Journal of Medical Imaging

MedicalImaging.SPIEDigitalLibrary.org

Assessing CT acquisition parameters with visual-search model observers

Zohreh Karbaschi
Howard C. Gifford

Assessing CT acquisition parameters with visual-search model observers

Zohreh Karbaschi and Howard C. Gifford*

University of Houston, Department of Biomedical Engineering, Houston, Texas, United States

Abstract. A principal difference between the channelized Hotelling (CH) and visual-search (VS) model observers is how they respond to noise texture in images. We compared the two observers in lesion-detection studies to evaluate linear and angular sampling parameters for CT. Simulated lung images were generated from a single two-dimensional mathematical torso phantom containing circular lesions of fixed radius and relative contrast. Projection datasets were produced for two detector pixel sizes and from 15 to 128 projections at 15 and 65 M counts per set. Filtered backprojection reconstructions were obtained with dimensions of 128×128 and 256×256 . A localization receiver operating characteristic study was conducted with two human observers, three single-feature VS observers, and a feature-adaptive VS observer. The effects of the sampling parameters on performance were similar for all of these observers. The CH observer, applied in location-known studies with and without background variability, was not affected by the variations in angular sampling. The two-stage VS framework was an effective modification of the CH observer for assessing the effects of noise texture on human-observer performance in this study. © 2018 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JMI.5.2.025501]

Keywords: computed tomography; model observer; visual search; image quality; data sampling; lesion detection; task-based assessment; tomographic imaging.

Paper 17145PRR received May 22, 2017; accepted for publication Mar. 12, 2018; published online Apr. 5, 2018.

1 Introduction

Noise texture is a primary factor in determining the quality of tomographic medical images. Chiefly comprised of quantum and anatomical noise, the texture is influenced by numerous components of the imaging chain, including patient habitus, data acquisition, the reconstruction method, and subsequent postprocessing. For example, aliasing artifacts¹ may occur due to undersampling while image smoothing can enhance noise correlations. Our focus is on how sampling affects image quality as quantified by lesion-detection performance. In this work, human and model observers read simulated two-dimensional (2-D) lung CT images for which the acquisition was configured by the number of angular projections and the linear detector dimension. Human performance was considered the gold standard for assessing the acquisition quality.

CT researchers have often examined aspects of the data acquisition using technical metrics, such as spatial resolution and contrast-to-noise ratio.^{2–5} In evaluating task performance instead, we tested two approaches for using model observers as stand-ins for human observers. One approach made use of the channelized Hotelling (CH) observer⁶ in location-known studies with a fixed lesion profile. This ideal linear observer includes spatial frequency channels to model the human visual system and reduce the computational dimensions. The CH observer has been used to investigate acquisition protocols for digital breast tomosynthesis^{7–9} (DBT) and CT.¹⁰ Modeling human observers may require the addition of internal noise to degrade the prior knowledge that the CH observer brings to a task. For the second approach, we applied visual-search (VS) observers.¹¹ These nonlinear observers, which implement a two-stage search-and-analysis process that approximates how radiologists read

medical images, have also been used to evaluate DBT acquisition protocols.¹² The human and VS observers both took part in a localization receiver operating characteristic (LROC) study, allowing for direct comparison of their results. The CH observer read images with noise levels and lesion contrasts that were consistent with this LROC study. The objective with the CH observer was to see whether performance trends with changes in the sampling parameters were consistent with those from human observers.

The VS observers operated with lesion-like features extracted from the images. These features are also relevant to the construction of linear models such as the CH observer. The lesion profile, the profile modified by an eye filter, and a profile gradient were each used to form a single-feature VS observer. A fourth, feature-adaptive visual-search (FAVS) observer implemented the same set of visual channels as the CH observer. This FAVS observer used training images to identify an effective subset of these channels for search.

