

RESEARCH ARTICLE

Human and model observer performance for lesion detection in breast cone beam CT images with the FDK reconstruction

Minah Han, Byeongjoon Kim, Jongduk Baek*

School of Integrated Technology and Yonsei Institute of Convergence Technology, Yonsei University, Incheon, South Korea

* jongdukbaek@yonsei.ac.kr



Abstract

We investigate the detectability of breast cone beam computed tomography images using human and model observers and the variations of exponent, β , of the inverse power-law spectrum for various reconstruction filters and interpolation methods in the Feldkamp-Davis-Kress (FDK) reconstruction. Using computer simulation, a breast volume with a 50% volume glandular fraction and a 2mm diameter lesion are generated and projection data are acquired. In the FDK reconstruction, projection data are apodized using one of three reconstruction filters; Hanning, Shepp-Logan, or Ram-Lak, and back-projection is performed with and without Fourier interpolation. We conduct signal-known-exactly and background-known-statistically detection tasks. Detectability is evaluated by human observers and their performance is compared with anthropomorphic model observers (a non-prewhitening observer with eye filter (NPWE) and a channelized Hotelling observer with either Gabor channels or dense difference-of-Gaussian channels). Our results show that the NPWE observer with a peak frequency of 7 cyc/degree attains the best correlation with human observers for the various reconstruction filters and interpolation methods. We also discover that breast images with smaller β do not yield higher detectability in the presence of quantum noise.

OPEN ACCESS

Citation: Han M, Kim B, Baek J (2018) Human and model observer performance for lesion detection in breast cone beam CT images with the FDK reconstruction. PLoS ONE 13(3): e0194408. <https://doi.org/10.1371/journal.pone.0194408>

Editor: Emil Sidky, University of Chicago Medical Center, UNITED STATES

Received: June 16, 2017

Accepted: February 19, 2018

Published: March 15, 2018

Copyright: © 2018 Han et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was supported by Ministry of Science, ICT and Future Planning (IITP-2017-2017-0-01015); National Research Foundation of Korea (2017M2A2A4A01070302, 2015R1C1A1A01052268, 2017M2A2A6A01019663). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Introduction

Since breast cancer is the second leading cause of cancer-related deaths among women [1], early detection of breast cancer is very important. Mammography is most commonly used for breast cancer screening, but superimposing the breast anatomy structure onto its two-dimensional (2D) mammographic image reduces the accuracy of lesion detection [2, 3]. Breast cone beam computed tomography (CBCT) has been developed with the expectation that it will improve detection performance by reducing tissue superimposition in volumetric images [4–6].

Because of the potential of CBCT for breast imaging, evaluation of the detectability of breast CBCT images has become an important issue to optimize the imaging system. Past studies

Competing interests: The authors have declared that no competing interests exist.

have investigated the effects of several imaging parameters on detectability, such as the X-ray tube voltage [7], scintillator thickness [8], detector pixel size [9], slice thickness [5, 10], and image plane [11]. Reconstruction algorithms producing optimal image quality in breast imaging have been explored [12–14], but the Feldkamp-Davis-Kress (FDK) algorithm [15] is the most widely used technique because of its simplicity, linearity, and computational efficiency. In the FDK reconstruction, optimal image quality can be acquired using various reconstruction filters and interpolation methods during back-projection depending on the task, and therefore, an investigation on the detectability for these parameters could help to optimize the FDK reconstruction in breast imaging.

In a previous study, it was reported that lesion detectability in breast CBCT images can be quantified by β (i.e. the slope of the breast anatomy power spectrum), because smaller β is considered as a predictor of better detectability [16–18]. A more desirable way to evaluate detectability in breast CBCT images is to conduct a human observer study because a human is the subject who makes diagnostic decisions. However, using human observers is time-consuming and expensive, thus it is desirable to use anthropomorphic model observers as surrogates [19–22]. In previous work [11], we evaluated detectability in breast CBCT images using β and model observers, especially for various reconstruction parameters considered in the FDK reconstruction, but without a human observer study.

In this study, we investigated the correlation between human observer and anthropomorphic model observer performance for lesion detection in breast CBCT images with various reconstruction filters and interpolation methods. The relationship between detectability by human observers and β was also investigated. For the investigations, we generated breast volumes with 50% volume glandular fraction (VGF) and a 2mm diameter lesion using computer simulation. Three reconstruction filters: Hanning, Shepp-Logan, and Ram-Lak [23] were considered in the FDK reconstruction, and projection data were filtered with and without Fourier interpolation before back-projection. We conducted signal-known-exactly and background-known-statistically (SKE/BKS) detection tasks. To evaluate detectability, we performed human observer experiments with four-alternative forced choice (4AFC) detection tasks [24]. For the model observer study, we used a non-prewhitening observer with eye filter (NPWE) [25] and a channelized Hotelling observer (CHO) with either Gabor channels (Gabor CHO) [26] or dense difference-of-Gaussian channels (D-DOG CHO) [27]. To present the Hotelling observer performance, we used Laguerre-Gauss channels (LG CHO) [28].

Methods

Image generation

Simulated breast volume. Breast volumes were generated in computer simulations using two main characteristics of breast anatomy: 1) a power law spectrum and 2) binary attenuation coefficients. The structure of the breast anatomy in mammograms has been characterized using the power law spectrum [16, 18]

$$P(f) = K/f^\beta \quad (1)$$

where f is the radial frequency (mm^{-1}), K is a constant and β is a power law exponent. Based on real clinical mammograms, the estimated value of β is ~ 3 [16]. The attenuation coefficient of breast anatomy is considered as binary to represent the dominant glandular and adipose tissues [29].

For the breast volume, we generated a volume with $512 \times 512 \times 512$ voxels of white Gaussian noise and transformed it into the frequency domain using the discrete Fourier transform (DFT). We computed a pointwise multiplication of the transformed white noise and a three-

dimensional (3D) filtering kernel (i.e. the square root of $1/f^3$), and then calculated the inverse DFT. Note that Eq (1) is a 2D power spectrum, and we simply extended it to the 3D power spectrum [6, 11] by the central slice theorem [30, 31]. As f approaches 0, the value of the filtering kernel becomes infinite; thus, the value at $f = 0$ was set to twice that of the first nonzero radial frequency component [17]. We extracted the central spherical volume with a diameter of 128 voxels from the filtered noise to avoid the wrap-around effect owing to the DFT filtering operation [27]. Afterwards, we sorted the voxel values of the spherical volume in descending order, and assigned the attenuation coefficient of glandular tissue to the top 50% of voxel values and that of adipose tissue to the remaining 50% [32–34], representing a 50% VGF breast volume.

For modelling a mass lesion, a 2mm diameter spherical signal was inserted near the center of the breast volume by replacing the attenuation coefficient of the breast tissue in the signal region with that of the signal. The attenuation coefficients of simulated glandular and adipose tissues, and signal region were 0.233, 0.194, and 0.238cm^{-1} , respectively, which were attenuation coefficients at 50keV energy [35]. The image voxel size was $0.16 \times 0.16 \times 0.16\text{mm}^3$ and the full volume size was $20.5 \times 20.5 \times 20.5\text{mm}^3$.

CBCT data acquisition. We computed the radiological path along the ray that connected the X-ray source and each of the detector pixels to acquire projection data of the simulated breast volume [36] (the CBCT simulation parameters are summarized in Table 1), and a detector quarter offset was used to avoid aliasing. For quantum noise, we generated uniform Poisson noise with 6914 photons per detector pixel, equivalent to the dose used in two-view mammography (i.e. 6.4mGy for a 14cm diameter breast with 50% VGF) [37]. Next, we applied log-normalization to the generated Poisson noise and added it to the noiseless projection data.

