Computational assessment of stereoscopic viewing a sequence of stereo pairs of breast tomosynthesis projection images

Gezheng Wen[1], *Student Member, IEEE* and Mia K. Markey[2], *Senior Member, IEEE*

***Abstract*— Digital breast tomosynthesis (DBT) is a 3D imaging technology in which an x-ray fan beam rotates around the breast, producing a series of projection images. The imaging geometry of DBT lends itself naturally to stereo viewing because a stereo pair can be easily formed by two projection images with a reasonable separation angle. Stereo viewing may reveal the 3D structures of breasts thus has the potential to increase the sensitivity and specificity of breast imaging. In this study, we conduct a simulation study that mimics the detection of breast lesions on stereoscopic viewing of DBT projections. The presentation approach we investigate here is one in which the reader is presented with a sequence of stereo pairs from a rotating point of view. We render voxel datasets that contain random 3D power-law noise to model normal breast tissues with different breast densities. A 3D Gaussian signal is inserted to some of the datasets to model the presence of a breast lesion. Sequences of stereo pairs of projection images are generated for each voxel dataset by varying the projection angles of the two views. The diagnostic performance, in terms of the accuracy of binary decisions on the presence of the simulated lesions, is evaluated with a numerical model observer.**

I. INTRODUCTION

Digital breast tomosynthesis (DBT) is a 3D imaging technology in which an x-ray fan beam rotates over a limited angular span around the breast, producing a series of projection images [1]. DBT yields volumetric, 3D data reducing tissue overlap encountered in conventional 2D mammography. In this study, we are interested in evaluating the efficacy of stereo viewing of DBT projection images because the imaging geometry of DBT lends itself naturally to stereo viewing. A stereo pair can be formed by selecting two projection images with a reasonable separation angle. With the aid of a stereo display, stereo viewing creates an impression of depth, which provides more details of anatomical structures and has the potential to increase the sensitivity and specificity of disease detection. A pilot human observer study [2] comparing stereoscopic and monoscopic detection of masses showed that stereo viewing could yield better detection performance. However, more empirical evidence is needed. As human observer studies are resource-demanding to conduct, it would be valuable to use a reliable numerical model observer as a surrogate. The goal is to provide an accurate prediction of human observer performance in the same clinical context. In addition to our work [3], we are aware of a few of investigations of model observers specifically designed for stereo viewing of 3D imaging data, particularly Zafar et al. [4]). A key feature of stereo matching based model observer is that it incorporates the characteristics of stereopsis. Given a stereo pair of images (i.e., left and right images separated by a small displacement), the model observer first finds the matching points in the two views, and then fuses them together to create a cyclopean view. Assuming that the cyclopean view extracts most of the 3D information presented in the stereo pair, a channelized Hotelling observer (CHO) is utilized to make decisions. We have shown in [3] that the model is able to generate decision statistics that are consistent with human observer performance of lesion detection as reported in prior studies .

In this paper, we present a simulation study that investigates the viewing approach in which a reader is presented with a sequence of stereo pairs of DBT projection images from a rotating point of view. Each of the stereo pairs in the sequence is created using the optimal separation angle as determined in [2], [3], [5] . For illustration, assume that the optimal separation angle is θ and there are six number-labeled raw projection images acquired from six different projection angles, in which image 1 and image 4 are separated by θ, image 2 and image 5 are separated by θ, and image 3 and image 6 are separated by θ (as shown in Fig. 1). Then a sequence of stereo pairs is created by presenting a stereo pair consisting of image 1 and image 4, followed by a stereo pair consisting of image 2 and image 5 and so on. The perceived effect of navigating through the sequence of stereo images is rotation of the breasts in the stereo display. It is expected that such presentations of DBT data may help radiologists to better perceive and separate structures within the breast in depth. Computational assessment is conducted with the stereo model observer to justify the assumption.

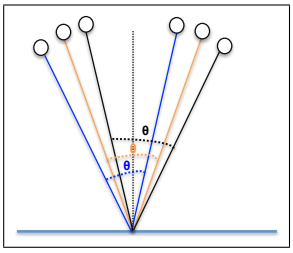


Fig. 1. Schematic of generating a sequence of stereo pairs of breast tomosynthesis projection images.

