Evaluation of clinical full field digital mammography with the task specific system-model-based Fourier Hotelling observer (SMFHO) SNR

Haimo Liu(a)
FDA Center for Devices and Radiological Health, Silver Spring, Maryland 20993-0002 and Department of Bioengineering, University of Maryland, College Park, Maryland 20742

Kish Chakrabarti and Richard V. Kaczmarek
US FDA Center for Devices and Radiological Health, Silver Spring, Maryland 20993-0002

Luis Benevides
Radiological Controls, Naval Sea Systems Command, Washington, DC 20376

Songxiang Gu
US FDA Center for Devices and Radiological Health, Silver Spring, Maryland 20993-0002

Iacovos S. Kyprianou(b)
US FDA Center for Devices and Radiological Health, Silver Spring, Maryland 20993-0002 and Department of Bioengineering, University of Maryland, College Park, Maryland 20742

(Received 3 October 2013; revised 7 February 2014; accepted for publication 17 March 2014; published 21 April 2014)

Purpose: The purpose of this work is to evaluate the performance of the image acquisition chain of clinical full field digital mammography (FFDM) systems by quantifying their image quality, and how well the desired information is captured by the images.

Methods: The authors present a practical methodology to evaluate FFDM using the task specific system-model-based Fourier Hotelling observer (SMFHO) signal to noise ratio (SNR), which evaluates the signal and noise transfer characteristics of FFDM systems in the presence of a uniform polymethyl methacrylate phantom that models the attenuation of a 6 cm thick 20/80 breast (20% glandular/80% adipose). The authors model the system performance using the generalized modulation transfer function, which accounts for scatter blur and focal spot unsharpness, and the generalized noise power spectrum, both estimated with the phantom placed in the field of view. Using the system model, the authors were able to estimate system detectability for a series of simulated disk signals with various diameters and thicknesses, quantified by a SMFHO SNR map. Contrast-detail (CD) curves were generated from the SNR map and adjusted using an estimate of the human observer efficiency, without performing time-consuming human reader studies. Using the SMFHO method the authors compared two FFDM systems, the GE Senographe DS and Hologic Selenia FFDM systems, which use indirect and direct detectors, respectively.

Results: Even though the two FFDM systems have different resolutions, noise properties, detector technologies, and antiscatter grids, the authors found no significant difference between them in terms of detectability for a given signal detection task. The authors also compared the performance between the two image acquisition modes (fine view and standard) of the GE Senographe DS system, and concluded that there is no significant difference when evaluated by the SMFHO. The estimated human observer efficiency was 30 ± 5% when compared to the SMFHO. The results showed good agreement when compared to other model observers as well as previously published human observer data.

Conclusions: This method generates CD curves from the SMFHO SNR that can be used as figures of merit for evaluating the image acquisition performance of clinical FFDM systems. It provides a way of creating an empirical model of the FFDM system that accounts for patient scatter, focal spot unsharpness, and detector blur. With the use of simulated signals, this method can predict system performance for a signal known exactly/background known exactly detection task with a limited number of images, therefore, it can be readily applied in a clinical environment. © 2014 American Association of Physicists in Medicine. [http://dx.doi.org/10.1118/1.4870377]

Key words: mammography, FFDM, Hotelling observer, SNR, CDMAM

1. INTRODUCTION AND BACKGROUND

According to the current Class II Special Controls Guidance Document for Full Field Digital Mammography (FFDM) (U.S. FDA, 2012), any manufacturer intending to market a new FFDM device can provide assurances to the FDA that the performance of their device has met the recommendations of this guidance. Digital detector performance is characterized by measuring the detector modulation transfer function (MTF) and noise power spectrum (NPS), which describe the
detector resolution and noise, respectively. Furthermore, image quality is evaluated by determining the threshold of smallest detectable thickness of gold disks of a given diameter that are included in the contrast-detail phantom for mammography (CDMAM) by experienced readers. By plotting the threshold thickness (contrast) vs disk diameter (detail), CD curves are generated and the FFDM system image quality can be evaluated. In this paper, we present a methodology that can be used to estimate the CD curves of a FFDM system using a model observer. The model observer uses the MTF and NPS measured with an uniform phantom placed in the field of view (FOV). By including the uniform phantom, we can summarize properties that contribute to image quality, including scatter blur from the phantom, focal spot unsharpness, grid performance, and signal magnification, in a single metric.