In a discrete-to-discrete projection procedure, discretization artifacts in the projection of the breast volume can be introduced when the image voxel size is larger than the detector pixel size [32]. To avoid these, the voxel size of the simulated breast volume was set to $0.16 \times 0.16 \times 0.16\text{mm}^3$, which was smaller than the detector pixel size magnified at the iso-center ($0.2028 \times 0.2028\text{mm}^2$) [11, 32].

Projection data were reconstructed using the FDK reconstruction [15] in which the voxel size was selected as half of the intrinsic voxel size at the iso-center. In the FDK reconstruction, the ramp filter was apodized using three filters: Hanning, Shepp-Logan, or Ram-Lak [23]; their frequency profiles are shown in Fig 1. To examine the effect of linear interpolation on detectability [38], projection data were filtered with and without 8-fold Fourier interpolation, and then voxel-driven back-projection with linear interpolation was performed. Note that the 8-fold Fourier interpolation was implemented by taking the inverse Fourier transform of the zero padded (seven times the length of the filtered projection data length) Fourier transform of the filtered projection data.

Table 1. CBCT simulation parameters.

Parameter	Value
Source to Iso-center Distance	460mm
Source to Detector Distance	880mm
Detector pixel size	$0.388 \times 0.388\text{mm}^2$
Detector array size	150×150 pixels (i.e., $58.2 \times 58.2\text{mm}^2$)
Number of Views	200
Reconstructed voxel size	$0.1014 \times 0.1014 \times 0.1014\text{mm}^3$
Reconstructed matrix size	200×200×200 voxels
Reconstructed volume size	$20.3 \times 20.3 \times 20.3\text{mm}^3$

<https://doi.org/10.1371/journal.pone.0194408.t001>

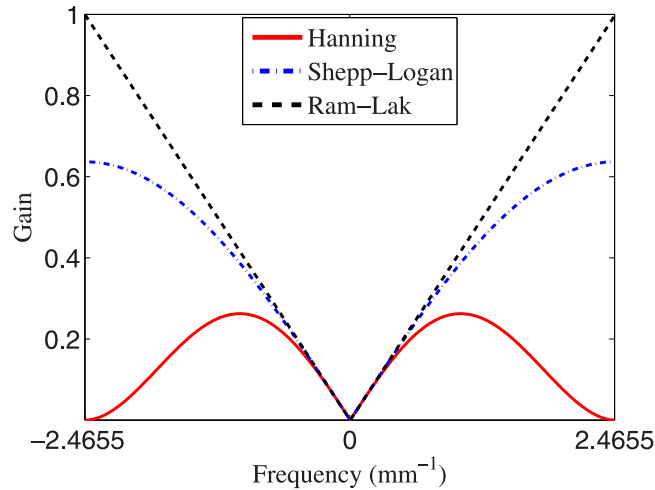


Fig 1. Profiles of the reconstruction filters.

<https://doi.org/10.1371/journal.pone.0194408.g001>

Image preparation. To evaluate detectability, we extracted a central volume of $128 \times 128 \times 128$ voxels from the full volume of $200 \times 200 \times 200$ voxels. Next, we used the central transverse (x-y) plane and the longitudinal (x-z) plane of the volume. With three reconstruction filters and two interpolation methods, evaluation of a total of six detection tasks was conducted for the transverse and longitudinal planes, as summarized in Table 2. The reconstructed breast images for the six tasks are shown in Fig 2.

4AFC detection task

To evaluate detectability by the human and model observers, we conducted 4AFC SKE/BKS detection tasks. The two hypotheses (i.e. H_0 and H_1 for signal-absent and signal-present, respectively) are given by [39]:

$$H_0 : \mathbf{g} = \mathbf{f}_b + \mathbf{f}_n \tag{2}$$

$$H_1 : \mathbf{g} = \mathbf{f}_s + \mathbf{f}_n \tag{3}$$

where \mathbf{f}_b is the breast background, \mathbf{f}_s is the breast background containing the $2mm$ signal, \mathbf{f}_n is the noise in the reconstructed CT image, and \mathbf{g} is the transverse or longitudinal plane reconstructed with various reconstruction filters and interpolation methods.

Table 2. Reconstructed image sets for detection task with different reconstruction filters and interpolation methods.

Task	Reconstruction filter	Interpolation method
1	Hanning	Linear
2	Shepp-Logan	Linear
3	Ram-Lak	Linear
4	Hanning	Fourier
5	Shepp-Logan	Fourier
6	Ram-Lak	Fourier

<https://doi.org/10.1371/journal.pone.0194408.t002>

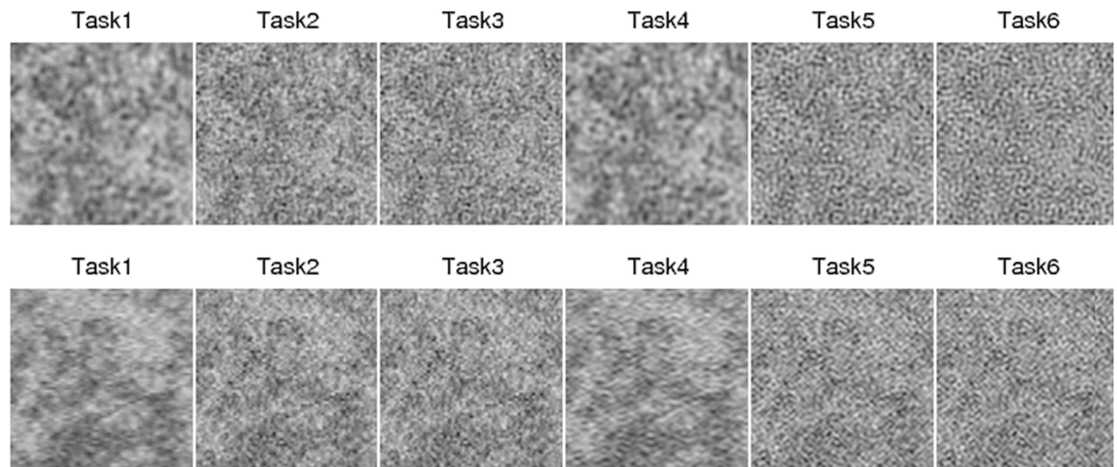


Fig 2. Reconstructed CBCT images. Reconstructed CBCT images for the transverse (upper) and the longitudinal plane (lower). The display window was $[0.10, 0.39] \text{ cm}^{-1}$.

<https://doi.org/10.1371/journal.pone.0194408.g002>

The human observer study

For the human observer study, seven human observers participated and performed six detection tasks for the transverse and longitudinal planes. In each trial, the observers were shown a signal image [20, 40], one signal-present image and three signal-absent images, as shown in Fig 3, and then asked to select the signal-present image. Images were displayed on a 21.3 inch Nio 3MP LED monitor (Barco, Kortrijk, Belgium) with a resolution of 2048×1536 pixels. To reduce distraction, the area around the test images was filled with a dark background. The locations of the signal-present image and signal-absent images were randomly switched in each trial. For all images, the area where the signal may have been present (i.e. the center of the image) was indicated with a red cross. For each task, the observers were trained with 30 trials with instant feedback and tested with 100 trials. We used unique 100 signal-present and 300 signal-absent image datasets for each task and the images were randomly selected from 400 image volumes and displayed in a random order. Note that the training dataset was generated independently from the test dataset for each task, and the training and test datasets were independent for each observer. No time limit was set for choosing an image. Although there was no restriction on viewing distance, it remained approximately $40 \sim 50 \text{ cm}$.

For the j -th trial of each observer, score o_j was recorded, where $j = 1, 2, \dots, 100$. If the answer was correct, o_j was 1, else it was 0. Subsequently, the percent correct, P_c , was computed by

$$P_c = \frac{1}{N_t} \sum_{j=1}^{N_t} o_j \tag{4}$$

where N_t is the number of test trials, and the variance of P_c was estimated by bootstrapping the scores 1,000 times [24].