2 Methods

2.1 Framework for Model Observers

We represent a 2-D reconstructed image as vector $\mathbf{f}$. The image may contain a single lesion centered at some pixel within a region of interest (ROI). With a total of J locations, an observer chooses from hypotheses $H_0, \dots, H_J$, where the null hypothesis H_0 is that $\mathbf{f}$ consists solely of a noise-free background plus quantum noise. With each remaining hypothesis, the image also contains a lesion centered at the j 'th location. A single-lesion profile $\mathbf{s}$ was used for this study. The instance of $J = 1$ describes the location-known task undertaken by the CH observer.

*Address all correspondence to: Howard C. Gifford, E-mail: hgifford@uh.edu

The model observers in our study generated a scalar rating λ indicating lesion presence or absence in a given image. For the location-known study, observer performance was measured by the detectability index

$$d' = \frac{\langle \lambda \rangle_1 - \langle \lambda \rangle_0}{[0.5\text{var}(\lambda)_1 + 0.5\text{var}(\lambda)_0]^{\frac{1}{2}}}, \quad (1)$$

that quantifies the separation between the conditional probability distributions of λ under the two hypotheses. The terms $\langle \lambda \rangle_i$ and $\text{var}(\lambda)_i$ ($i = 0, 1$) are, respectively, the mean and variance of λ under H_i .

For the LROC study, observers also indicated the maximally suspicious location in an image. The rating and probable location were, respectively, the max and argmax of a perception measurement λ_j computed over the ROI pixels. We used the area under the LROC curve (A_L) as the performance figure of merit. Further details on the observer studies are in Sec. 2.6.

2.2 Channelized Hotelling Observer

The CH observer computed an image rating as

$$\lambda_{\text{CH}} = (\bar{\mathbf{f}}_1 - \bar{\mathbf{f}}_0)^t \mathbf{U} \mathbf{K}^{-1} \mathbf{U}^t \mathbf{f}, \quad (2)$$

where $\bar{\mathbf{f}}_i$ ($i = 0, 1$) is the statistical mean image under hypothesis H_i and the superscript t denotes a transpose. The columns of matrix $\mathbf{U}$ contain the spatial responses for a group of spatial-frequency channels. A sparse set of three difference-of-Gaussian (DOG) channels originated by Abbey and Barrett¹³ was selected. The spatial response for a given channel was obtained as the discrete Fourier transform of the discretized channel. Also in Eq. (2), the 3×3 composite channel covariance matrix $\mathbf{K}$ was computed as the arithmetic average of the ensemble covariance matrices

$$\mathbf{K}_i = \langle \mathbf{U}^t [\mathbf{f} - \bar{\mathbf{f}}_i] [\mathbf{f} - \bar{\mathbf{f}}_i]^t \mathbf{U} \rangle_{\mathbf{f}|H_i}, \quad i = 0, 1, \quad (3)$$

under the two hypotheses. The bracket notation $\langle \cdot \rangle_{\mathbf{f}|H_i}$ indicates a conditional average over the images under hypothesis H_i . For a background-known-exactly (BKE) task, $\mathbf{K}$ is computed based on a single noise-free background. In a background-known-statistically (BKS) task, background variations factor into the matrix.¹¹

2.3 Visual-Search Observers

A VS observer applies a sequential search-and-analysis process to detect lesions. The search step finds suspicious locations (or candidates), possibly identifying false positives while missing actual lesions. The search is, therefore, implicitly affected by noise texture, and by extension, the choice of sampling parameters. Following the search, analysis of the candidates yields the most-probable lesion location and a confidence rating.

For one set of VS observers in this work, the search was guided by morphological features derived from knowledge of the lesion profile $\mathbf{s}$. Each feature generates a correlation map and local maxima in the map represent candidates. We define a generic correlation map by the 2-D cross correlation

$$\mathbf{z}_O = O\{\mathbf{s}\} ** O\{\mathbf{f}\}, \quad (4)$$

where the double asterisks denote the cross correlation and O represents an image processing operator applied to both the

lesion and the image. Specifically, $O\{\mathbf{s}\}$ represents a feature of interest. Three features were, respectively, defined by the identity operator, the 2-D gradient, and an approximate radial gradient represented by the Burgess eye filter.¹⁴ Figure 1 shows the candidate search process for a test image. Details on the test image generation are provided in Sec. 2.5.