[1]G. Wen is with Dept. of Electrical and Computer, The University of Texas at Austin, TX 78712 USA, and Dept. of Diagnostic Imaging, The University of Texas MD Anderson Cancer Center, Houston, TX 77030 USA wen.gezheng@utexas.edu

[2]M. K. Markey is with Dept. of Biomedical Engineering, The University of Texas at Austin, TX 78712 USA, and Dept. of Imaging Physics, The University of Texas MD Anderson Cancer Center, Houston, TX 77030 USA mia.markey@utexas.edu

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

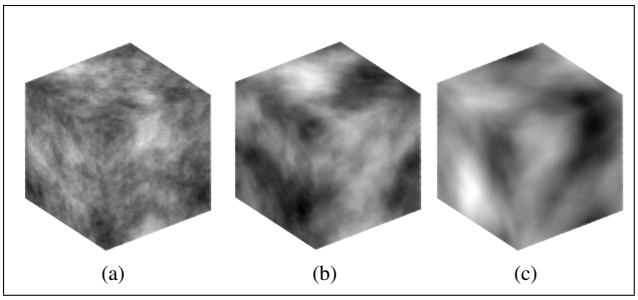


Fig. 2. Examples of simulated voxel datasets: (a) $\beta = 2$, (b) $\beta = 2.5$, (c) $\beta = 3$. The collections of light and dark areas correspond to pockets of fatty and glandular tissue. The larger the dark area is, the higher β is, and hence the higher the breast density is.

II. MATERIALS AND METHODS

A. Image Dataset

For each trial of the study, 1500 simulated voxel datasets are rendered in MATLAB (MathWork, Natick, MA) to mimic the characteristics of x-ray projection images of the breasts, in which 500 are used as the training dataset and the other 1000 are used as the testing dataset. The resolution of the images is $256 \times 256 \times 256$. We adapt the settings commonly used in prior observers studies (e.g., [6]–[8]) to simulate the anatomical structures of breast. Random power-law noise (i.e., $P(f) = \frac{1}{f^\beta}$, where β fits a power function to anatomical noise power spectrum [6]) is used as the image background of the simulated breast volumes. The value of β ranges from 2 to 3 in the datasets to mimic the variability of breast density, where breast density is quantified as the percentage of fibro glandular tissue in the breast volume. Previous research (e.g., [9], [10]) shows a general trend of increasing β with increasing breast density. It is well recognized that breast density affects the detectability of breast lesions (e.g., [3], [9], [10]). Here, we would like to check whether increasing β in the context of stereo viewing results in a reduction of detectability. A 3D grid of noise with the expected power spectrum is created by: 1). assign random samples from the standard Gaussian distribution (i.e., the mean $\mu = 0$ and the deviation $\sigma^2 = 1$) to each voxel; 2). filter the random noisy voxel with a digital finite impulse filter of the form [11]:

$$H(i,j,k) = \begin{cases} 1 & \text{if } i = j = k \\ \sqrt{i^{2\beta} + j^{2\beta} + k^{2\beta}} & \text{otherwise} \end{cases}, \qquad (1)$$

where $(i,j,k) \in [0,(I-1,J-1,K-1)]$. The filtering is conducted in the frequency domain by the use of the discrete Fourier transform. Examples of simulated voxel datasets with distinct β are shown in Fig. 2. For each of the signal-present cases, a 3D Gaussian signal is added at the center of the volume as a simulated lesion [3]. The contrast between the signal and the background is set to around $10 \pm 0.4\%$ [12] to simulate the appearance of a lesion embedded in normal breast tissues.

B. Stereo Model Observer

We follow the process of stereo matching as in [3] to generate a sequence of stereo pairs and the corresponding sequence of cyclopean views. A multislice channelized Hotelling observer (msCHO) is used on the cyclopean views to generate decision statistics.

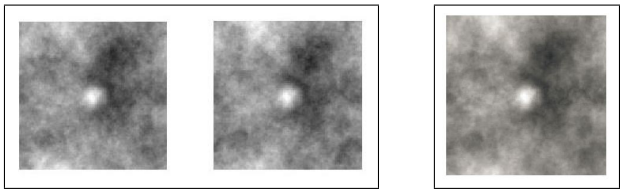


Fig. 3. (a) An example of a simulated stereo pair with a separation angle of 8 degrees; The left view is acquired at $+4^\circ$ apart from the central projection direction while the right view is acquired at -4° apart; (b) An example of a cyclopean view synthesized from the stereo pair shown in Fig. 3a.