It has been widely accepted that the detector performance can be quantitatively evaluated with the MTF and NPS. In order to take into account the effects of focal spot unsharpness and patient scatter, a number of scientists have contributed to generalizing the definitions of the MTF and NPS. Wagner et al. investigated the effects of patient scatter on system detectability recognizing the necessity of developing a comprehensive assessment approach. Muntz et al. studied focal spot size, magnification factor, and patient scatter and described their effects on image quality by combining them in a comprehensive function. Boone et al. introduced the scatter MTF that measures the spatial distribution of scatter. Following in the footsteps of Boone's research, Cooper et al. proposed an experimental methodology to measure the magnitude and spatial distribution of scattered radiation for FFDM systems. Doi and Rossman studied the effects of focal spot unsharpness and Shaw et al. introduced the focal spot MTF. In an effort to describe the complete system performance by considering both the focal spot unsharpness and scatter, Kyprianou et al. analytically generalized the MTF (GMTF) definition by separately defining the scatter MTF, the focal spot MTF, and the detector MTF. In the same papers, Kyprianou generalized the NPS (GNPS) definition by considering the magnification factor and noise contributions from scatter. In a parallel effort, Samei developed an experimental methodology for estimating the effective NPS and MTF, which accounts for the scatter magnitude, and evaluated its use in digital radiographic imaging systems.

2. MATERIALS AND METHODS

Figure 1 summarizes our approach for developing an experimental methodology for evaluating the performance of clinical FFDM image acquisition chain. The method uses an empirical model of the system that includes the effects of scatter blur and focal spot unsharpness as well as simulated signals to calculate the task specific system model-based Hotelling observer SNR in the spatial frequency domain (SMFHO SNR). The resulting SNR was used to generate CD curves in order to facilitate the comparison between systems and methods. We demonstrated the practicality and clinical applicability of the method on two clinical FFDM systems (the GE Senographe DS and the Hologic Selenia) and we compared our results with other published methods.

2.1. FFDM system descriptions

In this work, we evaluate two clinical FFDM systems with different underlying detector technologies and scatter removal grids: a GE Senographe DS located at the National Naval Medical Center, Bethesda, MD, and a Hologic Selenia at...
Sibley Memorial Hospital, Washington, DC. The schematic in Fig. 2 describes the experimental setup for both systems.

The GE Senographe DS FFDM system has two focal spots, 0.1 and 0.3 mm nominal, indirect CsI scintillator based digital detector, FOV of 19 × 23 cm (1920 × 2304 pixel), pixel size of 100 μm, kVp range 22–49, mAs range 4–500, and source to image distance (SID) of 66 cm. A 31 line pair/cm moving linear grid is built into the Bucky that covers the detector. The Bucky can be removed when required by our experiments. The system has two imaging modes. The fine view mode, uses proprietary software to filter the “for-processing” images (images before processing for display). The standard mode only performs the basic corrections for flat-field, dark-field bad pixels and gain. For simplicity, the fine view mode will be referred to as imaging mode A while the standard mode will be referred to as imaging mode B.

The Hologic Selenia FFDM system has two focal spots, 0.1 and 0.3 mm nominal, selenium based direct digital detector, FOV of 24 × 29 cm (3428 × 4096 pixel), pixel size of 70 μm, kVp range from 20 to 39, mAs range from 3 to 400, the LORAD HTC antiscatter grid (the HTC grid has a crosshatch design that reduces scatter in two directions) and SID of 66 cm. For our experiments, we used the “Phantom” mode which is typically used for QA purposes.

For both systems, we performed all of our experiments with the large focal spot, molybdenum target, molybdenum filter, 30 kVp, and three tube outputs: 20, 100, and 200 mAs. The uniform phantom we used consists of four polymethyl methacrylate (PMMA) plates (10 mm thick each), a 3 mm thick PMMA cover, and an aluminum sheet (0.5 mm thick) placed midway between the four PMMA plates. The phantom assembly matches the size and thickness of the CDMAM,26 as its attenuation for the typical mammography energy range. Both phantoms have 5 cm equivalent thickness of PMMA. According to Dance et al.,27 this PMMA thickness is equivalent to a 6 cm thick breast with 20% glandularity. We measured and compared the Al half value layer (HVL) for this phantom assembly and for the CDMAM phantom on the GE Senographe DS system set at 100 mAs. The CDMAM phantom HVL was 0.676 mm, while the uniform phantom HVL was 0.673 mm.