Model observer study

To compare the detectability by the human observer with that of the model observer, we considered three anthropomorphic model observers: NPWE [25], Gabor CHO [26] and D-DOG CHO [27] because the eye filter in NPWE, and the Gabor and D-DOG channels in CHO

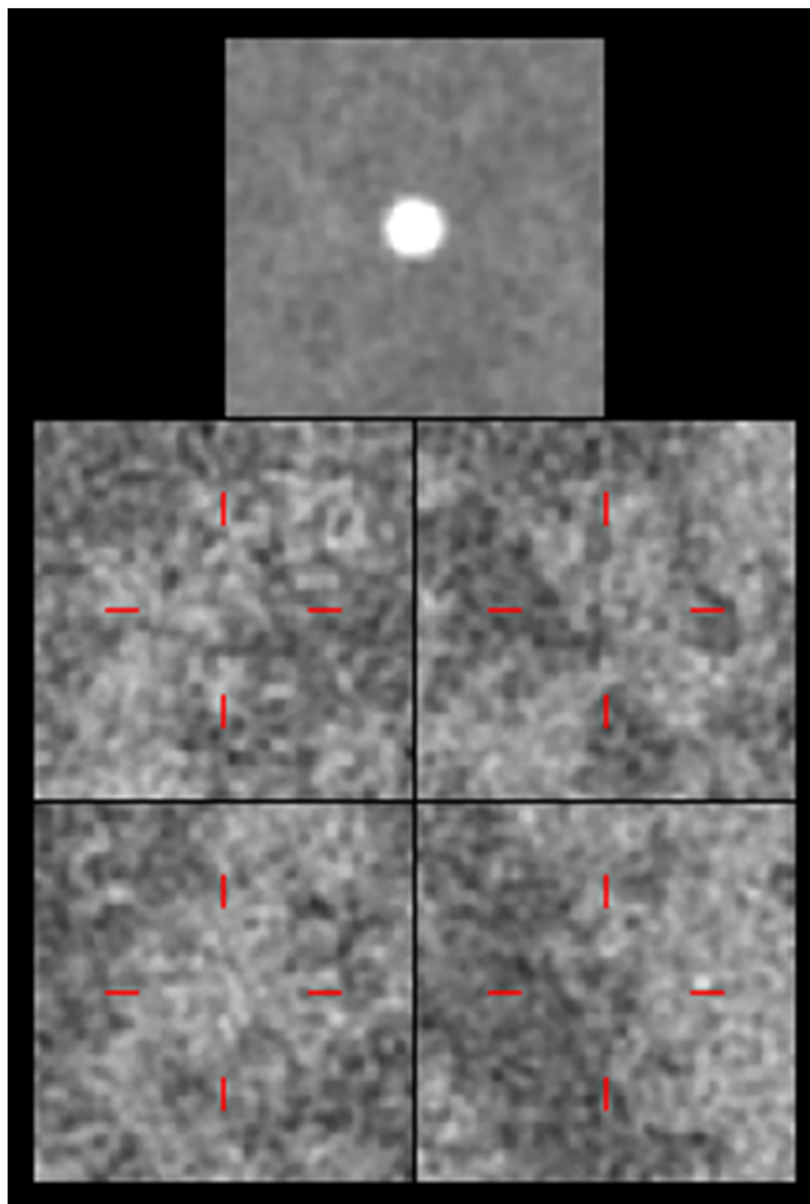


Fig 3. The 4AFC detection task for the human observer study.

<https://doi.org/10.1371/journal.pone.0194408.g003>

mimic the human visual system [19, 40]. Since our signal was rotational symmetric and the background was stationary within the field of view, we used LG CHO to represent the detectability of the Hotelling observer [19, 20].

The template for NPWE. NPWE applies a band pass filter (i.e. eye filter) to images which mimic the frequency selective characteristics of the human visual system [25]. The eye filter function is defined as [25, 27]

$$E(f) = f^{1.3} \exp(-cf^2) \quad (5)$$

where c is an eye filter parameter.

The template of NPWE is given by

$$\mathbf{w}_{\text{NPWE}} = F^{-1}\{\mathbf{E} \cdot F\{\Delta\mathbf{g}\}\} \tag{6}$$

where $F\{\cdot\}$ is a DFT operator, $F^{-1}\{\cdot\}$ is an inverse DFT operator, and $\Delta\mathbf{g}$ is the mean difference between signal-present and signal-absent images. To estimate $\Delta\mathbf{g}$, we used 400 signal-present and signal-absent image pairs.

The template for CHO. CHO applies multiple channels to images and generates a channelized image given by

$$\mathbf{v} = \mathbf{T}\mathbf{g} \tag{7}$$

where \mathbf{v} is the channelized image, \mathbf{T} is the channel matrix and \mathbf{g} is the given image.

For Gabor CHO, the Gabor function is defined as

$$Ch(x, y) = \exp[-4(\ln 2)((x - x_0)^2 + (y - y_0)^2)/w_s^2] \times \cos[2\pi f_c((x - x_0)\cos\theta + (y - y_0)\sin\theta) + \xi] \tag{8}$$

where (x_0, y_0) is the center of the channel and is set to the center of the signal, w_s is the spatial width of the channel, f_c is the center frequency of the channel, θ is the channel orientation, and ξ is a phase factor [26]. Channel matrix \mathbf{T} of the Gabor CHO is composed of discrete samples obtained from Eq (8).

For D-DOG CHO, the DOG function is in the form

$$C_j(\rho) = \exp[-\frac{1}{2}(\frac{\rho}{Q\sigma_j})^2] - \exp[-\frac{1}{2}(\frac{\rho}{\sigma_j})^2] \tag{9}$$

where ρ is the radial frequency (pixel^{-1}), Q is a multiplicative factor, and σ_j is the standard deviation of the j -th channel defined as $\sigma_j = \sigma_0 \alpha^j$. Channel matrix \mathbf{T} of the D-DOG CHO is composed of discrete samples of the inverse DFT of Eq (9).

For LG CHO, the LG function is defined as

$$u_p(\mathbf{r} | a_u) = \frac{\sqrt{2}}{a_u} \exp(\frac{-\pi\mathbf{r}^2}{a_u^2}) L_p(\frac{2\pi\mathbf{r}^2}{a_u^2}) \tag{10}$$

where \mathbf{r} represents a 2D spatial coordinate and a_u is the width of the Gaussian function. The Laguerre polynomials, $L_p(x)$, can be defined as

$$L_p(x) = \sum_{k=0}^p (-1)^k \binom{p}{k} \frac{x^k}{k!} \tag{11}$$

where p is the order of the polynomial. Channel matrix \mathbf{T} of the LG CHO is composed of discrete samples obtained from Eq (10).

The template of the CHO is given by

$$\mathbf{w}_{\text{CHO}} = (\mathbf{K}_v + \mathbf{K}_{\text{int}})^{-1} \Delta\mathbf{v} \tag{12}$$

where \mathbf{K}_v is the channelized covariance matrix, $\Delta\mathbf{v}$ is the channelized mean difference between the signal-present and signal-absent images, and \mathbf{K}_{int} is the covariance matrix of channel internal noise given as

$$\mathbf{K}_{\text{int}} = a \cdot (\mathbf{I} \circ \mathbf{K}_v) \tag{13}$$

where \circ is an elementwise multiplication operator, a is the internal noise level, and \mathbf{I} is an identity matrix. We used a nonuniform channel internal noise model since it showed good agreement between the human and model observers [41]. To estimate \mathbf{K}_v and $\Delta\mathbf{v}$, we used 400

signal-present and signal-absent image pairs. Note that the image sets used to estimate \mathbf{K}_v and $\Delta\mathbf{v}$ were independent of each other.

