Published in final edited form as:

Phys Med Biol. 2013 September 7; 58(17): 6011–6027. doi:10.1088/0031-9155/58/17/6011.

Validation of a power-law noise model for simulating small-scale breast tissue

I. Reiser, A. Edwards, and R. M. Nishikawa

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

I. Reiser: ireiser@uchicago.edu

Abstract

We have validated a small-scale breast tissue model based on power-law noise. A set of 110 patient images served as truth. The statistical model parameters were determined by matching the radially-averaged power-spectrum of the projected simulated tissue with that of the central tomosynthesis patient breast projections. Observer performance in a signal-known exactly detection task in simulated and actual breast backgrounds was compared. Observers included human readers, a pre-whitening observer model and a channelized Hotelling observer model. For all observers, good agreement between performance in the simulated and actual backgrounds was found, both in the tomosynthesis central projections and the reconstructed images. This tissue model can be used for breast x-ray imaging system optimization. The complete statistical description of the model is provided.

1. Introduction

Medical imaging is often performed to determine the presence or absence of a specific disease or diseases, such as the detection of breast cancer using mammography. Thus image quality, in such situations, should be defined as “the ability of an observer performing a task”. Task performance, then, becomes an objective measure of image quality and it can be used as a figure-of-merit for imaging system evaluation and optimization.

It is now well known that normal patient anatomy does affect observer performance. In particular, the impact of the normal breast parenchyma on tumor detectability has been found to be the primary limitation of the detection of tumors on mammograms, rather than quantum noise (Burgess et al., 2001). Park *et al.* have shown that task-performance trends in tomosynthesis are different for uniform and structured backgrounds (Park et al., 2010). Thus, faithful representation of the normal anatomy is essential for proper systems optimization.

With the development of 3D breast imaging systems (Chen and Ning, 2003; Poplack et al., 2007; Rafferty, 2007; Lindfors et al., 2008), there is a need for realistic 3D breast models that produce realistic backgrounds for system optimization through simulation. To address this need, a number of breast phantoms have been developed. Several models are derived from patient data such as dedicated breast CT images (Li et al., 2009) or mastectomy specimen (O'Connor et al., 2013). The drawback of using patient data is that the number and variation of models that can be created is limited by the available patient data, and that it may be difficult to create models with high spatial resolution, depending on the resolution and noise properties of the underlying patient image. Statistically defined phantoms do not have this limitation, but they may lack realism (Bakic et al., 2002a,b, 2003; Bliznakova et al., 2003; Ma et al., 2009; Chen et al., 2011; Pokrajac et al., 2012; Lau et al., 2012).

The breast tissue model evaluated in this work is based on power-law noise. This fundamental model has been proposed previously (Burgess et al., 2001) and is being used in a number of breast phantoms (Chen et al., 2012) including our own (Lau et al., 2012), and has been used in several tomosynthesis optimization studies (Reiser and Nishikawa, 2010; Richard and Samei, 2010a,b).

The purpose of this work was to validate a small-scale breast tissue model for simulating x-ray breast imaging, such as mammography or tomosynthesis. To validate the breast tissue model, statistical properties, such as gray value mean and gray value variance, as well as power-law exponent and magnitude of the radially averaged periodogram, were compared to those of actual breast backgrounds. Further, human and model observer performance in the task of detecting an exactly known signal in the simulated tissue backgrounds was compared to that in actual breast backgrounds, both for projections and reconstructed tomosynthesis images. To our knowledge, this work presents the first rigorous validation of an anthropomorphic breast phantom in terms of task performance.

2. Methods

2.1. Tomosynthesis system

The patient images used in this study were acquired on a tomosynthesis prototype, which is briefly described here. More detail can be found in (Wu et al., 2004). This tomosynthesis unit, which was located at the Massachusetts General Hospital, acquired 11 projection views over a 50° angular span. With the x-ray source at 0°, the source-to-detector distance was 66 cm. During the tomosynthesis scan, the x-ray source rotated about a pivoting point located at approximately 22 cm above a stationary detector. The detector pixel size was 100 $\mu\text{m} \times 100 \mu\text{m}$.

Tomosynthesis projections were reconstructed using an iterative maximum-likelihood expectation-maximization (ML-EM) algorithm with 8 iterations developed in-house (Sidky et al., 2009). Voxel dimension was 100 $\mu\text{m} \times 100 \mu\text{m}$ in-plane and 1 mm out-of-plane.

2.2. Actual breast images

Clinical tomosynthesis image data from 110 patient breasts was used in this study, 76 of which had no abnormal findings, and 24 had benign masses. The breast density distribution in this population in the 4 BI-RADS categories was as follows: 18% category I, 38% cat. II, 34% cat. III, 10% cat. IV.

For each patient breast, the central projection image was used, along with a reconstructed tomosynthesis image. The projection image was a log-transform of the “for-processing” image, which is also used for image reconstruction Reiser and Nishikawa (2010).

A research specialist with over 10 years of experience in mammography (A.E.) manually identified 384×384 pixel regions-of-interest (ROIs) in the central tomosynthesis projection image such that (1) ROIs did not contain any artifacts such as benign calcifications or benign masses; (2) ROIs were located within the uniform thickness region of the breast, i.e., at a distance from the skin line such that no intensity drop-off due to the breast edge was apparent; (3) the pectoralis muscle was not included in any ROI. The overlap between ROIs was at most 50% and ROIs were positioned so as to maximize the number of ROIs that could be extracted from a given breast image.