For both systems, we measured detector entrance exposures by placing the phantom as close as possible to the x-ray tube and placing an ionization chamber between the phantom and the detector surface, as shown in Fig. 2, to avoid backscatter. We fit a linear relation between tube output (mAs) and the measured exposures (mR), after correcting for the distance from the detector using the inverse square law. This way, we obtained the scatter free detector entrance exposures for a given tube output. Note that when the phantom is placed near the detector the detector entrance exposure will be higher due to scatter from the phantom. The detector entrance exposures are used to compare our evaluation results to those from the literature.

We estimated the mean glandular breast dose (MGD) by implementing the method described in Refs. 28 and 27, assuming a 6 cm thick standard breast with 20% glandularity. For the three tube outputs we considered, the MGDs are 0.53, 2.71, and 5.42 mGy for the GE Senographe DS and 0.49, 2.87, and 4.97 mGy for the Hologic Selenia system.

2.B. System characterization

We calculated the GMTF profiles along the x and y directions, where the x direction (parallel to the chest wall), and the y direction (perpendicular to the chest wall) following Ref. 10. The phrase “Generalized” is used to identify that the LRF was estimated with the uniform phantom in place and that it is different from the traditional detector MTF.

For both systems, an edge test object (2.5 × 2.5 cm) was placed midway between the four PMMA plates and 1.5 cm away from the chest side, where the incident x-ray beam is almost perpendicular to the detector surface. A 0.14 mm thick copper plate was chosen as the edge test object because it does not attenuate the beam completely. To reduce noise, an average of five images of the phantom assembly were taken at each exposure, corrected for the Heel effect. Following the procedure described in Refs. 23 and 29, the edge response function (ERF) and the LRF were estimated using the corrected average image while the presampled GMTF was
calculated from the LRF. For the GE Senographe DS system, the pixel size at the detector plane was estimated to be 100 μm, the magnification factor from the center of the phantom to the detector was 1.03, hence the pixel size at the object plane was 97 μm, and hence the Nyquist frequency was 5.15 mm⁻¹. For the Hologic Selenia system, the pixel size was 68 μm at the object plane and the Nyquist frequency was 7.35 mm⁻¹. The 2D GMTF was obtained by fitting cubic spline surface between the profiles along the two axes in a similar manner to Ref. 29. It should be noted that the results can be trusted only along the axis directions, and the interpolation between measured points provides an estimate of the 2D GMTF. The GMTF variance at each frequency was estimated from the multiple GMTFs obtained from each image of the edge. We estimated scatter fraction by the first age of the edge. We estimated scatter fraction by the first estimated from the multiple GMTFs obtained from each image-space method, Gagne et al. obtained both signal present and absent images using the CDMAM phantom. They generated the difference signal ∆S by subtracting sample averages of the two. They then estimated the covariance matrix which represents system noise and calculated the Hotelling observer SNR_{ISIB19} (note that the symbol SNR_{ISIB19} identifies the image-space, image-based SNR with 19 × 19 pixel ROIs using Gagne et al. method) defined as SNR_{ISIB19} = ∆SK⁻¹∆S. To implement this method, we acquired five images at each exposure with the CDMAM phantom placed in the middle of the four uniform PMMA plates (between the second and the third PMMA slabs replacing the 0.5 mm Al sheet of the uniform phantom) in order to obtain signal present ROIs. Five images of the uniform background with the aluminum sheet replacing the CDMAM phantom were acquired to obtain the signal absent ROIs (5048 ROIs in total). To generate ∆S for each disk size, we averaged five 19 × 19 pixel signal present ROIs and then subtracted the average signal absent ROI. All 5048 19 × 19 signal absent ROIs were used to estimate the covariance matrix for each exposure. In order to investigate differences between Gagne’s image-space, image-based method with Fourier-space, image-based methods, we calculated SNR_{FSIB19} and SNR_{FSIB256} by dividing the square of the Fourier transformation of 19 × 19 and 256 × 256 pixel ∆S, by the appropriate 2D GNNPS. Because of its larger size the 256 × 256 GNNPS was calculated with only 80 ROIs, while the 19 × 19 GNNPS was calculated with 5048 ROIs. The parameters of the aforementioned observer models are summarized in Table I for comparison.