The identity operator generated the matched-filter (MF) map [Fig. 1(b)] as the cross correlation of $\mathbf{s}$ and $\mathbf{f}$, whereas the 2-D gradient produced the gradient MF (GMF) map [Fig. 1(d)] as the sum

$$\mathbf{z}_\nabla = \left(\frac{\partial \mathbf{s}}{\partial x} ** \frac{\partial \mathbf{f}}{\partial x} \right) + \left(\frac{\partial \mathbf{s}}{\partial y} ** \frac{\partial \mathbf{f}}{\partial y} \right), \quad (5)$$

in which the right-hand terms denote the partial derivatives of $\mathbf{s}$ and $\mathbf{f}$ in the x and y directions, respectively. Last, an MFe (MF + eye filter) map [Fig. 1(c)] used the convolution of $\mathbf{s}$ with the eye filter as the cross-correlation kernel in Eq. (4). The eye filter resembles a 2-D ramp filter at low and midrange frequencies but rolls off at high frequencies.

To identify candidates, each map was segmented using the built-in MATLAB watershed routine.¹⁵ The routine was applied to the additive inverse of the map, so that each watershed segment was associated with a local maximum. Those maxima within the ROI defined the candidates for the map. Note that the ROI delineation in Fig. 1(a) omitted several lesion-like structures that the human observers learned to ignore during training. Example map segmentations and the resulting candidates for the three MFs are, respectively, shown in the third and fourth columns of Fig. 1.

How a VS observer analyzes the candidates to render LROC data depends on how many features are used. In this work, each of the three features was used to construct a single-feature VS observer, which we shall refer to as the MF, MFe, and GMF observers. The analysis for a single-feature model simply entailed selecting the candidate with the largest feature value. A feature-adaptive VS observer, described next, used multiple features, and the candidate analysis was carried out with a linear discriminant.

2.4 Feature-Adaptive Visual-Search Observer

Size, shape, and intensity are among the physical characteristics commonly recognized as significant for lesion detection. However, quantifying any one of these characteristics usually requires assumptions about the others. Similarly, the correlation maps for the morphological features discussed in Sec. 2.3 provide relative measures of these characteristics that are conditioned on knowledge of the lesion profile. The MF map computes a weighted sum of image intensity within lesion-sized regions, whereas the MFe and GMF maps quantify similarity with the lesion size and shape.

Yet humans might be expected to use more than just a single feature. Starting with a base set of features, an FAVS observer selectively chose a subset of features based on training performance. Although the model permits adaptive selection for both the search and the analysis, only the search adaption was applied for this work.

The FAVS model used radial gradient features, but instead of defining these features as operators applied to the known lesion profile as in Eq. (4), the k 'th feature instead took the form $O\{\hat{\mathbf{s}}_k\}$, where $\hat{\mathbf{s}}_k$ represented a presumed lesion profile. The correlation maps were thus of the form

Fig. 1 Search process for the VS observers. (a) A study image containing the lungs and heart. A circular lesion is indicated by the arrow. The dashed lines bound the search ROI. (b)–(d) The MF, MFe, and GMF correlation maps. (e)–(g) Watershed segmentations of the three maps. Each black segment represents a watershed; the white lines are dams. (h)–(j) The candidates defined as the local maxima of the watersheds.

$$\mathbf{z}_k = O\{\hat{\mathbf{s}}_k\} * O\{\mathbf{f}\}, \quad (6)$$

using 2-D Gaussians with standard deviations of 0.8, 1.4, and 3.0 cm as the presumed profiles. The choice of these profiles stemmed from how Eq. (6) was computed. An efficient way is to consider the equivalent equation

$$\mathbf{z}_k = O^\dagger O\{\hat{\mathbf{s}}_k\} * \mathbf{f}, \quad (7)$$

with the adjoint operator $O^\dagger$. When O is a spatial derivative, then $O^\dagger O$ is the negative second derivative, which can be approximated for a given Gaussian profile by a specific DOG function. For the three Gaussians described above, the DOG channels used by the CH observer are appropriate. The FAVS maps were thus computed by cross correlating $\mathbf{f}$ with the spatial responses of the DOG channels.