P_c of the model observers. Model observer performance was evaluated with the 4AFC task. Each trial was composed of one signal-present and three signal-absent images. A test consisting of 100 trials was performed independently seven times. For each trial, the decision variable for NPWE is computed by

$$t = \mathbf{w}_{\text{NPWE}}^t \mathbf{g}_{\text{eye}} + \varepsilon \tag{14}$$

where $\mathbf{g}_{\text{eye}} = F^{-1}\{\mathbf{E} \cdot F\{\mathbf{g}\}\}$ and ε is the internal noise.

The decision variable for the CHO is derived as

$$t = \mathbf{w}_{\text{CHO}}^t (\mathbf{v} + \mathbf{v}_{\text{int}}) \tag{15}$$

where \mathbf{v}_{int} is the channel internal noise. Note that the $j - th$ element of \mathbf{v}_{int} was sampled from random noise whose variance is equivalent to the (j, j) element of $\sqrt{\mathbf{K}_{\text{int}}}$.

For each trial, the image which produced the highest decision variable among the four images was determined as a signal-present image. If the answer from the model observer was correct, then $o_j = 1$, else $o_j = 0$. P_c was computed using Eq (4), and its variance was estimated by bootstrapping the scores 1,000 times. With internal noise ($\varepsilon > 0$ and $a > 0$), P_c was computed by averaging 10 repetitions of the P_c calculation.

Parameter selection for the model observers. The model observer have several parameters but adjusting all of them based on human observer data increases complexity and produces overfitting problems; thus, we adjusted only one parameter for each model observer.

For NPWE, we used two models. In the first model, denoted as NPWE4i, we adjusted the internal noise level ε with $c = 2$ such that the peak value of the eye filter occurred at $4\text{cyc}/\text{degree}$ because this is the value at which the human visual system is most sensitive [27]. In the second model, denoted as NPWEf, we adjusted the peak frequency of the eye filter by changing the c value and set $\varepsilon = 0$.

For Gabor CHO, we adjusted internal noise level a with $f_c = 3/128, 3/64, 3/32$ and $3/16$ ($w_s = 56.48, 28.24, 14.12$ and 7.06), $\theta = 0, 2\pi/5, 4\pi/5, 6\pi/5$ and $8\pi/5$, and $\xi = 0$ and $\pi/2$, as used in [26].

For D-DOG CHO, we adjusted a with $\sigma_0 = 0.005, \alpha = 1.4, Q = 1.67$, and $j = 1 \sim 10$, as used in [27].

Optimal values of ε, c , and a were selected to minimize the mean squared error (MSE) of P_c between the model and human observers. The searching interval for ε was 0.001 within a range of [0, 0.003], and the searching interval for c and a was 0.1 within a range of [0, 3] and [0, 5], respectively.

For the LG CHO, we used 10 channels with $a_u = 6$ and $a = 0$, because P_c became saturated when the number of channels was greater than 10, and was maximized with $a_u = 6$. The description of the model observers used in this study are summarized in Table 3.

Table 3. The model observers used in this study.

Model observer	Description
NPWE4i	NPWE with a peak frequency at $4\text{cyc}/\text{degree}$ and internal noise
NPWEf	NPWE without internal noise and adjusted peak frequency
Gabor CHOi	Gabor CHO with internal noise
D-DOG CHOi	D-DOG CHO with internal noise
LG CHO	LG CHO without internal noise

<https://doi.org/10.1371/journal.pone.0194408.t003>

Estimation of β

To estimate β , we first computed a 2D noise power spectrum (NPS) using 500 reconstructed signal-absent breast images. For each image, we subtracted the mean and multiplied the image by a spatial window W to avoid spectral leakage [18] owing to the finite length of the DFT.

$$W(r) = \begin{cases} 0.5 + 0.5\cos(\pi r/D) & r \leq D \\ 0 & r > D \end{cases} \tag{16}$$

Next, we performed the DFT on the window-applied images, and computed the 2D NPS by averaging the squared magnitudes of the DFT images. We computed a 1D NPS by radial averaging of the 2D NPS, and then applied the natural logarithm to the 1D NPS. To obtain β , we performed linear regression on the logarithm-applied 1D NPS by changing the fitting frequency range. The value of β was chosen to maximize the coefficient of determination (R^2) [18]. Note that R^2 measures the correspondence between a logarithm-applied 1D NPS and the corresponding linear regression model.

Results

Fig 4 shows the averaged P_c of the human observers and Fig 5 shows the measured β for each task. For β estimation, the fitting frequencies ranged from 0.1 to 0.48 cyc/mm with $R^2 > 0.95$. For the various reconstruction filters, the averaged P_c is highest (lowest) when the Hanning (Ram-Lak) filter is used because the Hanning (Ram-Lak) filter yields the lowest (highest) noise power, as shown in Fig 6. Note that the 1D NPS shown in Fig 6 is up to 1 cyc/mm because more than 99% signal power is concentrated below 1 cyc/mm. While the image with the Hanning filter produces the highest P_c , β is also at its highest for this image. Since the Hanning filter reduces high-frequency energy more than the other filters, the slope of the corresponding logarithm-applied 1D NPS increases, which yields a higher β .

The interpolation method producing higher P_c depends on reconstruction filters and image planes, while the Fourier interpolation reduces β for all reconstruction filters and image planes. The frequency response of linear interpolation ($\text{sinc}^2(f)$) reduces high-frequency energy, resulting in a higher β , but using Fourier interpolation minimizes the effect of linear interpolation.

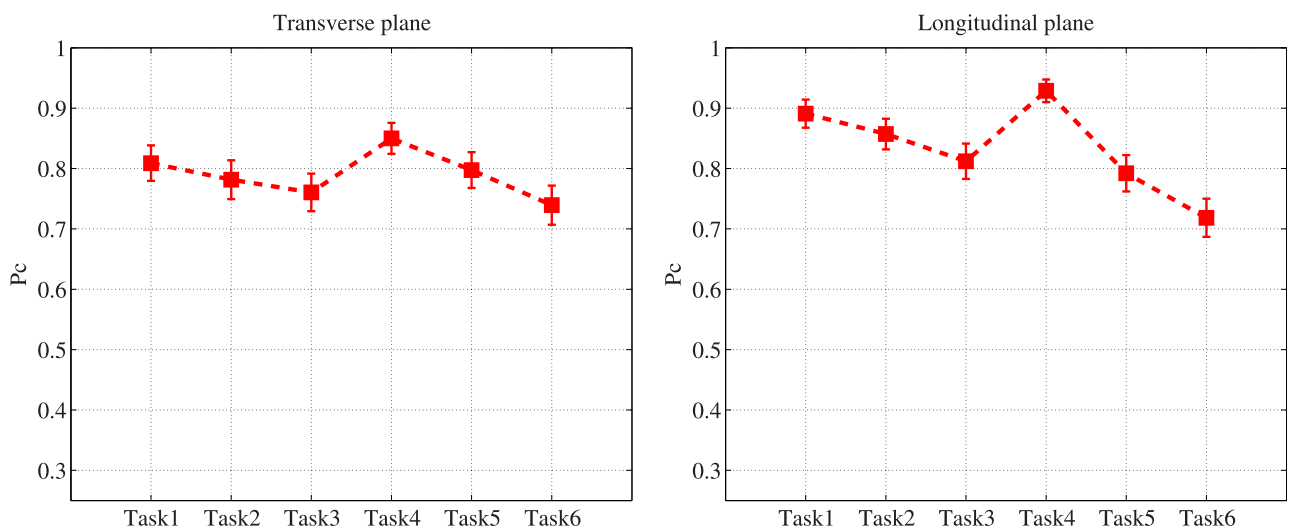


Fig 4. P_c by human observer. Averaged P_c values of human observers with 95% confidence intervals.