A total of 613 ROIs were extracted from the central tomosynthesis projection, along with the corresponding ROI from the reconstructed tomosynthesis slice at a depth of 1.6 cm above

the breast support. This depth was chosen because it was at the center of the thinnest breast in the data set (which had a compressed thickness of 3.2 cm).

2.3. Simulated breast tissue

Breast tissue was simulated using a power-law noise model similar to what we have described previously (Reiser and Nishikawa, 2010). Power-law noise does not have a defined correlation length, and has an appearance similar to breast tissue in a projection mammogram. In our tissue model, a power-law noise volume is initialized in the spatial frequency domain as a complex array $V(f_x, f_y, f_z)$ by combining a power-law magnitude with a random phase,

$$V(f_x, f_y, f_z) = \begin{cases} c_0 & \text{if } |f|=0 \\ c \cdot |f|^{-\beta/2} \cdot \exp(-i\tilde{\theta}) & \text{if } 0 < |f| \leq f_0 \\ 0 & \text{if } |f| > f_0 \end{cases} \quad (1)$$

where $\tilde{\theta} = (-\pi, \pi]$ and β is the power-law exponent that determines texture (Burgess and Judy, 2007). The division by two in the power-law exponent is for consistent notation, since β is often used to describe the exponent of the power-spectrum $W(f)$, given in Eq. 3, which involves the squared magnitude of the Fourier-transform (Burgess et al., 2001; Burgess and Judy, 2007; Metheany et al., 2008). Further, f_0 is a high-frequency cutoff to limit the detail-size of the simulated tissue and c and c_0 are arbitrary constants (as explained in more detail below).

An inverse 3D Fourier transform of $V(f_x, f_y, f_z)$ yields a gray-scale volume, which is converted to a binary volume by applying a threshold, as suggested by Abbey and Boone (Abbey and Boone, 2008). Modeling breast tissue as a binary distribution is based on the observation that breast tissue is primarily composed of two tissue types, namely adipose and fibroglandular tissues. The gray-value threshold was chosen such that the ratio of the total number of voxels in the two classes is g_F . Finally, each voxel is assigned the linear x-ray attenuation coefficient of their respective tissue class μ_{adi} (adipose tissue) or μ_{gland} (fibroglandular tissue).

Simulated breast tissue was generated as $256 \times 256 \times 256$ volumes with isotropic voxels ($200 \mu\text{m}$ side length). The simulated breast tissue was embedded into uniform tissue slab ($\mu = \mu_{adi}$) of 5.5 cm thickness and was aligned so that the x-ray source trajectory was along one of its sides, resulting in a half cone-beam projection.

This simulated tissue was projected onto a 2D plane with $100 \mu\text{m} \times 100 \mu\text{m}$ dels (detector elements) using the tomosynthesis geometry described in Sec. 2.1. Note that the detail size of the tissue structure is determined by the cutoff frequency $f_0 = 20 \text{cm}^{-1}$, which was chosen to be less than the Nyquist frequency of a $200 \mu\text{m}$ voxel ($f_{Ny} = 25 \text{cm}^{-1}$), which reduces projection artifacts observed in the previous version of the tissue model (Reiser and Nishikawa, 2010) that did not band-limit the power-law noise. The resulting pixel values represent line integrals through the simulated tissue, $p(x, y) = \int_{ray} \mu(r(\vec{r})) d\vec{r}$. Quantum noise was introduced by replacing each pixel's gray-value, $n(x, y)$, with a random sample $\tilde{n}(x, y)$ drawn from a Poisson distribution with mean

$$\bar{n}(x, y) = N_0 \exp[-p(x, y)], \quad (2)$$

where N_0 is the number of photons incident on the breast. The value of N_0 was chosen so as to match the high-frequency portion ($> 1 \text{cyc/mm}$) of the mammographic power-spectrum.

Finally, detector blur was simulated by convolving the photon distribution $\tilde{h}(x, y)$ with a Gaussian of width g . From each simulated center projection and tomosynthesis volume, a single 384×384 region-of-interest was extracted.

Values for a subset of model parameters was guided by published research ($\sigma, g, f_0, \mu_{gland} - \mu_{adi}, \mu_{adi}$), described in more detail below. The remaining model parameters (σ, N_0, g) were chosen so as to match the radially averaged power-spectrum of the simulated projections to those of actual breast ROIs. The power spectrum $W(\vec{f})$ was computed using

$$W(\vec{f}) = \langle |\mathcal{F}\{\mathcal{R} - \langle \mathcal{R} \rangle\}|^2 \rangle \quad (3)$$

where $\mathcal{R}$ was a 384×384 pixel region-of-interest that was multiplied with a Hann window to alleviate spectral leakage (Percival and Walden, 1993). $\mathcal{F}\{\cdot\}$ is the Fourier transform and $\langle \cdot \rangle$ denotes average.

Within this model, σ (texture) and g (glandular fraction) are random variables. Based on our previous studies (Engstrom et al., 2009; Reiser et al., 2011), σ is sampled from a Gaussian distribution with a standard deviation $\sigma = 0.3$ and $\mu = 3.3$, where the latter was chosen so as to match the power-spectrum of the clinical projections.

The glandular fraction is modeled by a beta distribution (Weisstein, n.d.), which produces a continuous random variable in the interval (0,1) and is therefore suited to represent ratios. The two distribution parameters ($a=3, b=15$) produce a mean and standard deviation of 0.167 and 0.09, respectively, which is similar to volumetric glandular fractions measured in dedicated breast CT images (mean = 0.14, $\sigma = 0.103$) (Yaffe et al., 2009).