2.C. Contrast-detail analysis

For the contrast detail analysis, we first estimate the detection probability of a four alternative forced choice (4AFC) detection task that is based on the visual detection task performed by human readers of CDMAM images. In that detection test readers choose the correct signal location out of four possibilities within each cell of the CDMAM phantom. The SKE/BKE Hotelling observer SNR and the 4AFC imaging (1)

\[
\text{SNR}^2 = \Delta x \Delta y \sum_{k=1}^{256} \sum_{l=1}^{256} \frac{\text{GMTF}^2(f_x, f_y) \text{GNNPS}(f_x, f_y) \Delta S^2(f_x, f_y)}{\Delta S^2(f_x, f_y)},
\]

where \(\Delta S_f(f_x, f_y)\) is the Fourier transformation of an object of interest, here a set of gold disks with thicknesses from 0.03 to 2.0 μm and diameters from 0.06 to 2.0 mm as input signals to match the targets in the CDMAM phantom. In the simulation, each disk signal was specified by signal size and thickness, taking into account the energy-spectrum-dependent linear attenuation coefficient following the procedure of Ref. 15. The standard deviation of the Hotelling observer SNR was determined from the propagation of errors of the GMTF and the GNNPS.

In order to compare the SMFHO SNR with other published methods, we implemented three additional observer models on the GE Senographe DS system. In their image-based, image-space method, Gagne et al. obtained both signal present and absent images using the CDMAM phantom. They estimated scatter fraction generated with the commonly used beam-stop method for the GE Senographe DS system. Note that the scatter fraction generated with the commonly used beam-change in slope of the GMTF curve and compared it with "strapping" with replacement was used to randomly sample ROIs using Gagne et al. method, Gagne et al. obtained both signal present and absent images using the CDMAM phantom. They generated the difference signal ∆S by subtracting sample averages of the two. They then estimated the covariance matrix which represents system noise and calculated the Hotelling observer SNR_{ISIB19} (note that the symbol SNR_{ISIB19} identifies the image-space, image-based SNR with 19 × 19 pixel ROIs using Gagne et al. method) defined as SNR_{ISIB19} = ∆SK⁻¹∆S. To implement this method, we acquired five images at each exposure with the CDMAM phantom placed in the middle of the four uniform PMMA plates (between the second and the third PMMA slabs replacing the 0.5 mm Al sheet of the uniform phantom) in order to obtain signal present ROIs. Five images of the uniform background with the aluminum sheet replacing the CDMAM phantom were acquired to obtain the signal absent ROIs (5048 ROIs in total). To generate ∆S for each disk size, we averaged five 19 × 19 pixel signal present ROIs and then subtracted the average signal absent ROI. All 5048 19 × 19 signal absent ROIs were used to estimate the covariance matrix for each exposure. In order to investigate differences between Gagne’s image-space, image-based method with Fourier-space, image-based methods, we calculated SNR_{FSIB19} and SNR_{FSIB256} by dividing the square of the Fourier transformation of 19 × 19 and 256 × 256 pixel ∆S, by the appropriate 2D GNNPS. Because of its larger size the 256 × 256 GNNPS was calculated with only 80 ROIs, while the 19 × 19 GNNPS was calculated with 5048 ROIs. The parameters of the aforementioned observer models are summarized in Table I for comparison.

| Observer model   | ROI number | ROI size       | Domain    | ∆S        |
|------------------|------------|----------------|-----------|-----------|
| SMFHO SNR        | 80         | 256 × 256      | Fourier   | Simulated |
| SNR_{FSIB19} (Ref. 21) | 5048      | 19 × 19        | Spatial   | Image based |
| SNR_{FSIB256}    | 80         | 256 × 256      | Fourier   | Image based |
| SNR_{ISIB19}     | 5048       | 19 × 19        | Fourier   | Image based |
task detection probability $p$ are connected with an analytical relation.\textsuperscript{32}

The CD curves were generated by determining the disk thickness-diameter pairs that result in a fixed detection probability. The detection probability of 62.5\% was chosen as a threshold because it is the midway between chance (25\%) and 100\% detection (typical for 4AFC tests). In practice, we converted the threshold detection probability to a threshold SNR $\mu = 1.24$.\textsuperscript{32} We simulated a series of gold disks with thicknesses from 0.03 to 2.0 $\mu$m and diameters from 0.06 to 2.0 mm as input signals to match the targets in the CDMAM phantom. For each disk diameter, we compared the 16 SNRs (corresponding to 16 thicknesses) to the threshold SNR. The thicker one of the two disks that produced a SNR closest to the threshold SNR was chosen as the threshold thickness. For example, if the threshold SNR is equal to 0.5, the 1.42 $\mu$m disk has a SNR of 0.3 and the 2 $\mu$m disk has a SNR of 0.51, then 2 $\mu$m will be chosen as a threshold detectable thickness. CD curves were obtained by linear interpolation between all 16 threshold thicknesses.