1) Stereo Matching:

a) A Sequence of Stereo Pairs: We generate a sequence of stereo pairs of projection images from the voxel datasets (as shown in Fig. 1). The two views of one stereo pair, denoted as $\mathbf{g}_{left}$ and $\mathbf{g}_{right}$, are separated by 8°, which we previously demonstrated [3] to be the optimal separation angle for depth perception of breast. One interesting question we explore here is whether the angular span of projections impacts radiologists' diagnostic performance. As one of the main design considerations in DBT systems [13], it is reasonable to expect that increasing the number of stereo pairs increases diagnostic accuracy. On the other hand, there may be some threshold beyond which adding more pairs no longer improves accuracy. Thus, we generate different sets of stereo pairs of the same volume by changing the number of stereo pairs in the sequence (i.e., angular span $\Theta = 10^\circ, 15^\circ, 20^\circ$, the interval between two pairs $= 1^\circ$). An example of a simulated stereo pair is shown in Fig. 3a.

b) Cyclopean Views: This mental image, created in the brain by fusing two images received from the two eyes [14], appears as seen from a virtual eye placed midway between the two eyes. Stereo matching is used to incorporate this property into the model observer. First, a disparity map $\mathbf{d}$, defined as the difference in the location of an object in $\mathbf{g}_{left}$ and $\mathbf{g}_{right}$ [15], [16] is computed by the belief propagation based stereo matching [15], [17]. Second, $\mathbf{g}_{right}$ is shifted based on $\mathbf{d}$ to match with the corresponding pixels in $\mathbf{g}_{left}$: $\mathbf{g'}_{right}(x,y) = \mathbf{g}_{right}(x + \mathbf{d}(x,y)_x, y + \mathbf{d}(x,y)_y)$, where (x,y) is the pixel location and $\mathbf{d}(x,y)$ is the direction of shift in x and y coordinates. As a result, a 2D cyclopean view $\mathbf{g}_{cyclo}$ is constructed as the average of the left image $\mathbf{g}_{left}$ and the disparity-compensated right image $\mathbf{g'}_{right}$: $\mathbf{g}_{cyclo} = \frac{1}{2} \times (\mathbf{g'}_{right} + \mathbf{g}_{left})$. An example of a cyclopean view is shown in Fig. 3b.

2) Detection Task: Given a sequence of simulated stereo pairs, the detection task of the model observer is to make a binary decision regarding the presence or absence of a simulated lesion. Denote $\mathbf{s}$ as the signal to be detected, $\mathbf{b}$ as the noise-less image background, and $\mathbf{n}$ as the noise in the image. Then, the image data under the signal-present and

signal-absent hypotheses are given by:

$$\begin{aligned} \mathrm{H}_0 : \mathbf{g} &= \mathbf{b} + \mathbf{n}, \\ \mathrm{H}_1 : \mathbf{g} &= \mathbf{b} + \mathbf{n} + \mathbf{s}, \end{aligned}$$

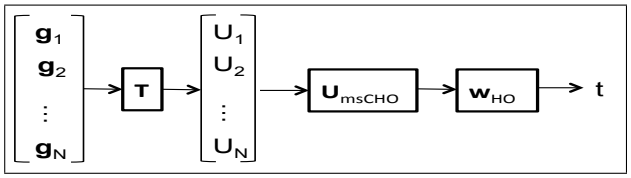
Through the process of stereo matching discussed in Section II-B.1, most of the 3D information conveyed in a stereo pair $(\mathbf{g}_{left}, \mathbf{g}_{right})$ is captured by the corresponding 2D cyclopean view $\mathbf{g}_{cyclo}$. Thus, the detection task with the sequence of stereo pairs can be simplified into finding the lesion (if any) in the sequence of cyclopean views. It is reasonable for the stereo model observer to adapt the design of msCHO [18] for computing the decision statistics.