<https://doi.org/10.1371/journal.pone.0194408.g004>

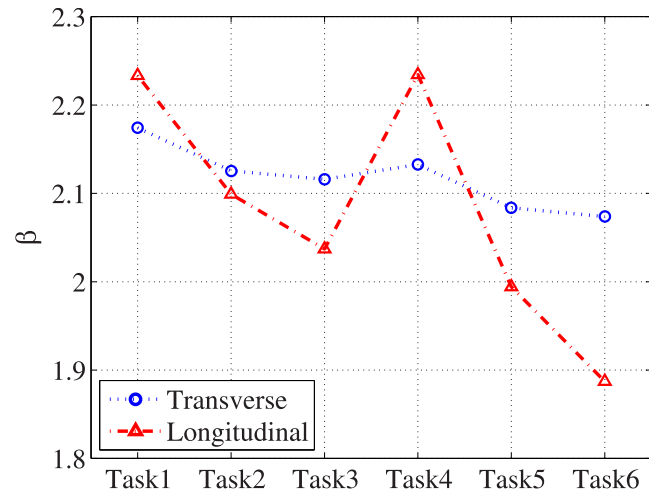


Fig 5. Estimated β . β of the transverse and longitudinal planes.

<https://doi.org/10.1371/journal.pone.0194408.g005>

For the image planes, the longitudinal plane produces higher P_c than the transverse plane from tasks 1 to 4, and the transverse plane produces higher P_c for tasks 5 and 6, because the noise power of the longitudinal plane is higher than that of the transverse plane for these two tasks. The variation of P_c for the various tasks is larger in the longitudinal plane than in the transverse plane because the noise power varies more significantly in the longitudinal plane, as shown in Fig 6. As a result, the β value varies more in the longitudinal plane for the various tasks, as shown in Fig 5.

For both image planes, P_c is highest in task4 and lowest in task6, and Table 4 contains a summary of P_c and β for these tasks. This shows that a task with higher P_c has higher β , which is different from the conclusion of the previous study (i.e., higher β implies lower detectability [16, 17]). This contradiction mainly comes from using apodization filters in the FDK algorithm and the presence of quantum noise, which have not been considered in traditional mammography studies. Quantum noise is a limiting factor in CT image quality because the quantum noise level of each CT projection is much higher than that of the mammography

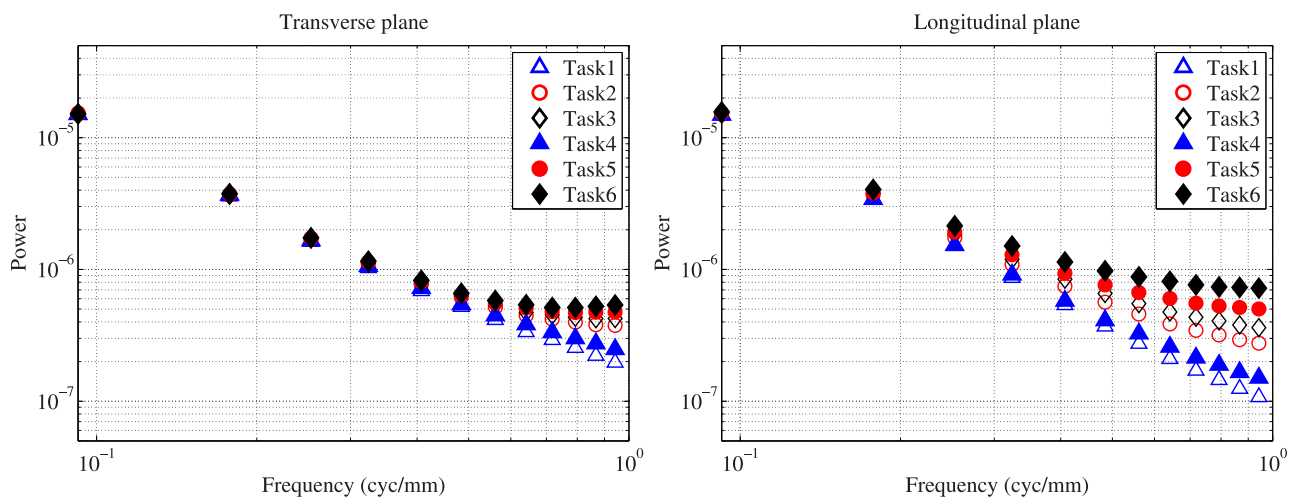


Fig 6. 1D NPS. Log-log plots of 1D NPS for the transverse and longitudinal planes.

<https://doi.org/10.1371/journal.pone.0194408.g006>

Table 4. P_c and β of the task which produces the highest P_c (task4) and lowest P_c (task6).

	Transverse		Longitudinal	
	P_c	β	P_c	β
Task4	0.85	2.13	0.93	2.23
Task6	0.74	2.07	0.72	1.89

<https://doi.org/10.1371/journal.pone.0194408.t004>

with an equivalent dose assumption. For a qualitative comparison, Fig 7 shows example reconstructed images with the signal for the various tasks.

Fig 8 shows the P_c values for the human and model observers. Note that P_c for the anthropomorphic model observers (NPWE4i, NPWEf, Gabor CHOi and D-DOG CHOi) in Fig 8 is calculated using the optimal ϵ , c , and a values summarized in Table 5. LG CHO shows the highest P_c value for all tasks and image planes. Fig 9 shows scatter plots of P_c between the human and anthropomorphic model observers, and corresponding linear regression lines.

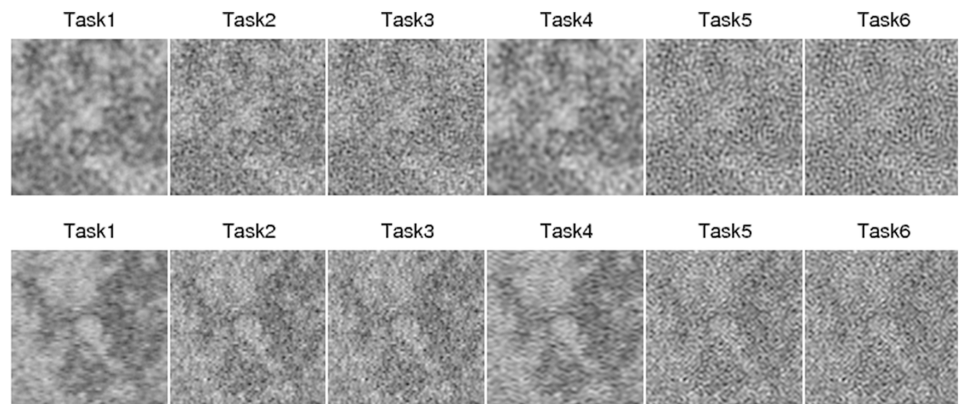


Fig 7. Example CBCT images. Example images with a 2mm diameter signal for the transverse (upper) and longitudinal planes (lower).