The high-frequency cutoff $f_0 = 20\text{cm}^{-1}$ limits the detail size, that is being simulated by power-law noise, to about 1mm. The constant c_0 avoids division by zero at $f_r = 0$ and affects the mean of the power-law noise volume in the spatial domain. Since the tissue volume voxel values are assigned after binarization, the value of c_0 is arbitrary. Likewise, c controls the gray scale of the inverse Fourier transform of $V(f)$ before binarization, its value is therefore arbitrary. Tissue contrast $\mu_{gland} - \mu_{adi} = 0.1\text{cm}^{-1}$ is chosen according to Johns and Yaffe (Johns and Yaffe, 1987), and μ_{adi} was set to 0.5cm^{-1} . Since the contrast (and thereby the magnitude the power spectrum) is determined by the difference in attenuation between adipose and glandular tissues, actual values used are unimportant, except that they will determine the average pixel value in the projections and reconstructed images.

Figure 1 shows the radially averaged power spectrum estimated from the 613 clinical projection ROIs, and from 512 simulated projection ROIs. The breast model parameters used to create this set of simulated tissue volumes are listed in Table 1.

The agreement between both spectra is excellent. These 512 breast tissue models were used for statistical analyses and in the observer experiments in the remainder of this study. Figures 2 and 3 show examples of noisy projections and reconstructed tomosynthesis images of this simulated breast tissue.

2.4. Signals and test images

Test images for the observer studies were generated by adding computer-generated signals to actual breast or simulated tissue backgrounds. As in a previous study (Reiser and Nishikawa, 2010), signals were 3D designer nodules, which are spherically symmetric with a $(1 - r^2)$ intensity profile (r is the signal radius). The projection of such a signal with amplitude A is

$$s(r) = \frac{4}{3}AR \left[1 - \left(\frac{r}{R} \right)^2 \right]^{3/2} \pi_R. \quad (4)$$

where R is the signal radius and π_R is a rect function with $\pi_R = 1$ for $|r| < R$ and 0 elsewhere. The use of a 3D signal in conjunction with an analytical projection allows signal projections to be directly computed from Eq. 4, rather than from a discrete signal, thereby eliminating artifacts due to the discrete-to-discrete projection. In this study, signals had radii of 0.025 cm, 0.05 cm, 0.1 cm, 0.2 cm, 0.4 cm, and 0.8 cm.

Signal templates for projections were created by projecting the designer nodules using the tomosynthesis geometry described above. Tomosynthesis signal templates were created by reconstructing signal projections added to a uniform background, which was subsequently subtracted from the reconstruction. We have verified that these signal templates are identical to those produced if signals are added to breast projections rather than uniform backgrounds. This indicates that our reconstruction algorithm exhibits a linear behavior. No noise or blur was added to signal projections, because the purpose of these experiments was to compare performance in simulated breast tissue backgrounds to that in real breast backgrounds, rather than to determine performance of a tomosynthesis system per se. For observer studies with actual breast images, where ROIs could be centered anywhere within the breast, signal locations were matched to that of the corresponding ROI center. That is, signals were generated at multiple locations corresponding to ROI centers in the tomosynthesis image.

2.5. Human performance

Human performance in detecting an exactly-known signal was determined through 2-alternative forced choice (2-AFC) experiments. The screen layout of the 2-AFC experiments has been described previously (Reiser et al., 2010). Briefly, in each trial, the observer is presented with a pair of test images, one of which contained a signal that is to be identified by the observer. A high-contrast copy of the signal is shown along with whisker cues to show the two possible signal locations. This is done to ensure that the observer only considers the possibility of a signal being present at an exact location. Uncertainty as to the location of the signal can degrade human performance in a signal-known-exactly detection task (Reiser and Nishikawa, 2006). The images were rescaled linearly to 8 bit and an offset was added to each image such that the mean gray value of the test image center, within a region of 90 pixels from the ROI center, was equal to 128. The scaling factor was held constant for a given condition. The scaling factor (i.e., the display window width) has little effect on observer performance (Burgess et al., 2001).

For each experimental condition, the percentage of correct responses (PC) was computed from 128 trials and converted to d'_A using

$$d'_A = 2 \cdot \text{erf}^{-1}(2 \cdot \text{PC} - 1). \quad (5)$$

For each signal size, d'_A was obtained for three signal amplitudes and a linear fit was used to estimate the amplitude threshold required to obtain $d'_A = 2$. During the experiments, signal amplitudes were chosen so as to obtain a PC in the vicinity of 0.8, where the power of 2-AFC experiments is maximized.

Three readers participated in the experiments. Two readers were medical physicists, and one was a research specialist. All readers had previously participated in a pilot study (Reiser et al., 2010) and were experienced in mammography research.

2.6. Observer Models

Signal-known exactly (SKE) detection performance was estimated using two approximations to the Hotelling observer, one a “pre-whitening” observer (PWO), the second a channelized Hotelling observer with Laguerre-Gauss channels (LG-CHO) (Barrett and Myers, 2004; Burgess and Judy, 2007).