The adaptive feature selection was conducted by computing the fraction of lesion-present training images for which the lesion was included as a candidate. Candidates within a radius R_{CL} (the radius of correct localization) about a lesion center were associated with the lesion. (This same radius was also used to judge observer localizations in the LROC study; how R_{CL} was set is described in Sec. 2.6.) A feature was discarded for the testing phase of the study when this fraction was below a selection threshold δ . Values of δ between 0.75 and 0.95 were tested. The preserved features generated candidates as described in Sec. 2.3. With multiple features, the set of candidates for an image was formed by aggregating the candidates from the correlation maps.

The candidate analysis used all three features to construct a shift-invariant linear discriminant. The discriminant was calculated from candidate statistics in the lesion-absent and lesion-present training images, based on the assumption that the candidate feature values were distributed according to location-marginalized multinormal probabilities. We let the three-element vectors $\bar{\mathbf{p}}_1$ and $\bar{\mathbf{p}}_0$ represent the respective mean feature vectors for lesion-present and lesion-absent candidates. With $\mathbf{K}_p$ the composite feature covariance matrix calculated as the average of the covariance matrices from the lesion-absent and lesion-present candidates, a Hotelling test statistic at the j 'th candidate can be calculated as

$$\lambda_{FAVS,j} = (\bar{\mathbf{p}}_1 - \bar{\mathbf{p}}_0)' \mathbf{K}_p^{-1} \mathbf{p}_j, \quad (8)$$

where $\mathbf{p}_j$ is the feature vector for the j 'th candidate. For this study, Eq. (8) was applied with the full covariance matrix and also with only the diagonal values as computed from lesion-absent candidates alone. The results presented herein were acquired using the diagonal matrix as little difference was found between the two approaches. Maximizing $\lambda_{FAVS,j}$ over the candidates produced the LROC data for the image.

2.5 Imaging Simulation

Our studies explored data acquisition and reconstruction for 2-D images. We started with a portion of the three-dimensional extended cardiac torso (XCAT) phantom,¹⁶ with the subvolume of interest containing the lungs and heart in the torso. This

Fig. 2 FBP study images obtained from a single lesion-present phantom for count level 65 M and $N = 256$. (a)–(c) The number of projection angles was 20, 30, and 64. The lesion location is indicated by the arrow in (c).

subvolume was generated as a 2048^3 array with a voxel size of 0.026 cm. A transverse 2-D XCAT slice that included both the lungs and the heart was then extracted as a background phantom. To prepare for the LROC study, circular soft-tissue lesions of radius 1 cm and a relative intensity of 2.0 were added to the lungs to form 500 single-lesion phantoms. Each phantom had a unique lesion center location randomly sampled from a map of the lung pixels. (This map also represented the search region for the VS observers, as shown in Fig. 1.) The lesion locations were equally divided between the left and right lungs.

For the location-known study, the same XCAT slice was used as the background phantom. Phantoms were also created from an additional five slices containing lung. Reconstructions from all six slices were used to train the CH observer for a BKS task.

The MATLAB radon routine was used to create noise-free sets of 1-D projections with $P = 15, 20, 30, 64$, and 128 projections over 180 deg. Note that the lesion projections were generated independently and were then combined with the background projections. The projections were then down-sampled to detector bin dimensions of $N = 128$ (pixel size of 0.42 cm) and 256 (pixel size of 0.21 cm). Poisson noise based on mean total acquisitions of 15 and 65 million counts was then added to the projections, and 2-D reconstructions were generated with filtered backprojection (FBP). As the 128×128 reconstructions for $N = 128$ were difficult for humans to read, these were resized to 256×256 using linear interpolation. Sample FBP images are shown in Fig. 2 for $N = 256$ and $P = 20, 30$, and 64 projection angles.