3) Multislice channelized Hotelling Observer (msCHO): The Hotelling observer uses the means and covariance of image data to differentiate between classes of images [19]. The CHO is widely used to reduce the high dimensionality of image data. A channelized image $\mathbf{U}$ is represented as: $\mathbf{U} = \mathbf{T}\mathbf{g}$, where $\mathbf{g}$ is the image, $\mathbf{T}$ is an $N_c \times N_p$ matrix that represents a set of channels, and U is an $N_c \times 1$ channelized image. N_c is the number of channels and N_p is the dimension of $\mathbf{g}$. Generally,$N_c \ll N_p$. The test statistic t is in the form of: $t(\mathbf{U}) = \mathbf{w}_{\mathbf{U}}^T \mathbf{U}$, where $\mathbf{w}_{\mathbf{U}}$ is the template in the form of:

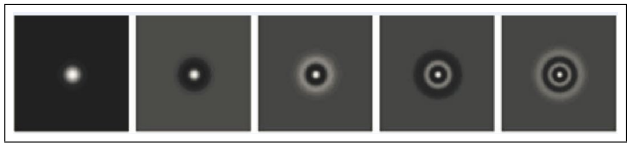
$$\begin{aligned} \mathbf{w}_{\mathbf{U}} &= \mathbf{K}_{\mathbf{U}}^{-1} \overline{\mathbf{U}_{\mathbf{s}}}, \\ \overline{\mathbf{U}_{\mathbf{s}}} &= \overline{\mathbf{U}_1} - \overline{\mathbf{U}_0}, \quad \overline{\mathbf{U}_j} = \mathbb{E}[\mathbf{U}|H_j], j = 0, 1 \\ \mathbf{K}_{\mathbf{U}} &= \frac{1}{2}[\mathbf{K}_{\mathbf{U},0} + \mathbf{K}_{\mathbf{U},1}], \\ \mathbf{K}_{\mathbf{U},j} &= \mathbb{E}[(\mathbf{U} - \overline{\mathbf{U}_j})(\mathbf{U} - \overline{\mathbf{U}_j})^T|H_j], j = 0, 1 \end{aligned}$$

where $\overline{\mathbf{U}_{\mathbf{s}}}$ is the mean difference channelized signal, and $\mathbf{K}_{\mathbf{U}}$ is the mean covariance of channelized images [18], [19]. The values of t determines whether H_0 or H_1 is more likely to be true.

In this study, we have a sequence of cyclopean views as the image data; thus, it is necessary to efficiently integrate the information from multiple images. Plastivsa et al. [18] compared three designs of msCHOs and found that the one that feeds the channelized data directly into a HO for a final observer score (as shown in Fig. 4a) performed best. Thus, we use a similar design for our msCHO and consider the cyclopean views in the sequence as correlated data. We apply 5 Laguerre-Gauss channels [20](as shown in Fig. 4b) on the sequence of cyclopean views. The templates $\mathbf{w}_{\mathrm{HO}}$ are estimated from the training set, then applied to the unseen testing set for computing the decision statistics t. The performance of the model is evaluated with Receiver Operating Characteristic (ROC) analysis. Decision thresholds in the range of $[0, 1]$ on the continuous statistic t are used to generate binary decisions for each sequence of stereo pairs being tested. If t is lower than the threshold, the image is decided as signal-absent; otherwise, it is decided as signal-present. ROC curves, showing the fraction of true positives (TPF) versus the fraction of false positives (FPF) at various thresholds, are plotted. The accuracy in terms of the area under the ROC (AUC) is calculated as the figure of merit.



(a)



(b)

Fig. 4. (a) msCHO: a sequence of cyclopean views $\{\mathbf{g}_1, \mathbf{g}_2, ...\mathbf{g}_N\}$ is channelized then used as input $\mathbf{U}_{\mathrm{msCHO}}$ to the integration stage for calculate decision statistics t; (b) Images of the first 5 Laguerre-Gauss channels [18].

III. Results

A. Varying the angular span Θ

Results of ROC analysis conducted on the testing sets are shown in Table I. For each of the testing voxels, three sequences of stereo pairs were generated with the angular spans Θ of $10°, 15°$, and $20°$. Table I shows that the model observer achieves the highest AUCs of 0.890 when $\Theta = 15°$. However, the AUCs achieved for different Θs are not statistically different, even given the moderate sample size used in the study. For instance, the two-sided test for the null hypothesis $\mathrm{AUC}(\Theta = 10) = \mathrm{AUC}(\Theta = 15)$ results in a p-value of 0.5013. The sample size of the study was sufficient since the post hoc power analysis suggests that a difference of 0.035 in the AUCs should have been detected with a power of 0.8. The AUCs of the model observer are slightly lower than those reported in prior observer studies (e.g., [2], [3], [12]) with single stereo pair. This may due to the instabilities in estimating the template and high-dimensional covariance matrix from the correlated image data [18], which affects the msCHO in the calculation of decision statistics. The conclusions regarding the angular span are consistent with the choice used in clinical practice: a typical DBT tomosynthesis scan with Hologic Selenia system produces 15 projection images taken about $1.08°$ apart, which covers an approximate angular range of 15 degrees [13]. ROC curves for an example voxel dataset are shown in Fig. 5.