PWO detection performance was estimated using

$$d_{PW}'^2 = \int_{f>0} \frac{|S(\vec{f})|^2}{W(\vec{f})} d\vec{f}, \quad (6)$$

where $W(\vec{f})$ is the background power-spectrum defined in Eq. (3) and $S(\vec{f})$ is the Fourier-spectrum of the signal. For multiple signal locations, as was the case for mammographic ROIs, $S(\vec{f})$ is the Fourier-spectrum of the average signal. For two-dimensional images, this integral can be carried out as 2D integration, $f = (f_x, f_y)$, or, assuming rotational symmetry, in polar coordinates as

$$d_{PW_s}'^2 = 2\pi \int_{f_r} \frac{|S_r(f_r)|^2}{W_r(f_r)} f_r df_r, \quad (7)$$

where S_r and W_r are radial averages of S and W , respectively.

Within the CHO, the image data is represented with respect to basis functions (so-called channels). This formalism is briefly outlined here: First, the $N \times N$ image data is reshaped (“vectorized”) into a $[1 \times N^2]$ column vector $\mathbf{g}$. Channelizing the image data consists of a basis transform

$$\mathbf{g}_u = \mathbf{U}^t \mathbf{g}. \quad (8)$$

where each column of the channel matrix $\mathbf{U}$ contains one basis function and $\mathbf{U}$ has $n_{chan} \times N^2$ elements where n_{chan} is the number of channels. In this study, $n_{chan} = 10$, and Laguerre-Gauss (LG) basis functions $\Phi_n(r)$ were used,

$$\Phi_n(r) = \frac{2\sqrt{\pi}}{a} \exp\left[-\frac{\pi r^2}{a^2}\right] L_n\left(\frac{2\pi r^2}{a^2}\right) \quad (9)$$

where $L_n(\cdot)$ are Laguerre polynomials of order n and the support $a = 0.85 * r_{sig}$.

LG-CHO detection performance is then estimated as

$$d_{CHO}'^2 = \Delta \mathbf{g}_u^t \mathbf{K}_u^{-1} \Delta \mathbf{g}_u \quad (10)$$

where $\mathbf{K}_u$ is the covariance matrix, $\mathbf{K}_u = \langle \mathbf{g}_u - \langle \mathbf{g}_u \rangle \rangle \langle \mathbf{g}_u^t - \langle \mathbf{g}_u^t \rangle \rangle$ (Beutel et al., 2000, Chapt. 11). For signals centered at different locations, the signal shape could exhibit

variations due to the projection geometry. In this case, the average signal $\bar{\mathbf{g}}_{\mathbf{u}}$ replaces $\mathbf{g}_{\mathbf{u}}$ in Eq. 10, and $d_{CHO}^{\prime 2}$ is the performance of the ideal linear observer (Barrett and Myers, 2004).

Note that both observer models provide estimates for ideal observer performance (Barrett and Myers, 2004). However, the two observer models make assumptions: The pre-whitening observer model assumes that the image data are statistically stationary, in which case the Fourier-space representation of the covariance matrix is diagonal with diagonal elements given by the power spectrum. The channelized Hotelling observer assumes that the channel functions form a complete basis that can represent the signal.

For comparison with human performance, a constant efficiency of human observers was assumed,

$$d_{human}^{\prime 2} = \eta d_{ideal}^{\prime 2}, \quad (11)$$

where $\eta = 0.69$ was used throughout this study. This value is similar to what has been reported in the literature (Burgess et al., 1981, 2001; Burgess and Judy, 2007).

3. Results

3.1. Comparison of ROI statistics

Figure 4 compares the radially averaged power-spectra in reconstructed tomosynthesis images, for actual breast and simulated tissue ROIs. Excellent agreement of both spectra is observed, similar to what was seen in the projection domain (Fig. 1). This verifies that the agreement between radially averaged power-spectra of both backgrounds persists in the reconstructed tomosynthesis images. Note that all ROIs were extracted from a single tomosynthesis slice at 1.6 cm above the breast support.

ROI statistics were analyzed by computing ROI mean, ROI variance and periodogram power-law exponent and magnitude and $\log(c)$ (Engstrom et al., 2009; Reiser et al., 2011). Mean and standard deviations of the resulting distributions for breast projections and simulated breast tissue projections are listed in Table 2. Note that isotropic symmetry was assumed for periodogram analysis.

Previously, we have observed a strong correlation of $\rho = 0.924$ between $\log(c)$ and σ in breast ROIs (Fig. 5a) (Engstrom, 2007). A similarly strong correlation ($\rho = 0.90$) was observed for the simulated tissue projections (Fig. 5b). The origin of this correlation is illustrated in Fig. 6: a large σ value produces a coarser texture (Fig. 6a,c), compared to a smaller σ value (Fig. 6b,d), both in the original 3D volume (Fig. 6a,b) and the corresponding projections (Fig. 6c,d). The large-scale variations produced by higher values of σ produce more spectral density at low frequencies, than those produced by smaller σ values (Fig. 6e,f). Note that σ in a 3D volume and the corresponding projection is expected to be approximately equal due to the Fourier slice theorem (Kak and Slaney, 1988; Gong et al., 2006; Metheany et al., 2008).