2.6 Observer Studies

Two nonradiologist imaging scientists read high-count LROC image sets for all P . For each combination of N and P , there were 50 training images followed by 100 test images. These training and test sets both contained equal numbers of lesion-present and lesion-absent images. The image-set reading order varied with observer and the reading order of the images within a set was randomized. The observers provided confidence ratings on a four-point scale.

Correct localizations were scored using the same localization radius R_{CL} that was used for feature selection by the FAVS observer. This threshold was set by an empirical process involving localizations from the human observers such that small changes in R_{CL} would not substantially alter an observer's calculated performance.¹⁷ Individual observer performance was computed for a given $N - P$ combination by numerical integration of the empirical LROC curve, and overall performance was

recorded as the average over the two observers. Uncertainties in overall performance are reported as $\pm$ one standard error.

The VS observers also read the high-count images, using 200 training images and 800 test images for each parameter pair, with both sets equally divided between lesion-present and lesion-absent images. The training images were used to estimate the lesion profile s and (for the FAVS observer) the feature covariance matrix of Eq. (8). The observer rating data were preserved as floating-point values and the same R_{CL} value was used as in the human-observer study. A three-way analysis of variance (ANOVA) was conducted on the results, with projection angles, detector dimension, and observers (the VS and individual human observers) as fixed factors. Pairwise differences between the observers were examined with the Tukey honest significant differences (HSD) test.¹⁸

The CH observer was applied for both the low-count and high-count images in separate BKE and BKS studies. The observer read 800 test images per parameter combination. A separate set of 2000 images was used to compute the means and covariance matrix of Eq. (2). More training images were used than for the FAVS observer, which only required estimation of the feature variances while extracting more information per image. The testing and training sets were equally split between lesion-present and lesion-absent images. For the BKS task, the training images represented six different XCAT slices. Note that the BKS test images featured the same single background as the BKE study and the LROC study.

3 Results

3.1 Human Observer Results

The LROC results for the human observers are summarized in Table 1. An R_{CL} value of six pixels (1.3 cm, roughly equal to the lesion radius) was used for the calculations. The uncertainties in individual performance ranged from 0 to ± 0.03 , increasing for lower values of P . Average performance was largely the same for both values of N , varying substantially only for 20 projection angles. The lowest performance was obtained for $P = 15$ projection angles, increasing monotonically to a plateau of about 1.0 for $P \geq 64$.

3.2 Single-Feature VS Observer Results

As shown in Fig. 3, the performance trends for the MF, GMF, and MFe observers tracked the trends in average human performance, although none of the models agreed exactly with the humans. The linear correlation coefficients between the average human and the model observers for $N = 128$ were 0.81

Table 1 Individual and average human-observer LROC performance (A_L) as a function of the number of projections P and the detector dimension N . The test images were based on acquisitions with 65 M counts.

Observer	P				$N = 128$
	15	20	30	64	128
Human #1	0.72	0.77	0.97	0.99	1.00
Human #2	0.57	0.63	0.92	0.99	1.00
Average	0.65 ± 0.07	0.70 ± 0.07	0.95 ± 0.03	0.99 ± 0.01	1.00
$N = 256$					
Human #1	0.70	0.89	0.92	1.00	1.00
Human #2	0.62	0.78	0.95	1.00	1.00
Average	0.66 ± 0.04	0.84 ± 0.05	0.94 ± 0.02	1.00	1.00

Fig. 3 Average human and single-feature VS observer LROC performance for (a) $N = 128$ and (b) $N = 256$.

(MF observer), 0.97 (GMF observer), and 0.98 (MFe observer). For $N = 256$, the respective coefficients were 0.94, 0.99, and 0.99. For a given P , the MFe and GMF observers, both sensitive to edge structure, consistently scored higher using the larger detector dimension. The effect was less pronounced for the MF observer. Overall, the MF observer showed the best quantitative agreement with the average human scores.