To link the SMFHO CD curves to human reader performance, we estimated the human efficiency, defined as the human observer SNR $\text{SNR}_{\text{Human}}^2$ divided by the Hotelling observer SNR $\text{SNR}_{\text{SMFHO}}^2$. We obtained the estimated human performance by implementing an empirical method developed by Young \textit{et al.}\textsuperscript{33, 34} Young’s empirical model links the CD curves generated by the CDCOM software (CDCOM (Ref. 26) version 1.4.2) with the CDMAM phantom version 3.4 (Capintec, Inc., NJ) for scoring CDMAM images. It works by detecting the two disk signals within each cell (one in the center and the other one at either one of the four corners). The program compares the detected location to the true location of the disks and returns the detection probability] and those by human readers from three different imaging centers. We implemented Young’s method on the GE Senographe DS system with imaging mode B for the three exposure levels (20, 100, and 200 mAs). For each exposure level, we acquired eight images of the CDMAM phantom and used CDCOM software output and Young’s model to obtain the estimated human CD curves. From the CD curves, we obtained the estimated human observer SNR and compared it to the Hotelling observer SNR. The human efficiency was calculated by averaging over the three exposures.

3. RESULTS

Figure 3 shows the comparison between the detector MTF (without phantom) and the GMTF (with the phantom in place) with and without the grid of the GE Senographe DS FFDM system (mode B) measured at 100 mAs along the x-direction (parallel to the chest wall). The detector MTF was slightly higher when the grid was not installed. The low frequency drop observed when the grid was in place was due to scatter within the grid that degraded the detector MTF. The grid, however, significantly improved the GMTF when the phantom was in place because, as evidenced by the low frequency drop, it reduced the scatter fraction from 38.6\% to 18.4\%.

Figure 4 shows the 1D GMTF derived from the LRF measured along the x-direction, for the GE Senographe DS system [Fig. 4(a)], and the Hologic Selenia [Fig. 4(b)] for different MGDs. In Fig. 4(a), solid lines with filled symbols represent imaging mode A while dashed lines with empty symbols represent mode B. The symbols are shown to help identify the curves and they do not correspond to specific points. In mode B, the GMTF did not significantly change with MGD. In mode A, the GMTFs were all higher than those in mode B for all MGDs, possibly due to image sharpening employed by mode A. Furthermore, the GMTF increased with higher MGD, indicating that a higher degree of sharpening was employed when the dose was increased. In Fig. 4(b), for the Hologic Selenia system, the GMTFs at three MGDs were identical within error bars. The scatter fraction of the GE Senographe DS estimated from the GMTF was 18.4\% ± 0.4\% was very similar to the one estimated with the beam stop method 17.8\% ± 0.2\%. For the Hologic Selenia, the scatter fraction was estimated to be 9.8\% ± 0.4\%.

Figure 5 shows the 1D GNNPS for the GE Senographe DS system [Fig. 5(a)], and the Hologic Selenia system [Fig. 5(b)]. For both systems, the GNNPS decreases with

![Fig. 3. Detector MTF (without phantom) and GMTF (with phantom in place) measurements for the GE Senographe DS system at 100 mAs tube output with imaging mode B. Results are shown with and without the grid installed on the system.](image-url)
FIG. 5. Profiles of GNNPS along the x direction (parallel to the chest wall), for the two systems. Three MGDs are denoted by different colors/symbols. The symbols that correspond to every tenth point are shown to help identify the curves. Error bars are twice the standard deviation. (a) Profiles of GNNPS for the GE Senographe DS system along the x direction. Solid lines with filled symbols represent imaging mode A while dashed lines with empty symbols represent mode B. (b) Profiles of GNNPS for the Hologic Selenia system along the x direction.

In order to compare the SNR estimated using the SMFHO method with other available methods, we determined the SNR of a disk signal (0.63 mm diameter and 1 μm thick) using the three methods described in Sec. 2.B. Figure 6 shows the comparison between (a) SMFHO SNR, (b) SNRISIB19, (c) SNRFSIB256, and (d) SNRFSIB19 for image mode B of the GE Senographe DS FFDM system for 2.71 mGy MGD. The SNR values are similar between methods; further discussion and explanation on the small differences observed will be given in Sec. 4.