3.2. Comparison of observer performance

Human observer performance is shown in Fig. 7 by plotting the signal amplitude threshold required by each observer to reach a detection performance of $d_A' = 2$ (Burgess et al., 2001), scaled by signal radius. This scaling converts the amplitude threshold for the 3D signal to the corresponding threshold for a projected 2D signal. In this representation, the threshold curve exhibits a minimum where the dominant noise source transitions from quantum noise

(smaller signals) to anatomic noise (larger signals). In this figure, the lines represent average human performance. Overall, human observer performance in the actual and simulated breast backgrounds is very similar, and, importantly, performance trends as a function of signal size are well reproduced in the simulated backgrounds. Further, a Wilcoxon signed-rank test failed to show statistical significance for the difference of average human observer performance in the two backgrounds (projections: $p=0.16$, slices: $p=0.31$) (Corder and Foreman, 2009).

Figure 8 shows model observer performance in the actual breast and simulated breast tissue backgrounds. Average human observer performance is included as a reference. Model observer performance was scaled by an efficiency factor of $\eta = 0.69$ for all observer models and experimental conditions (see Eq. 11). In human observer experiments with actual breast backgrounds, signal locations were matched with the ROI centers. Thus, for model observer performance calculations, signals from multiple locations were averaged (solid line), or, a single signal located at the center of the power-law noise volume was used (dotted line). For simulated tissue backgrounds, all ROIs were extracted from the same location relative to the tomosynthesis system, so there was only a single signal location.

Figure 8 compares model observer performance with human performance. The top panel shows PWO performance estimated using 2D integration (Eq. 6). Overall, human performance is well predicted for both real and simulated backgrounds. For small signals, the PWO overestimates human performance, as human observers require a higher amplitude threshold for detection than what is predicted by the model. This is true for both projections and reconstructed tomosynthesis images. In addition, a large difference in PWO performance for small signals in actual breast backgrounds is observed when a single signal is used, or when the average signal is used. Further, the PWO predicts a small performance difference in actual breast and simulated backgrounds for larger signal sizes, that is not observed in the human performance data.

The bottom panel shows performance of the LG-CHO, which also predicts human performance well and shows approximately similar performance in both background types for all signal sizes. Again, the use of the average signal when estimating LG-CHO performance for small signals in mammographic backgrounds dramatically reduces detectability (which results in an increase in amplitude threshold as the signal becomes “harder” to detect).

4. Discussion

For both observer models, use of the average signal increased amplitude threshold for small signals. This effect was most pronounced for the LG-CHO observer in the reconstructed slices. This is due to the location dependence of the reconstructed signal across the reconstructed slice. Initially, an identical 3D signal was placed at the center location of each ROI in the breast volume, from where it was projected and then reconstructed. The actual appearance of the signal in the reconstructed tomosynthesis image depends on its location relative to the x-ray source path. This is expected and was the reason for matching signal and ROI location in the first place, because differences due to mismatched geometry are noticeable by human observers, in particular, if the 3D tomosynthesis volume is displayed (Reiser and Nishikawa, 2012). To make this more clear, one can inspect a tomosynthesis plane perpendicular to the detector and observe that objects are blurred and slightly tilted towards the center of the x-ray source path. In addition to the geometry mismatch, sampling produces differences in the observed signals, which also particularly affects small signals. Figure 9 shows projections and reconstructions of single signals (a,c), and the respective average signal (b,d), along with the corresponding channelized signal (e–h). Of all signals

shown, the average reconstructed signal (Fig. 9d) is least symmetric, and therefore likely the most difficult to be represented by the radially symmetric LG channel functions.

Overall, performance of both observer models in the actual breast backgrounds is similar. However, the PWO is a frequency-domain observer model that assumes wide-sense stationarity of the image statistics (Bendat and Piersol, 2010). The LG-CHO makes no such assumption (but assumes radial symmetry of the system). Our results indicate that the stationarity assumption may be justified when averaging across locations. Future studies will investigate this in more detail.

In Fig. 8, amplitude thresholds for signal-known exactly detection performance are shown for tomosynthesis projections and reconstructed tomosynthesis slices. However, caution is warranted when using these data to infer performance differences of 2D mammography and tomosynthesis: (i) the tomosynthesis projections used in this study differ from mammograms because the x-ray exposure is much lower and no anti-scatter grid is used; (ii) higher energy x-ray spectra tend to be used for tomosynthesis than for mammography (Wu et al., 2004; Feng and Sechopoulos, 2012). (iii) signals were not degraded by focal spot and detector blur, as they would be in an actual imaging system; (iv) the effect of adding a signal to a projection image rather than embedding it into the tissue structure, has not been investigated. These factors likely affect smaller signals more than larger ones. Comparison studies that account for these differences have been performed and found that for masses, contrast levels in tomosynthesis were by a factor 2–4 lower than in mammography, but for microcalcifications, only a small or no difference was found (Ruschin et al., 2007; Timberg et al., 2010). Note also that these performance differences depend on the tomosynthesis geometry, most notably the scan angle (Reiser and Nishikawa, 2010; Gang et al., 2010; Richard and Samei, 2010b; Sechopoulos and Ghetti, 2009).

Our previous studies have shown that mammographic power-spectra are not radially symmetric, but exhibit elliptical symmetry caused by the “directionality” of fibroglandular tissues, which tend to radiate towards the nipple (Reiser et al., 2011). The PWO is sensitive to this symmetry, which causes a performance difference between the two background types that can be observed for larger signals. This difference is not observed for human observers. Enforcing radial symmetry by integrating Eq. 6 in polar coordinates removes this difference. Human observers may not be sensitive to directionality present in the background when tasked with the detection of a spherically symmetric signal.