Figure 4 shows how the average number of candidates per image varied as a function of the numbers of acquisition counts and projection angles P for the three VS observers. The test images for this plot were generated with $N = 256$, but similar results were observed for $N = 128$. The shape of the curves corresponds with the shape of the performance curves of Fig. 3, with the numbers of candidates reaching a plateau between 30 and 64 angles. As would be expected, increasing the acquisition counts decreased the number of candidates. For a given P , the GMF observer generated the most candidates, whereas the MFe observer generated the fewest.

3.3 Channelized Hotelling Observer Results

The outcomes from the location-known studies with the CH observer are shown in Fig. 5. These plots summarize the impact of sampling parameters N and P , the task paradigm (BKE/BKS), and the acquisition counts on observer performance.

At both count levels with the BKE task [Fig. 5(a)], the effects of N and P on performance were negligible. For the BKS task [Fig. 5(b)], performance steadily increased with larger N , but there was no discernible trend with respect to P .

Fig. 4 Mean number of candidates per image as a function of P for the single-feature observers with $N = 256$. Both low and high-count images were analyzed, although only the latter were used in the LROC study.

Fig. 5 Detectability index values for the CH observer as a function of acquisition counts and parameters P and N for (a) the BKE task and (b) the BKS task.

3.4 Feature-Adaptive VS Observer Results

Figure 6 shows the LROC results for the FAVS observer. The plots show observer performance as a function of N and P for selection threshold values $\delta = 0.80, 0.85$, and 0.90 . The performance of the individual human observers is also indicated. For δ in the range $[0.75, 0.95]$, model observer performance fell with decreasing δ as less-effective features contributed to the search. (Recall that all of the FAVS observers used a three-element discriminant to analyze candidates.) Thresholds outside of the above interval did not further affect performance.

At lower numbers of projection angles, the radial gradient of the widest Gaussian profile was most often selected as an efficient feature. The FAVS models outperformed the single-feature observers for $P = 15$, whereas the range of scores for other P values was similar for both types of VS observers. With the two lowest P values, the FAVS observers performed best with $N = 256$, but N had less effect with higher values of P .

3.5 ANOVA Results

Ten values of A_L (five P values per N value) for each of the eight human and VS observers were input into the three-way ANOVA. The results are shown in Table 2. All of the primary and two-way interaction effects were found to be significant at the $\alpha = 0.05$ level. The primary effects were particularly strong. A posthoc Tukey HSD test found a significant difference

between the two human observers: observer #2 was most similar to the FAVS observers with the two lower δ values, whereas observer #1 compared most favorably with the remaining FAVS observer and the MF observer. As shown in Figs. 3 and 6, all of the observers had the same general response to increases in P and N ; the significant interactions involving the observers were due to variations in the magnitude of the response.

4 Discussion

We tested two approaches for using model observers to evaluate basic tomographic acquisition protocols. The premise for the CH study is that outcomes with location-known tasks are predictive of results with more-complex tasks.^{19,20} However, no particular performance trends were noted with respect to sampling parameter P for this observer, even when increased quantum noise and background uncertainty were incorporated into the task. With the BKS task alone, the CH observer did show consistently better performance with increased detector dimension N . The VS observers, applied to the LROC task of interest, demonstrated overall agreement with how the human observers responded to changes in N and P .

At issue is the substantial prior knowledge that a model observer may bring to a task. Information about lesion geometry and location, as well as background structure, is central to ideal observers such as the CH model. Internal noise can provide some counterbalance but is not meant to compensate for excessive prior knowledge. This problem can persist with search

Fig. 6 Comparison of individual human and FAVS observer performance as a function of P for (a) $N = 128$ and (b) $N = 256$. The model observer used selection threshold values $\delta = 0.80, 0.85$, and 0.90 .