We performed CD analysis as a comprehensive way to evaluate the performance of the two systems. The SMFHO CD curves shown in Fig. 7 were referenced to 1 mGy MGD to allow comparisons of the two systems. This was possible because SMFHO SNR is linearly proportional to MGD since the phantom is uniform and the MGD is high enough that generates detector signal significantly higher than the electronic noise base. Note that because the CDMAM disk thicknesses are discrete, small fluctuations in the threshold SNR can cause the CD curve to jump to the next available disk thickness, accentuating the differences between the two systems. The CD curves of GE mode A and mode B are essentially identical. The percent differences in the threshold disk thickness between GE and Hologic range between 0% and 60% well within the estimated uncertainty of the two curves.

Figure 8 shows the intermediate CD curves used to estimate the human observer efficiency and the human efficiency adjusted SMFHO CD curves: The CD curves of the GE Senographe DS system acquired at 239.4 μGy detector entrance exposure using the SMFHO SNR, while the SMFHO CD curves adjusted by 30% human efficiency. We obtained the estimated human efficiency as described in Sec. 2.C for each disk signal and all exposures and calculated the mean human efficiency 30% ± 5%. The CD curves of the CDMAM were estimated with the CDCOM software, while the predicted human CD curve from the CDMAM/CDCOM curve was estimated using Young’s model. Two human observer CD curves (for the GE Senographe 2000D system at 70 and 140 μGy detector exposures, cited from Ref. 35) are shown for comparison. Note that error bars are not shown to improve the readability of the plots.
Fig. 9. CD curves of the GE Senographe DS system acquired at 44.6, 239.4, and 468.3 μGy detector entrance exposure using the SMFHO SNR adjusted by 30% human efficiency. Two threshold thicknesses for disk diameter 0.1 and 0.25 mm at 54.2 μGy, cited from Ref. 24. Another two human observer CD curves (for the GE Senographe 2000D system at 70 and 140 μGy detector exposures, cited from Ref. 35) are shown for comparison.

shown for comparison. Note that error bars are not shown to improve the readability of the plots. For this and the following plots, we used units of detector entrance exposure in order to compare our results with published literature.

Figure 9 shows the comparison between CD curves acquired with different methods. We show the SMFHO CD curves adjusted by the mean estimated human efficiency for the GE Senographe DS system, mode A, at 44.6, 239.4, and 468.3 μGy detector entrance exposure. The adjusted SMFHO CD curve at 44.6 μGy detector entrance exposure shows good agreement to that obtained by Monnin24 at a similar entrance exposure (54.2 μGy) for the same FFDM system for two disk diameters (0.1 and 0.25 mm). In the same figure, we also show two human observer CD curves obtained at 70 and 140 μGy detector entrance exposures for the GE Senographe 2000D FFDM system (a slightly older FFDM model) by Rivetti35 for comparison. Even though the human CD analysis was performed on an older version of the GE FFDM system, the results are close to the adjusted SMFHO CD curve (within error bars).

4. DISCUSSION

The SMFHO method uses an empirical model of system noise and signal transfer in a uniform breast phantom to calculate the Hotelling observer SNR for clinically relevant FFDM imaging tasks. The SMFHO SNR was estimated for different disk signals with different system settings allowing for a practical and comprehensive evaluation of two clinical FFDM systems, including detector and scatter rejection performance, focal spot unsharpness, and patient MGD. To simulate signals, we have accounted for the linear attenuation coefficient of signals,36 x-ray spectrum, signal size, shape and thickness of the background phantom. Furthermore, using a previously published method,33,34 we were able to determine the SMFHO observer efficiency compared to an average human observer, thereby allowing us to compare our results to human observer performance.

In order to make the SMFHO methodology practical to evaluate FFDM systems in clinical settings, we made the following assumptions: the imaging systems were locally linear (without nonlinear or adaptive image processing) shift invariant and wide-sense cyclostationary and the scatter was uniform. It is widely known that the shift invariance and cyclostationarity assumptions do not hold in general. The x-ray angle of incidence affects the detector point spread function.37–39 the projected focal spot shape varies with the location, and the intensity of small signals varies depending on their location within a pixel,22,28 while scatter is typically larger in the center of the image as compared to the edges. In order to ensure that our assumptions did not affect our results significantly, we took the following measures: We chose a relatively small area (about 6 × 6 cm²) in the center of a uniform phantom for the evaluation of the noise.40 The images we used were flat-field corrected and detrended with a second order polynomial ensuring that the pixel intensity and gain were uniform across our small ROI. The LRF was determined from an edge section 2 cm long, translating to less than 1° x-ray incidence angle span eliminating any angle of incidence effects on our estimates of the MTF. The ROI used for the disk signal SNR estimation was much larger than the largest disk diameter, while the smallest disk we simulated was one pixel in diameter. The simulated signals were always positioned in the center of the ROI and the center of the pixel and were subsampled about ten times smaller than the pixel size.