Tomosynthesis is a quasi-3D breast imaging modality that produces a breast volume that partially resolves breast structure. Yet, in this study, human and model observer performance were determined for 2D tomosynthesis images only. Other researchers report 3D performance measures for tomosynthesis (Hu and Zhao, 2011; Gang et al., 2011). We have performed human observer studies to investigate whether readers’ performance increases when the entire volume is shown, rather than the 2D image only, but so far we failed to find clear evidence that human observers can retrieve additional information from the volume, for the task of detecting exactly known signals in a 50°-scan tomosynthesis system (Reiser and Nishikawa, 2012).

A simplification implemented in our study was that for our tissue model, thickness was kept constant, while there is variation in patient breast thickness. Also, the tissue model did not include any skin. In addition, the physics model was simplistic in that x-ray energy spectra and scatter were ignored, and that the detector had perfect efficiency. We chose these simplifications because the purpose of this work was to investigate whether this tissue model is equivalent to actual breast tissue in terms of task-performance in projections and in reconstructed images. Our goal was to have few parameters that fully describe our tissue

model, which are listed in Table 1). We have co-developed a statistical breast model by embedding this small-scale tissue into a breast phantom that includes large-scale breast features such as skin, variable breast volume, density, and texture (Lau et al., 2012).

On the other hand, we found that ensembles of simulated tissue volumes, whose projections had equal radially averaged power-spectra, resulted in equal task-performance both in the projections (as one might expect) but also in the reconstructed volume. This model was designed and validated for mammography and breast tomosynthesis and is hoped to be useful for assessing task-performance of tomosynthesis acquisition strategies and reconstruction algorithms. We will explore in the future whether this model is applicable to other breast imaging geometries such as CT.

5. Conclusion

We presented a statistical model for small-scale breast tissue, representative of a population of 110 patients. We have validated this model in terms of detection task-performance, both by human observers and model observers. Our results showed that there is very little difference in measured performance of human observers reading either projections or reconstructed tomosynthesis images (Figure 7). We also tested the performance of two different observer models, a pre-whitening observer (PWO), and a channelized Hotelling observer with Laguerre-Gauss channels (LG-CHO) on the real and simulated backgrounds (Figure 8). With the exception of LG-CHO for small objects, there was little difference in measured performance with either backgrounds. We therefore conclude that simulated backgrounds based on power-law noise is a suitable replacement for actual breast backgrounds, at least for tasks that do not involve search. Further research is needed to validate the simulated backgrounds for free search experiments.

Acknowledgments

The authors wish to thank Dr. Beverly Lau for participating in the reader study, and Dr. Daniel Kopans and Richard Moore from the Massachusetts General Hospital for providing them with patient tomosynthesis images. They also thank Dr. Art Burgess for insightful discussions on power-law noise and observer performance studies. This work was funded in part by NCI R21 EB008801. RM Nishikawa is a shareholder in, a consultant to, and receives royalties from Hologic, Inc. He is also a consultant to iCAD, Inc.

References

- Abbey, CK.; Boone, JM. Proc IWDIM. Springer; 2008. An Ideal Observer for a Model of X-Ray Imaging in Breast Parenchymal Tissue; p. 393-400. Vol. LNCS 5116
- Bakic P, Albert M, Brzakovic D, Maidment A. Mammogram synthesis using a 3D simulation. II. Evaluation of synthetic mammogram texture. Medical Physics. 2002a; 29(9):2140. [PubMed: 12349936]
- Bakic PR, Albert M, Brzakovic D, Maidment ADA. Mammogram synthesis using a 3D simulation. I. Breast tissue model and image acquisition simulation. Medical Physics. 2002b; 29(9):2131. [PubMed: 12349935]
- Bakic PR, Albert M, Brzakovic D, Maidment ADA. Mammogram synthesis using a three-dimensional simulation. III. Modeling and evaluation of the breast ductal network. Medical Physics. 2003; 30(7): 1914. [PubMed: 12906210]
- Barrett, HH.; Myers, KJ. Foundations of Image Science. Wiley-Interscience; Hoboken, New Jersey: 2004.
- Bendat, J.; Piersol, A. Random data: Analysis and measurement procedures. John Wiley & Sons; Hoboken, New Jersey: 2010.
- Beutel, J.; Kundel, HL.; Van Metter, RL., editors. Handbook of Medical Imaging: Physics and Psychophysics. SPIE Press; Bellingham, WA: 2000.

