proc.SPIE 3661: 642 (1999).
17. Glick, S. J., Vedantham, S., and Karellas, A. "Investigation of Optimal KVp Setting for CT Mammography Using a Flat-Panel Imager," Proc.SPIE 4682: 392-402 (2002).
18. Gong, X., Glick, S. J., Liu, B., Vedula, A. A., and Thacker, S. "A Computer Simulation Study Comparing Lesion Detection Accuracy With Digital Mammography, Breast Tomosynthesis, and Cone-Beam CT Breast Imaging," Med.Phys. 33[4]: 1041-1052 (2006).
19. Heine, J. J., Deans, S. R., Velthuizen, R. P., and Clarke, L. P. "On the Statistical Nature of Mammograms," Med.Phys. 26[11]: 2254-2265 (1999).
20. Heine, J. J. and Velthuizen, R. P. "Spectral Analysis of Full Field Digital Mammography Data," Med.Phys. 29[5]: 647-661 (2002).
21. Richard, S., Siewerdsen, J. H., Jaffray, D. A., Moseley, D. J., and Bakhtiar, B. "Generalized DQE Analysis of Radiographic and Dual-Energy Imaging Using Flat-Panel Detectors," Med.Phys. 32[5]: 1397-1413 (2005).
22. Richard, S. and Siewerdsen, J. H. "Optimization of Dual-Energy Imaging Systems Using Generalized NEQ and Imaging Task," Med.Phys. 34[1]: 127-139 (2007).
23. Yoon, S., Gang, J. G., Tward, D. J., Siewerdsen, J. H., and Fahrig, R. "Analysis of Lung Nodule Detectability and Anatomical Clutter in Tomosynthesis Imaging of the Chest," Proc.SPIE 7258: 72581M-72581M-11 (2009).
24. Engstrom, E., Reiser, I., and Nishikawa, R. "Comparison of Power Spectra for Tomosynthesis Projections and Reconstructed Images," Med.Phys. 36[5]: 1753-1758 (2009).
25. Metheany, K. G., Abbey, C. K., Packard, N., and Boone, J. M. "Characterizing Anatomical Variability in Breast CT Images," Med.Phys. 35[10]: 4685-4694 (2008).
26. ICRU Report 54. "Medical Imaging - The Assessment of Image Quality," (1996). Bethesda, MD, International Commission on Radiation Units and Measurements.
27. Siewerdsen, J. H. and Jaffray, D. A. "Optimization of X-Ray Imaging Geometry (With Specific Application to Flat-Panel Cone-Beam Computed Tomography)," Med Phys. 27[8]: 1903-1914 (2000).
28. Chawla, A. S., Samei, E., Saunders, R. S., Lo, J. Y., and Baker, J. A. "A Mathematical Model Platform for Optimizing a Multiprojection Breast Imaging System," Med Phys. 35[4]: 1337-1345 (2008).
29. Tapiovaara, M. J. "Efficiency of Low-Contrast Detail Detectability in Fluoroscopic Imaging," Med Phys. 24[5]: 655-664 (1997).
30. Burgess, A. E., Li, X., and Abbey, C. K. "Visual Signal Detectability With Two Noise Components: Anomalous Masking Effects," J.Opt.Soc.Am.A Opt.Image Sci.Vis. 14[9]: 2420-2442 (1997).
31. Beutel, J., Kundel, H. L., and Van Metter, R. L. "*Handbook of Medical Imaging, Volume 1. Physics and Psychophysics*," [1] (2000). SPIE Press.
32. Gang, G. J., Tward, D. J., Lee, J., and Siewerdsen, J. H. "Generalized NEQ and Detectability Index in Cone-Beam CT: Influence of Anatomical Background Clutter," Med Phys. In press. (2010).
33. Albert, M. and Maidment, A. D. "Linear Response Theory for Detectors Consisting of Discrete Arrays," Med Phys. 27[10]: 2417-2434 (2000).