B. Varying the exponential coefficient β

The results of ROC analysis conducted on the testing sets are shown in Table II. For all the testing voxels, the stereo pairs were generated with an angular span of $15°$. Table II shows an overall trend that larger values of β yield smaller AUCs. The model observer achieves the highest AUC of 0.900 when $\beta = 2$, while the lowest AUC of 0.821 when $\beta = 3$. Over the range of β values investigated, larger differences in β resulted in statistically different AUCs whereas smaller differences in β resulted in statistically equivalent AUCs. For instance, the one-sided test for the null hypothesis $\mathrm{AUC}(\beta = 2) = \mathrm{AUC}(\beta = 3)$ results in a p-value of 0.003. On the other hand, the one-sided testing for the null hypothesis AUC($\beta =$

TABLE I

AVERAGE AUCS WITH DIFFERENT ANGULAR SPANS Θ

Θ	10	15	20
AUC	0.878 ± 0.012	0.890 ± 0.011	0.887 ± 0.013

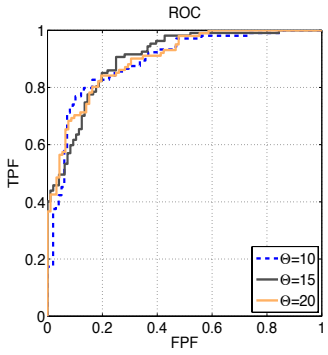


Fig. 5. ROC curves with angular span Θ of 10,15,and 20 degrees. The AUCs for these three angular spans are statistically indistinguishable in this study, implying that there is a range of suitable angular spans.

$2) = \text{AUC}(\beta = 2.25)$ results in a p-value of 0.7792. Again, the sample size was sufficient as the post hoc power analysis suggests that a 0.03 difference in AUCs should have been detected with a power of 0.8. Given that larger βs correspond to denser breasts, the results suggest that increasing breast density results in a reduction of breast lesion detectability in this presentation approach. This is consistent with human observer studies with 2D mammograms (e.g., [6], [9]) and our prior model observer study with DBT [3]. Example ROC curves for different β are shown in Fig. 6.

IV. CONCLUSIONS

We conduct a simulation study to computationally assess the efficacy of stereo viewing a rotating sequence of DBT projections. We show that the model observer can generate reliable decision statistics with this presentation approach.

ACKNOWLEDGMENT

We would like to thank David J. Getty for suggesting to us stereo viewing of a rotating sequence of projections.

REFERENCES

[1] A. Smith, "Fundamentals of breast tomosynthesis."
[2] G. Muralidhar, M. Markey, A. Bovik, T. Haygood, T. Stephens, W. Geiser, N. Garg, B. Adrada, B. Dogan, S. Carkaci, R. Khisty, and G. Whitman, "Stereoscopic interpretation of low-dose breast tomosynthesis projection images," *Journal of Digital Imaging*, vol. 27, no. 2, pp. 248–254, 2014.
[3] G. Wen, M. K. Markey, and G. S. Muralidlhar, "A stereo matching model observer for stereoscopic viewing of 3d medical images," in *SPIE Medical Imaging*, vol. 9037, 2014, pp. 90 370Z–90 370Z–8.
[4] F. Zafar, J. Dorband, and A. Badano, "Computational observer approach for the assessment of stereoscopic visualizations for 3d medical images," in *SPIE Medical Imaging*, 2012, pp. 831 806–831 806.
[5] C. J. DOrsi, D. J. Getty, R. M. Pickett, I. Sechopoulos, M. S. Newell, K. R. Gundry, S. R. Bates, R. M. Nishikawa, E. A. Sickles, A. Karellas *et al.*, "Stereoscopic digital mammography: Improved specificity and reduced rate of recall in a prospective clinical trial," *Radiology*, vol. 266, no. 1, pp. 81–88, 2013.
[6] A. E. Burgess, F. L. Jacobson, and P. F. Judy, "Human observer detection experiments with mammograms and power-law noise," *Medical physics*, vol. 28, p. 419, 2001.