An additional element that determines the clinical applicability of an imaging system evaluation methodology is the total number of images required to evaluate the system under typical use conditions. The reason is that in many clinical centers mammography x-ray tube replacements are linked to the total number of exposures. Having this constraint in mind we determined that five images for each parameter we evaluated were adequate to estimate the GNNPS that was used to determine the SNR. By plotting the SNR as a function of the number of ROIs, for both imaging systems we evaluated, we determined that after 80 ROIs (five images) the SNR reaches its steady state for all the exposures and gold disk sizes we used. Similarly, we wanted to minimize the number of images used for the collection of images to generate the CD curves using the CDCOM software and we verified that the minimum number of images (eight) recommended by the CDCOM software manual was sufficient to reach a steady state.

We compared the SMFHO with three other observer models (see Table I) using different ROI sizes, different ROI numbers, in the Fourier and spatial domains in Fig. 6. The small differences observed in the mean SNRs were well within error bars. The difference between $\text{SNR}_{\text{ISIB19}}$ and $\text{SNR}_{\text{FSIB19}}$ is mainly due to the fact that the two models were calculated in different domains (spatial vs frequency). $\text{SNR}_{\text{ISIB19}}$ is the Fischer linear discriminant, which is designed to be optimal among all linear observers. $\text{SNR}_{\text{ISIB19}}$ therefore must be greater than or equal to $\text{SNR}_{\text{FSIB19}}$.

We noticed that for the specific signals we evaluated the ROI size ($256 \times 256$ vs $19 \times 19$ pixels) does not affect the SNR significantly as can be seen when comparing $\text{SNR}_{\text{FSIB256}}$ to $\text{SNR}_{\text{FSIB19}}$.

One of the greatest advantages of the SMFHO method is that it provides a means to evaluate the scatter removal method of the FFDM system. The SMFHO method uses the
to evaluate the complete system response (including the detector, phantom, focal spot, and grid) because it accounts for realistic scatter magnitude and blur as well as focal spot unsharpness typically present in clinical images. For the GE Senographe DS system, we estimated the detector MTF with and without a grid (without using a phantom) and the GMTF with and without a grid (with the uniform phantom on the detector). The detector MTF is slightly worse with the grid (see Fig. 3), while the GMTF is much higher with the grid, as the grid removes a significant amount of scatter (scatter fraction reduction by 20%). The grid itself contributes slightly to the detector blur, which explains the small low frequency drop of the detector MTF with the grid in place. However, comparing the GMTF with and without the grid we see that the grid contribution to the blur is insignificant compared to the benefit gained when the phantom is imaged.

For the GE Senographe DS, we have evaluated the performance of two different image acquisition modes (i.e., raw images before processed for display): the fine view (imaging mode A) and the standard (imaging mode B). In mode B other than the standard flat-field and bad pixel corrections, no other preprocessing is done by the system before images are saved, while for mode A images are edge enhanced with a linear filter. Due to the edge enhancement we observed that the GMTF is higher than that of mode B, and that it is also dependent on the MGD, indicating the edge enhancement is stronger at higher doses. Consequently, the edge enhancement contributed to higher noise, as evidenced in the NNPS. Note however that the CD curves, generated using the linear Hotelling observer (SMFHO), for the two modes are effectively identical, indicating that the observer can effectively undo the effects of linear edge enhancement. Any effects, that this type of edge enhancement might have on a human reader, have not been evaluated by this study, however it appears that no information is lost by such preprocessing.

As we illustrated in Fig. 8, using Young’s model, we were able to make estimates of an average human reader performance in order to compare the CD curves we estimated with published literature. In Fig. 9, the SMFHO CD curves (adjusted by the estimated human efficiency) were compared to Rivetti’s CD analysis. Although Rivetti’s study was performed on an older version of the GE FFDM system, the results are still within the error bars of the SMFHO CD curve at 44.6 μGy detector entrance exposure. The two data points available from Monnin’s study on the same system model are even closer to the human performance we estimated using the SMFHO and Young’s model for human efficiency.