Table 2 Three-way ANOVA results for the human and VS observer scores. All factors were significant at the $\alpha = 0.05$ level.

Factor	SS	df	F	Pr(> F)
Detector dimension (<i>N</i>)	0.02	1	41.26	<10 ⁻⁴
Observers (<i>O</i>)	0.17	7	45.7	<10 ⁻⁴
Projections (<i>P</i>)	1.85	4	893.69	<10 ⁻⁴
<i>N</i> × <i>O</i>	0.01	7	3.01	0.017
<i>N</i> × <i>P</i>	0.01	4	5.89	0.0014
<i>O</i> × <i>P</i>	0.22	28	15.28	<10 ⁻⁴
Error	0.01	28		
Total	2.29	79		

tasks.¹¹ Alternatively, sources of noise and uncertainty are more easily integrated into the dual stages of the quasi-ideal VS observer. For example, the impact of noise texture is consistently defined through the candidate search, whereas inserting anatomical noise for the CH observer required a supposition of background variations. An explicit search task is not required, as the VS observer can also be applied for location-known tasks.²¹

The VS observers operated with features that are commonly used to construct other model observers.²² The features engendered by the MF, GMF, and MFe operations relate to relevant lesion characteristics such as intensity and size of image structures. Moreover, eye-tracking experiments with simulated DBT images²³ have found high correlations between human-observer dwell times at fixation points and the corresponding feature values.

Testing of the FAVS observer is another important contribution of this work. For images from a given *N* – *P* pair, this observer was trained to select relatively effective search features from a base set of features. (Allowing different features to be used for the search and analysis acknowledges that human observers may do the same.) Changing the selection threshold δ for the FAVS observer produced performance variations on par with the variations between our human observers. In assessing this result, one should realize that with the exception of the R_{CL} threshold estimation, the human-observer data played no role in the VS observer training. Such data would be required were the model observers expected to mimic specific human observers.

Substantial differences between the CH and FAVS observers might be obscured by the fact that they used the same DOG channels and implemented linear discriminants based on channel statistics. That was practically the extent of their similarity. For the CH observer, the channels simulated the human visual system. With the FAVS observer, the channels implemented gradient features for three presumed lesion profiles, adding uncertainty to the detection task that the CH and single-feature observers did not contend with.

Overall, this work provides basic insights into the use of model observers for task-based studies of medical image quality, but as an assessment of CT acquisition protocols, there are clear limitations. The physics of the imaging simulation was relatively simple compared with the realities of CT imaging.

Also, case variability was minimal and the LROC study had but two participants. Future studies with an upgraded imaging simulation will employ additional observers with the goal of investigating adaptive model observer correlations across a wider spectrum of human abilities and experience levels.

5 Conclusions

This work has compared two approaches to assessing essential CT acquisition parameters with model observers. Variations between the human observers matched those obtained with a feature-adaptive version of the VS observer. These results demonstrate the usefulness of a two-stage observer model for assessing the effects of noise textures on the detection-task performance of humans.

Disclosures

The authors have no conflicts of interest, financial or otherwise, to declare.

Acknowledgments

The authors thank Dr. Anando Sen for his contributions to the statistical analysis in this study.