- Bliznakova K, Bliznakov Z, Bravou V, Kolitsi Z, Pallikarakis N. A three-dimensional breast software phantom for mammography simulation. *Phys Med Biol*. 2003; 48(22):3699–719. [PubMed: 14680268]
- Burgess AE, Jacobson FL, Judy PF. Human observer detection experiments with mammograms and power-law noise. *Medical Physics*. 2001; 28(4):419. [PubMed: 11339738]
- Burgess AE, Judy PF. Signal detection in power-law noise: effect of spectrum exponents. *JOSA A*. 2007; 24(12):B52–60. [PubMed: 18059914]
- Burgess A, Wagner R, Jennings R, Barlow H. Efficiency of human visual signal discrimination. *Science*. 1981; 214:93–94. [PubMed: 7280685]
- Chen B, Shorey J, Saunders RS, Richard S, Thompson J, Nolte LW, Samei E. An anthropomorphic breast model for breast imaging simulation and optimization. *Acad Radiol*. 2011; 18(5):536–46. [PubMed: 21397528]
- Chen L, Abbey CK, Nosrateih A, Lindfors KK, Boone JM. Anatomical complexity in breast parenchyma and its implications for optimal breast imaging strategies. *Medical physics*. 2012; 39(3):1435–1441. [PubMed: 22380376]
- Chen Z, Ning R. Why should breast tumour detection go three dimensional? *Phys Med Biol*. 2003; 48(14):2217–28. [PubMed: 12894980]
- Corder, GW.; Foreman, DI. *Nonparametric Statistics for Non-Statisticians*. Wiley Hoboken; New Jersey: 2009.
- Dobbins JT III, Godfrey DJ. Digital x-ray tomosynthesis: current state of the art and clinical potential. *Phys Med Biol*. 2003; 48(19):R65–106. [PubMed: 14579853]
- Engstrom, E. Master's thesis. Royal Institute of Technology; Stockholm, Sweden: 2007. Comparison of the breast tissue power spectra for x-ray imaging of the breast.
- Engstrom E, Reiser I, Nishikawa R. Comparison of power spectra for tomosynthesis projections and reconstructed images. *Medical Physics*. 2009; 36(5):1753. [PubMed: 19544793]
- Feng SSJ, Sechopoulos I. Clinical Digital Breast Tomosynthesis System: Dosimetric Characterization. *Radiology*. 2012; 263(1):35–42. [PubMed: 22332070]
- Gang GJ, Lee J, Stayman JW, Tward DJ, Zbijewski W, Prince JL, Siewerdsen JH. Analysis of Fourier-domain task-based detectability index in tomosynthesis and cone-beam CT in relation to human observer performance. *Medical Physics*. 2011; 38(4):1754–1768. [PubMed: 21626910]
- Gang GJ, Tward DJ, Lee J, Siewerdsen JH. Anatomical background and generalized detectability in tomosynthesis and cone-beam CT. *Medical Physics*. 2010; 37(5):1948–1965. [PubMed: 20527529]
- Gong X, Glick SJ, Liu B, Vedula AA, Thacker S. A computer simulation study comparing lesion detection accuracy with digital mammography, breast tomosynthesis, and cone-beam CT breast imaging. *Medical Physics*. 2006; 33(4):1041. [PubMed: 16696481]
- Hu YH, Zhao W. The effect of angular dose distribution on the detection of microcalcifications in digital breast tomosynthesis. *Medical Physics*. 2011; 38(5):2455. [PubMed: 21776781]
- Johns PC, Yaffe MJ. X-ray characterisation of normal and neoplastic breast tissues. *Phys Med Biol*. 1987; 32(6):675–95. [PubMed: 3039542]
- Kak, AC.; Slaney, M. *Principles of Computerized Tomographic Imaging*. IEEE press New York; New York: 1988.
- Lau BA, Reiser I, Nishikawa RM, Bakic PR. A statistically defined anthropomorphic software breast phantom. *Medical Physics*. 2012; 39(6):3375–3385. [PubMed: 22755718]
- Li CM, Segars WP, Tourassi GD, Boone JM, Dobbins JT. Methodology for generating a 3D computerized breast phantom from empirical data. *Medical Physics*. 2009; 36(7):3122. [PubMed: 19673211]
- Lindfors KK, Boone JM, Nelson TR, Yang K, Kwan ALC, Miller DF. Dedicated Breast CT: Initial Clinical Experience. *Radiology*. 2008; 246(3):725–733. [PubMed: 18195383]
- Ma AKW, Gunn S, Darambara DG. Introducing DeBRA: a detailed breast model for radiological studies. *Phys Med Biol*. 2009; 54(14):4533–45. [PubMed: 19556683]
- Metheany KG, Abbey CK, Packard N, Boone JM. Characterizing anatomical variability in breast CT images. *Medical Physics*. 2008; 35(10):4685. [PubMed: 18975714]