<https://doi.org/10.1371/journal.pone.0194408.g007>

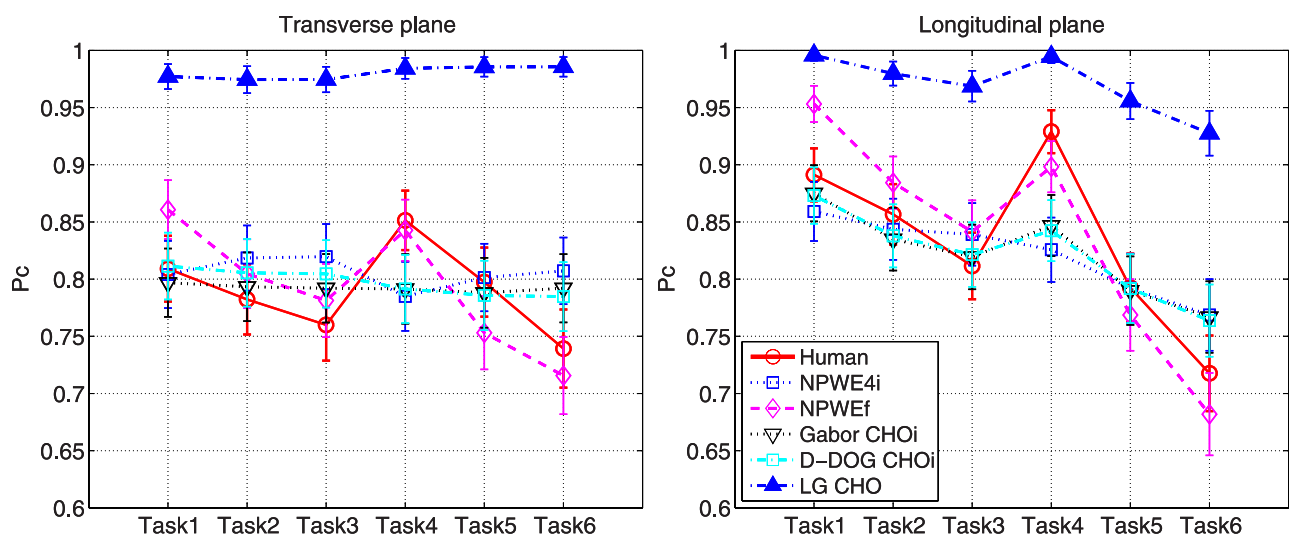


Fig 8. P_c by human and model observers. The P_c value of the human and model observers with 95% confidence intervals.

<https://doi.org/10.1371/journal.pone.0194408.g008>

Table 5. Optimal parameters (i.e., ϵ , c and a) for the anthropomorphic model observers.

	NPWE4i	NPWEf	Gabor CHOi	D-DOG CHOi
Transverse	$\epsilon = 0.001$	$c = 0.7$	$a = 2.9$	$a = 3.8$
Longitudinal	$\epsilon = 0.001$	$c = 0.6$	$a = 2.1$	$a = 2.6$

<https://doi.org/10.1371/journal.pone.0194408.t005>

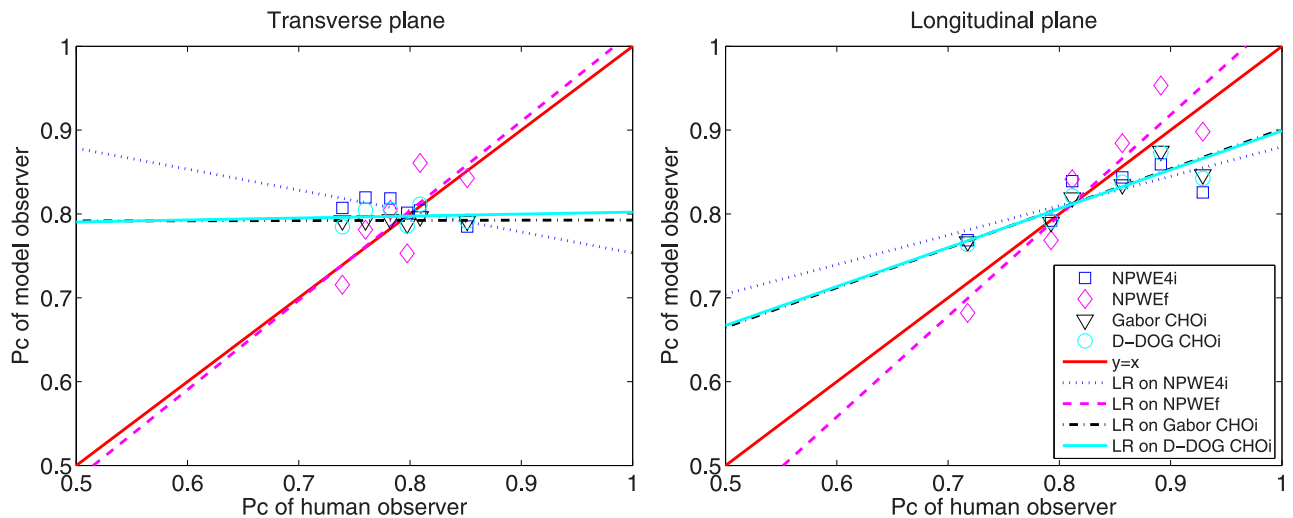


Fig 9. Correlation between human and model observers. Scatter plots of P_c between human and model observers, and corresponding linear regression (LR) results.

<https://doi.org/10.1371/journal.pone.0194408.g009>

The greater the similarity between the linear regression line and $y = x$ line, the higher the correlation between the human and model observers. NPWEf shows the best correlation with a human observer in both the transverse and longitudinal planes; the eye filter profiles for NPWEf with optimal c are shown in Fig 10. The eye filter has its peak value at $1\text{cyc}/\text{mm}$, which translates to $7\text{cyc}/\text{degree}$ for a 400mm viewing distance belonging to the sensitive frequency range of the human visual system [42].

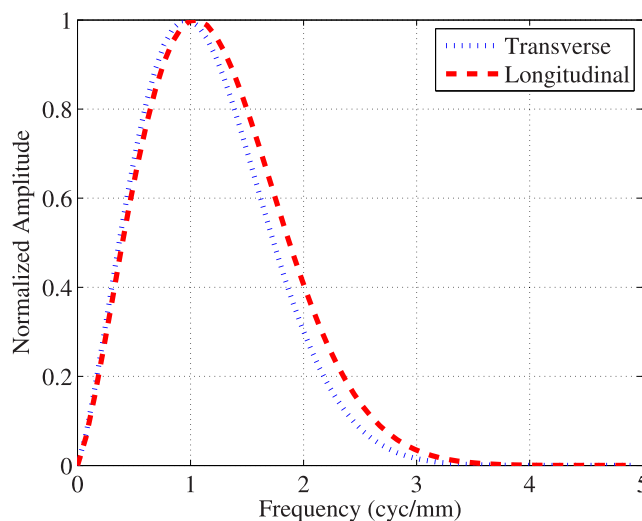


Fig 10. Eye filter. Eye filter profiles with the optimal value of c .

<https://doi.org/10.1371/journal.pone.0194408.g010>

Conclusion and discussion

We evaluated detectability in breast CBCT images using human and model observers, and β with various reconstruction filters, interpolation methods, and image planes. In this study, NPWEf with a peak frequency of 7 cyc/degree showed the highest correlation with human observer performance for the various reconstruction filters, interpolation methods, and image planes. For both image planes, detectability was highest for the Hanning filter and lowest for the Ram-Lak filter both with Fourier interpolation. However, higher detectability was not related to a smaller β value in the presence of quantum noise.

Although the main focus of this work is to evaluate the detectability in breast CBCT images, we evaluated detectability for the same tasks with a uniform background. Note that the 2 mm diameter spherical signal was also used for the uniform background case, and the intensity of the signal was adjusted to produce P_c values ranging from 0.5 to 1 for all tasks. As shown in Fig 11, the NPWEf still showed a high correlation with the human observers for both image planes. Note that the optimal eye filter for a uniform background had a peak value at around 1.5 cyc/mm , which is different from the breast anatomy background (1 cyc/mm as shown in Fig 10). This result indicates that the peak frequency of an eye filter depends on image statistics. In contrast to the results with breast CBCT images, both Gabor CHOi and D-DOG CHOi showed good correlation with human observers for the longitudinal plane.

Of the model observers, NPWEf showed the highest correlation with human observers, but its performance was sensitive to the peak frequency of the eye filter. Thus, optimization for one signal size might not have been optimal for other signal sizes. In these cases, D-DOG CHOi would be more useful to predict human observer performance, as shown in [20].

We used a central single slice of the CBCT images, but in order to take signal information along the direction in which the image slices are stacked into account, evaluating multiple slices of CBCT images is required, which is a subject for future research.