TABLE II

AVERAGE AUCS WITH DIFFERENT VALUES OF β. THE ANGULAR SPAN OF STEREO PAIRS IS SET TO 15° FOR ALL THE SEQUENCES.

β	2	2.25	2.5	2.75	3
AUC	0.900 ± 0.011	0.899 ± 0.012	0.878 ± 0.011	0.842 ± 0.007	0.821 ± 0.012

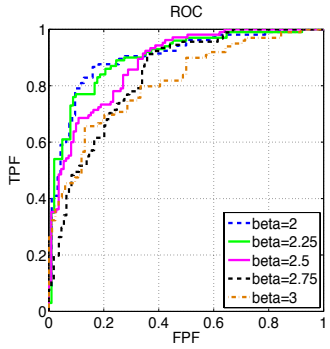


Fig. 6. ROC with varying beta and fixed separation angle 8 degree. The ROC with $\beta = 2$ has the highest AUC among the five ROC curves, while the ROC with $\beta = 3$ has the lowest AUC.

[7] I. Andersson, D. M. Ikeda, S. Zackrisson, M. Ruschin, T. Svahn, P. Timberg, and A. Tingberg, "Breast tomosynthesis and digital mammography: a comparison of breast cancer visibility and birads classification in a population of cancers with subtle mammographic findings," *European radiology*, vol. 18, no. 12, pp. 2817–2825, 2008.
[8] X. Gong, S. J. Glick, B. Liu, A. A. Vedula, and S. Thacker, "A computer simulation study comparing lesion detection accuracy with digital mammography, breast tomosynthesis, and cone-beam ct breast imaging," *Medical physics*, vol. 33, p. 1041, 2006.
[9] J. G. Mainprize, A. H. Tyson, and M. J. Yaffe, "The relationship between anatomic noise and volumetric breast density for digital mammography," *Medical Physics*, vol. 39, p. 4660, 2012.
[10] J. G. Mainprize and M. J. Yaffe, "A breast density-dependent power-law model for digital mammography," in *Breast Imaging*, 2012, pp. 761–768.
[11] J. Shorey, "Stochastic simulations for the detection of objects in three," Ph.D. dissertation, Duke University, 2007.
[12] L. J. Webb, E. Samei, J. Y. Lo, J. A. Baker, S. V. Ghate, C. Kim, M. S. Soo, and R. Walsh, "Comparative performance of multiview stereoscopic and mammographic display modalities for breast lesion detection," *Medical physics*, vol. 38, p. 1972, 2011.
[13] A. Smith, "Design considerations in optimizing a breast tomosynthesis system." [Online]. Available: http://www.hologic.com/data/Design_Considerations_Optimizing_Breast_Tomo.pdf
[14] W. Wright, "Foundations of cyclopean perception," *Journal of Modern Optics*, vol. 19, no. 6, pp. 550–550, 1972.
[15] A. Klaus, M. Sormann, and K. Karner, "Segment-based stereo matching using belief propagation and a self-adapting dissimilarity measure," in *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, vol. 3, 2006, pp. 15–18.
[16] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms," *International journal of computer vision*, vol. 47, no. 1-3, pp. 7–42, 2002.
[17] J. Sun, N.-N. Zheng, and H.-Y. Shum, "Stereo matching using belief propagation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 25, no. 7, pp. 787–800, 2003.
[18] L. Platiša, B. Goossens, E. Vansteenkiste, S. Park, B. D. Gallas, A. Badano, and W. Philips, "Channelized hotelling observers for the assessment of volumetric imaging data sets," *JOSA A*, vol. 28, no. 6, pp. 1145–1163, 2011.
[19] A. E. Burgess, "Visual perception studies and observer models in medical imaging," in *Seminars in Nuclear Medicine*, vol. 41, no. 6, 2011, pp. 419–436.
[20] H. H. Barrett, C. K. Abbey, B. D. Gallas, and M. P. Eckstein, "Stabilized estimates of hotelling-observer detection performance in patient-structured noise," in *Medical Imaging'98*, 1998, pp. 27–43.