- O'Connor JM, Das M, Didier CS, Mah'd M, Glick SJ. Generation of voxelized breast phantoms from surgical mastectomy specimen. *Medical Physics*. 2013 (in press).
- Park S, Jennings R, Liu H, Badano A, Myers K. A statistical, task-based evaluation method for three-dimensional x-ray breast imaging systems using variable-background phantoms. *Medical Physics*. 2010; 37(12):6253. [PubMed: 21302782]
- Percival, DB.; Walden, AT. *Spectral Analysis for Physical Applications*. Cambridge University Press; Cambridge: 1993.
- Pokrajac DD, Maidment A, DaBakic PR. Optimized generation of high resolution breast anthropomorphic software phantoms. *Medical physics*. 2012; 39(4):2290–2302. [PubMed: 22482649]
- Poplack SP, Tosteson TD, Kogel CA, Nagy HM. Digital breast tomosynthesis: initial experience in 98 women with abnormal digital screening mammography. *AJR*. 2007; 189(3):616–23. [PubMed: 17715109]
- Rafferty EA. Digital mammography: novel applications. *Radiologic Clinics of North America*. 2007; 45(5):831–843. [PubMed: 17888772]
- Reiser I, Lee S, Little K, Nishikawa RM. Toward validation of a 3D structured background model for breast imaging. *Proc SPIE*. 2010; 7627:762716.
- Reiser I, Lee S, Nishikawa RM. On the orientation of mammographic structure. *Medical Physics*. 2011; 38(10):5303–5307. [PubMed: 21992347]
- Reiser I, Nishikawa RM. Identification of simulated microcalcifications in white noise and mammographic backgrounds. *Medical Physics*. 2006; 33(8):2905. [PubMed: 16964867]
- Reiser I, Nishikawa RM. Task-based assessment of breast tomosynthesis: Effect of acquisition parameters and quantum noise. *Medical Physics*. 2010; 37(4):1591–1600. [PubMed: 20443480]
- Reiser I, Nishikawa RM. Signal-known exactly detection performance in tomosynthesis: does volume visualization help human observers? *Proc SPIE*. 2012; 8318:83180K.
- Richard S, Samei E. Quantitative breast tomosynthesis: From detectability to estimability. *Medical Physics*. 2010a; 37(12):6157–6168. [PubMed: 21302772]
- Richard S, Samei E. Quantitative imaging in breast tomosynthesis and CT: Comparison of detection and estimation task performance. *Medical Physics*. 2010b; 37(6):2627–2637. [PubMed: 20632574]
- Ruschin M, Timberg P, Bath M, Hemdal B, Svahn T, Saunders RS, Samei E, Andersson I, Mattsson S, Chakraborty DP, Tingberg A. Dose dependence of mass and microcalcification detection in digital mammography: Free response human observer studies. *Medical Physics*. 2007; 34(2):400–407. [PubMed: 17388156]
- Sechopoulos I, Ghetti C. Optimization of the acquisition geometry in digital tomosynthesis of the breast. *Medical Physics*. 2009; 36(4):1199. [PubMed: 19472626]
- Sidky EY, Pan X, Reiser IS, Nishikawa RM, Moore RH, Kopans DB. Enhanced imaging of microcalcifications in digital breast tomosynthesis through improved image-reconstruction algorithms. *Medical Physics*. 2009; 36(11):4920. [PubMed: 19994501]
- Timberg P, Båth M, Andersson I. In-plane visibility of lesions using breast tomosynthesis and digital mammography. *Medical physics*. 2010; 37(11):5618–5626. [PubMed: 21158273]
- Weisstein, EW. Beta Distribution. n.d.
- Wu T, Moore RH, Rafferty EA, Kopans DB. A comparison of reconstruction algorithms for breast tomosynthesis. *Medical Physics*. 2004; 31(9):2636. [PubMed: 15487747]
- Yaffe MJ, Boone JM, Packard N, Alonzo-Proulx O, Huang SY, Peressotti CL, Al-Mayah A, Brock K. The myth of the 50–50 breast. *Medical Physics*. 2009; 36(12):5437. [PubMed: 20095256]

Figure 1. Radially averaged power-spectra estimated from ROIs extracted from actual breast projections (n=613) and simulated breast tissue projections (n=512).

Figure 2.
Examples of noisy projections of simulated breast tissue.

Figure 3.
Examples of tomosynthesis slices of simulated breast tissue, reconstructed from noisy simulated projections.

Figure 4. Radially averaged power-spectra of ROIs extracted from reconstructed tomosynthesis images of actual breasts (n=613) and simulated breast tissue (n=512).

Figure 5. Correlation between power-law exponent β and magnitude $\log(c)$ in projections. (a) breast projection ROIs ($n=613$) and (b) simulated tissue projection ROIs ($n=512$).

Figure 6.

(a–b): Cuts through two simulated tissue volumes with different β_{vol} (c–d): Simulated noisy tissue projections (e): Line profiles through the projections. (f): Radially averaged projection periodograms.

Figure 7.

Comparison of human observer detection-task performance in actual breast (open symbols) and simulated breast tissue backgrounds (closed symbols). The three symbols represent each of the three observers participating in the experiments. The lines represent amplitude thresholds averaged across observers.

Figure 8.

Comparison of human performance (symbols) and model observers performance in actual breast (solid lines) and simulated tissue (dashed line) backgrounds. For actual breast backgrounds, performance was estimated based on the average signal (solid line) or a signal located at the center of the simulated tissue volume (dotted line). Top panel: Pre-whitening observer performance. Bottom panel: Channelized Hotelling observer performance with Laguerre-Gauss basis functions. The symbols indicate average human observer performance.

Figure 9.

Comparison of single and average signals. Top row: Actual signal. Bottom row: Channelized signal, using LG basis function and the same parameters as were used in the performance estimate. (a): Single signal projection. (b): Average projected signal. (c): Single reconstructed signal. (d): Average reconstructed signals. (e–h): Corresponding channelized signals. The same display window was used for all signals.

Table 1

Breast tissue simulation parameters. Note that c and c_0 in Eq. 1 can take any value since the gray values in the volume are determined by μ_{adi} and μ_{gland} .

μ : normal distribution($\mu = 3.3$, $\sigma = 0.3$)
g_f : beta distribution ($a = 3$, $b = 15$)
$\mu_{adi} = 0.5 \text{ cm}^{-1}$
$\mu_{gland} - \mu_{adi} = 0.1 \text{ cm}^{-1}$
$f_0 = 20 \text{ cm}^{-1}$
$N_0 = 1.5 \cdot 10^5 \text{ pixels}^{-1}$
$\delta_g = 0.5 \text{ pixels}$

Table 2

Pixel statistics (mean, variance) and periodogram power-law coefficients (, $\log(c)$) for breast projection ROIs (n=613) and simulated tissue projection ROIs (n=512). Values in parentheses are standard deviation.

	mean	var		$\log(c)$
clinical	2.90(0.39)	6.4e-4(6e-3)	2.6(.39)	-4.37(0.38)
simulation	2.58(0.046)	4.1e-3(2.9e-3)	2.57(0.4)	-4.42(0.46)