Since our study is a preliminary investigation of detectability in breast CBCT images, we modelled the breast volume using simple characteristics of breast anatomy without considering detailed morphological features such as tissue orientation [43], other tissue types [44], various VGF values [45], and other shapes of lesion [46]. In our CBCT simulation, we did not

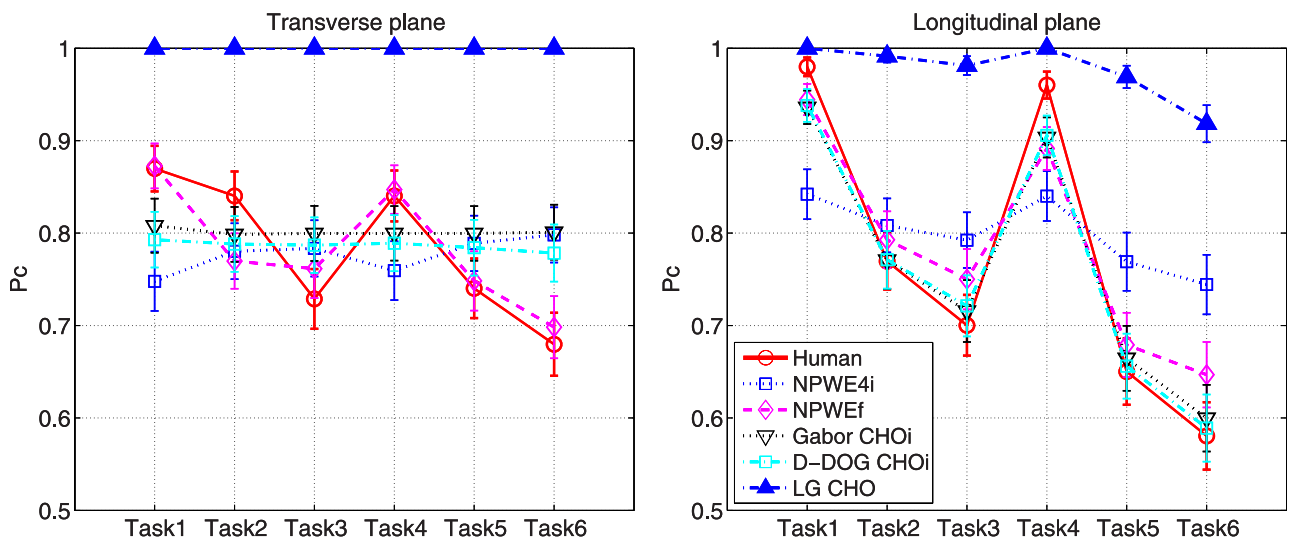


Fig 11. P_c by human and model observers for uniform background. P_c of human and model observers with 95% confidence intervals for uniform backgrounds.

<https://doi.org/10.1371/journal.pone.0194408.g011>

consider the effect of beam hardening and scatter because their effects on image quality are dependent on X-ray energy, object size, and data acquisition geometry. In the presence of beam hardening and scatter, detectability will be reduced due to the decreased contrast, especially for low contrast small objects. However, the effects of the apodization filter and interpolation method on detection task would remain the same. We also assumed an ideal detector response. Correlation between detector pixels and nonuniform detector pixel response degrade the overall image quality, and thus provide lower detectability in our tasks. In future studies, we aim to extend this work to clinical breast CBCT images, for which we will consider these physical factors.

Supporting information

S1 File. P_c by each of the seven human observers. P_c values by each of the seven human observers for six detection tasks. (XLSX)

Author Contributions

Conceptualization: Minah Han.

Formal analysis: Minah Han, Jongduk Baek.

Funding acquisition: Jongduk Baek.

Investigation: Minah Han, Jongduk Baek.

Methodology: Minah Han, Byeongjoon Kim, Jongduk Baek.

Validation: Minah Han, Jongduk Baek.

Visualization: Minah Han, Byeongjoon Kim.

Writing – original draft: Minah Han.

Writing – review & editing: Minah Han, Jongduk Baek.

References

1. Seigel R, Naishadham D, Jemal A. Cancer statistics, 2012. *Ca Cancer J Clin.* 2012; 62(1):10–29. <https://doi.org/10.3322/caac.20138>
2. Mushlin AI, Kouides RW, Shapiro DE. Estimating the accuracy of screening mammography: a meta-analysis. *Am J Prev Med.* 1998; 14(2):143–153. [https://doi.org/10.1016/S0749-3797\(97\)00019-6](https://doi.org/10.1016/S0749-3797(97)00019-6) PMID: 9631167
3. Evans KK, Birdwell RL, Wolfe JM. If you don't find it often, you often don't find it: why some cancers are missed in breast cancer screening. *PLoS One.* 2013; 8(5):e64366. <https://doi.org/10.1371/journal.pone.0064366> PMID: 23737980
4. Chen B, Ning R. Cone-beam volume CT breast imaging: Feasibility study. *Med Phys.* 2002; 29(5):755–770. <https://doi.org/10.1118/1.1461843> PMID: 12033572
5. Packard NJ, Abbey CK, Yang K, Boone JM. Effect of slice thickness on detectability in breast CT using a prewhitened matched filter and simulated mass lesions. *Med Phys.* 2012; 39(4):1818–1830. <https://doi.org/10.1118/1.3692176> PMID: 22482604
6. 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. *Med Phys.* 2006; 33(4):1041–1052. <https://doi.org/10.1118/1.2174127> PMID: 16696481
7. Lai CJ, Shaw CC, Chen L, Altunbas MC, Liu X, Han T, et al. Visibility of microcalcification in cone beam breast CT: Effects of x-ray tube voltage and radiation dose. *Med Phys.* 2007; 34(7):2995–3004. <https://doi.org/10.1118/1.2745921> PMID: 17822008