References

1. A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*, Classics in applied mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia (2001).
2. J. M. Boone, "Determination of the presampled MTF in computed tomography," *Med. Phys.* **28**, 356–360 (2001).
3. P. F. Judy, "The line spread function and modulation transfer function of a computed tomographic scanner," *Med. Phys.* **3**, 233–236 (1976).
4. R. Grimmer et al., "Assessment of spatial resolution in CT," in *IEEE Nuclear Science Symp. and Medical Imaging Conf. Record*, pp. 5562–5566 (2008).
5. K. Tang et al., "Effect of low tube voltage on image quality, radiation dose, and low-contrast detectability at abdominal multidetector CT: phantom study," *J. Biomed. Biotechnol.* **2012**, 1–6 (2012).
6. K. J. Myers and H. H. Barrett, "Addition of a channel mechanism to the ideal-observer model," *J. Opt. Soc. Am. A* **4**, 2447–2457 (1987).
7. A. S. Chawla et al., "Optimized image acquisition for breast tomosynthesis in projection and reconstruction space," *Med. Phys.* **36**, 4859–4869 (2009).
8. S. Park, G. Z. Zhang, and R. Zeng, "Comparing observer models and feature selection methods for a task-based statistical assessment of digital breast tomosynthesis in reconstruction space," *Proc. SPIE* **9037**, 90370M (2014).
9. E. Samei et al., "A case for wide-angle breast tomosynthesis," *Acad. Radiol.* **22**, 860–869 (2015).
10. D. Racine et al., "Objective assessment of low contrast detectability in computed tomography with channelized Hotelling observer," *Phys. Med.* **32**, 76–83 (2016).
11. A. Sen and H. C. Gifford, "Accounting for anatomical noise in search-capable model observers for planar nuclear imaging," *J. Med. Imaging* **3**, 015502 (2016).
12. H. C. Gifford, Z. Liang, and M. Das, "Visual-search observers for assessing tomographic x-ray image quality," *Med. Phys.* **43**, 1563–1575 (2016).
13. C. K. Abbey and H. H. Barrett, "Human and model-observer performance in ramp spectrum noise: effects of regularization and object variability," *J. Opt. Soc. Am. A* **18**, 473–488 (2001).
14. A. E. Burgess, "Statistically defined backgrounds: performance of a modified nonprewhitening observer model," *J. Opt. Soc. Am. A* **11**, 1237–42 (1994).
15. J. B. Roerdink and A. Meijster, "The watershed transform: definitions, algorithms and parallelization strategies," *Fundam. Inf.* **41**, 187–228 (2001).

16. W. Segars et al., "4D XCAT phantom for multimodality imaging research," *Med. Phys.* **37**, 4902–4915 (2010).
17. R. G. Wells et al., "Effect of filtering on the detection and localization of small Ga-67 lesions in thoracic single photon emission computed tomography images," *Med. Phys.* **26**, 1382–1388 (1999).
18. J. L. Myers and A. D. Wahl, *Research Design and Statistical Analysis*, Lawrence Erlbaum Associates, Mahwah, New Jersey (2003).
19. A. E. Burgess and H. Ghandeharian, "Visual signal detection. II. Signal-location identification," *J. Opt. Soc. Am. A* **1**, 906–910 (1984).
20. M. P. Eckstein and C. K. Abbey, "Model observers for signal-known-statistically tasks (SKS)," *Proc. SPIE* **4324**, 91–102 (2001).
21. H. C. Gifford et al., "Visual-search model for location-known detection tasks," *Proc. SPIE* **10136**, 1013612 (2017).
22. J. Beutel, H. L. Kundel, and R. L. Van Metter, *Handbook of Medical Imaging: Physics and Psychophysics*, Vol. **1**, Chapter 10, SPIE Press, Bellingham, Washington (2000).
23. Z. Jiang, M. Das, and H. C. Gifford, "Analyzing visual-search observers using eye-tracking data for digital breast tomosynthesis images," *J. Opt. Soc. Am. A* **34**, 838–845 (2017).

Zohreh Karbaschi received her master's degree in biomedical engineering from the University of Toledo in 2012. She worked in an independent imaging core lab, Medical Metrics Inc., from 2012 to 2014 as a technical services associate. She started her PhD in the Department of Biomedical Engineering at the University of Houston in 2015. As a PhD candidate, her research includes task-based image quality assessment and optimization methods for CT imaging.

Howard C. Gifford received his PhD in applied mathematics from the University of Arizona in 1997. He was a postdoctoral fellow in the Radiology Department at the University of Massachusetts Medical School (UMMS) until 2000 and a UMMS faculty member through 2011. He is an associate professor in the Department of Biomedical Engineering at the University of Houston. His research interests include task-based image quality assessment and optimization methods for nuclear medicine and x-ray imaging.