8. Glick SJ, Vedantham S, Karellas A. Investigation of optimal kVp settings for CT mammography using a flat-panel imager. In: Proc. SPIE. International Society for Optics and Photonics; 2002. p. 392–402.
9. Gong X, Vedula AA, Glick SJ. Microcalcification detection using cone-beam CT mammography with a flat-panel imager. *Phys Med Biol*. 2004; 49(11):2183–2195. <https://doi.org/10.1088/0031-9155/49/11/005> PMID: 15248571
10. Chen L, Abbey CK, Nosratieh A, Lindfors KK, Boone JM. Anatomical complexity in breast parenchyma and its implications for optimal breast imaging strategies. *Med Phys*. 2012; 39(3):1435–1441. <https://doi.org/10.1118/1.3685462> PMID: 22380376
11. Han M, S, Baek J. Effect of anatomical noise on the detectability of cone beam CT images with different slice direction, slice thickness, and volume glandular fraction. *Opt Express*. 2016; 24(17):18843–18859. <https://doi.org/10.1364/OE.24.018843> PMID: 27557168
12. Bian J, Yang K, Boone JM, Han X, Sidky EY, Pan X. Investigation of iterative image reconstruction in low-dose breast CT. *Phys Med Biol*. 2014; 59(11):2659. <https://doi.org/10.1088/0031-9155/59/11/2659> PMID: 24786683
13. Makeev A, Glick SJ. Investigation of statistical iterative reconstruction for dedicated breast CT. *Med Phys*. 2013; 40(8):081904. <https://doi.org/10.1118/1.4811328> PMID: 23927318
14. Beister M, Kolditz D, Kalender WA. Iterative reconstruction methods in X-ray CT. *Physica medica*. 2012; 28(2):94–108. <https://doi.org/10.1016/j.ejmp.2012.01.003> PMID: 22316498
15. Feldkamp LA, Davis LC, Kress JW. Practical cone-beam algorithm. *J Opt Soc Am*. 1984; 1(6):612–619. <https://doi.org/10.1364/JOSAA.1.000612>
16. Burgess AE, Jacobson FL, Judy PF. Human observer detection experiments with mammograms and power-law noise. *Med Phys*. 2001; 28(4):419–437. <https://doi.org/10.1118/1.1355308> PMID: 11339738
17. Burgess AE, Judy PF. Signal detection in power-law noise: effect of spectrum exponents. *J Opt Soc Am A*. 2007; 24(12):B52–B60. <https://doi.org/10.1364/JOSAA.24.000B52>
18. Chen L, Abbey CK, Boone JM. Association between power law coefficients of the anatomical noise power spectrum and lesion detectability in breast imaging modalities. *Phys Med Biol*. 2013; 58(6):1663–1681. <https://doi.org/10.1088/0031-9155/58/6/1663> PMID: 23422272
19. He X, Park S. Model observers in medical imaging research. *Theranostics*. 2013; 3(10):774–786. <https://doi.org/10.7150/thno.5138> PMID: 24312150
20. Han M, Lee C, Park S, Baek J. Investigation on slice direction dependent detectability of volumetric cone beam CT images. *Opt Express*. 2016; 24(4):3749–3764. <https://doi.org/10.1364/OE.24.003749> PMID: 26907031
21. Huang J, Zhang Y, Ma J, Zeng D, Bian Z, Niu S, et al. Iterative image reconstruction for sparse-view CT using normal-dose image induced total variation prior. *PloS one*. 2013; 8(11):e79709. <https://doi.org/10.1371/journal.pone.0079709> PMID: 24260288
22. Mananga ES, El Fakhri G, Schaefferkoetter J, Bonab AA, Ouyang J. Myocardial defect detection using PET-CT: phantom studies. *PloS one*. 2014; 9(2):e88200. <https://doi.org/10.1371/journal.pone.0088200> PMID: 24505429
23. Jang SY, Kim HK, Youn H, Cho S, Cunningham IA. Fourier analysis of noise characteristics in cone-beam microtomography laboratory scanners. *IEEE Trans Biomed Eng*. 2017; 64(1):173–183. <https://doi.org/10.1109/TBME.2016.2552496> PMID: 27093307
24. Timberg P, Båth M, Andersson I, Mattsson S, Tingberg A, Ruschin M. Visibility of microcalcification clusters and masses in breast tomosynthesis image volumes and digital mammography: a 4AFC human observer study. *Med Phys*. 2012; 39(5):2431–2437. <https://doi.org/10.1118/1.3694105> PMID: 22559613
25. Burgess A. Statistically defined backgrounds: performance of a modified nonprewhitening observer model. *J Opt Soc Am A*. 1994; 11(4):1237–1242. <https://doi.org/10.1364/JOSAA.11.001237>
26. Wunderlich A, Noo F. Image covariance and lesion detectability in direct fan-beam x-ray computed tomography. *Phys Med Biol*. 2008; 53(10):2471. <https://doi.org/10.1088/0031-9155/53/10/002> PMID: 18424878
27. Abbey CK, Barrett HH. Human-and model-observer performance in ramp-spectrum noise: effects of regularization and object variability. *J Opt Soc Am A*. 2001; 18(3):473–488. <https://doi.org/10.1364/JOSAA.18.000473>
28. Gallas BD, Barrett HH. Validating the use of channels to estimate the ideal linear observer. *J Opt Soc Am A*. 2003; 20(9):1725–1738. <https://doi.org/10.1364/JOSAA.20.001725>
29. Metheany KG, Abbey CK, Packard N, Boone JM. Characterizing anatomical variability in breast CT images. *Med Phys*. 2008; 35(10):4685–4694. <https://doi.org/10.1118/1.2977772> PMID: 18975714
30. Kak AC, Slaney M. Principles of Computerized Tomographic Imaging; 1988.

31. Cao ZJ, Tsui BMW. A fully three-dimensional reconstruction algorithm with the nonstationary filter for improved single-orbit cone beam SPECT. *IEEE T Nucl Sci.* 1993; 40(3):280–287. <https://doi.org/10.1109/23.221052>
32. Reiser I, Nishikawa RM. Task-based assessment of breast tomosynthesis: Effect of acquisition parameters and quantum noise. *Med Phys.* 2010; 37(4):1591–1600. <https://doi.org/10.1118/1.3357288> PMID: 20443480
33. Richard S, Samei E. Quantitative imaging in breast tomosynthesis and CT: Comparison of detection and estimation task performance. *Med Phys.* 2010; 37(6):2627–2637. <https://doi.org/10.1118/1.3429025> PMID: 20632574
34. Abbey CK, Boone JM. An ideal observer for a model of x-ray imaging in breast parenchymal tissue. In: *International Workshop on Digital Mammography*. Springer; 2008. p. 393–400.
35. Johns PC, Yaffe MJ. X-ray characterisation of normal and neoplastic breast tissues. *Phys Med Biol.* 1987; 32(6):675–695. <https://doi.org/10.1088/0031-9155/32/6/002> PMID: 3039542
36. Siddon RL. Fast calculation of the exact radiological path for a three-dimensional CT array. *Med Phys.* 1985; 12(2):252–255. <https://doi.org/10.1118/1.595715> PMID: 4000088
37. Boone JM, Kwan AL, Seibert JA, Shah N, Lindfors KK, Nelson TR. Technique factors and their relationship to radiation dose in pendant geometry breast CT. *Med Phys.* 2005; 32(12):3767–3776. <https://doi.org/10.1118/1.2128126> PMID: 16475776
38. Baek J, Pelc NJ. The noise power spectrum in CT with direct fan beam reconstruction. *Med Phys.* 2010; 37(5):2074–2081. <https://doi.org/10.1118/1.3378673> PMID: 20527540
39. Barrett HH, Abbey CK. Bayesian detection of random signals on random backgrounds. In: *Biennial International Conference on Information Processing in Medical Imaging*. Springer; 1997. p. 155–166.
40. Ba A, Eckstein MP, Racine D, Ott JG, Verdun F, Kobbe-Schmidt S, et al. Anthropomorphic model observer performance in three-dimensional detection task for low-contrast computed tomography. *Journal of Medical Imaging.* 2016; 3(1):011009–011009. <https://doi.org/10.1117/1.JMI.3.1.011009> PMID: 26719849
41. Zhang Y, Pham BT, Eckstein MP. Evaluation of internal noise methods for Hotelling observer models. *Med Phys.* 2007; 34(8):3312–3322. <https://doi.org/10.1118/1.2756603> PMID: 17879795
42. Mannos J, Sakrison D. The effects of a visual fidelity criterion of the encoding of images. *IEEE T Inform Theory.* 1974; 20(4):525–536. <https://doi.org/10.1109/TIT.1974.1055250>
43. Reiser I, Lee S, Nishikawa RM. On the orientation of mammographic structure. *Med Phys.* 2011; 38(10):5303–5306. <https://doi.org/10.1118/1.3633905> PMID: 21992347
44. Wu X, Chen G, Lu J, Zhu W, Qiu J, Chen J, et al. Label-free detection of breast masses using multiphoton microscopy. *PloS one.* 2013; 8(6):e65933. <https://doi.org/10.1371/journal.pone.0065933> PMID: 23755295
45. Gubern-Mérida A, Kallenberg M, Platel B, Mann RM, Martí R, Karssemeijer N. Volumetric breast density estimation from full-field digital mammograms: a validation study. *PLoS One.* 2014; 9(1):e85952. <https://doi.org/10.1371/journal.pone.0085952> PMID: 24465808
46. An YY, Kim SH, Kang BJ. Differentiation of malignant and benign breast lesions: Added value of the qualitative analysis of breast lesions on diffusion-weighted imaging (DWI) using readout-segmented echo-planar imaging at 3.0 T. *PloS one.* 2017; 12(3):e0174681. <https://doi.org/10.1371/journal.pone.0174681> PMID: 28358833