We summarized the advantages of using the SMFHO method to evaluate the performance of clinical FFDM systems. Comparing to the pixel SNR, the SMFHO method provides an objective assessment of the system performance based on a SKE/BKE detection task. It uses an uniform phantom when estimating the system noise and deterministic properties to take into account the scatter from the phantom and focal spot unsharpness. The SMFHO method can be used to evaluate the performance of the antiscatter grid in clinical FFDM. It provides equivalent evaluation results with Gagne’s image-based method, by using an empirical model of the system (using the MTF and a simulated signal) to predict system performance.

5. CONCLUSIONS

A clinically practical SMFHO assessment methodology for evaluating FFDM system performance has been presented. This method uses an empirical model of the FFDM imaging system, incorporating a uniform phantom assembly to account for scatter blur and focal spot unsharpness. For simulated disk signals inspired from the CDMAM phantom and a SKE/BKE detection task, SMFHO CD curves were generated, and adjusted for an estimated human reader efficiency using Young’s model. We used this method to compare the performance of the GE Senographe DS and Hologic Selenia FFDM systems, and demonstrated that there is no significant difference between them. The SMFHO SNR compares well with other observer models and the resulting CD curves adjusted for the estimated human efficiency are consistent with published literature.

---

1. J. M. Boone, B. A. Arnold, and J. A. Seibert, “Characterization of the point spread function and modulation transfer function of scattered radiation using a digital imaging system,” Med. Phys. 13(2), 254–256 (1986).
2. N. Cooper III, J. M. Boone, J. Anthony Seibert, and C. J. Pellet-Baratik, “An edge spread technique for measurement of the scatter-to-primary ratio in mammography,” Med. Phys. 27(5), 845–853 (2000).
3. K. Doi and K. Rossman, “The effect of radiographic magnification on blood vessel imaging with various screen-film systems,” Med. Phys. 1(5), 257–261 (1974).
4. C. C. Shaw, X. Liu, M. Lennacks, J. X. Rong, and G. J. Whitman, “Optimization of MTF and DQE in magnification radiography: A theoretical analysis,” Proc. SPIE 3977, 466–475 (2000).
5. S. Kyprianou, “A method for total x-ray imaging system evaluation application to a microangiographic detector for neurovascular procedures,” Ph.D. thesis (University of New York at Buffalo, NY, 2004).
6. S. Kyprianou, S. Rudin, D. R. Bednarek, and K. R. Hoffmann, “Generalizing the MTF and DQE to include x-ray scatter and focal spot unsharpness: Application to a new microangiographic system,” Med. Phys. 32(2), 613–626 (2005).
11 R. D. Fiete, H. H. Barrett, W. E. Smith, and K. J. Myers, “Hotelling trace criterion and its correlation with human-observer performance,” J. Opt. Soc. Am. A 4(5), 945–953 (1987).

12 R. M. Gagne, B. D. Gallas, and K. J. Myers, "Toward objective and quantitative evaluation of imaging systems using images of phantoms," Med. Phys. 33(1), 83–95 (2006).

13 I. S. Kyprianou, A. Badano, S. Park, H. Liu, and K. J. Myers, “Noise and signal detection in digital x-ray detectors using the spatial definition of SNR,” Proc. SPIE 7258, 72581D (2009).

14 H. Liu, I. S. Kyprianou, A. Badano, K. J. Myers, R. J. Jennings, S. Park, R. V. Kaczmarek, and K. Chakrabarti, “SKE/BKE task-based methodology for calculating Hotelling observer SNR in mammography,” Proc. SPIE 7258, 72581D (2009).

15 P. Monnin, N. W. Marshall, H. Bosmans, F. O. Bochud, and F. R. Verdun, “Image quality assessment in digital mammography: Part II. NPWE as a validated alternative for contrast detail analysis,” Phys. Med. Biol. 56, 4221–4238 (2011).

16 J. E. Gray and J. A. Princehorn, “HTC grids improve mammography contrast,” White Paper W-BI-HTC(9/04), 2004.

17 R. Visser and N. Karssemeijer, “Manual CDCOM version 1.5: Software for automated readout of CDMAM 3.4 images,” 2007.

18 D. R. Dance, C. L. Skinner, K. C. Young, J. R. Beckett, and C. J. Kotre, “Additional factors for the estimation of mean glandular breast dose using the UK mammography dosimetry protocol,” Phys. Med. Biol. 45, 3225–3240 (2000).

19 D. R. Dance, “Monte Carlo calculation of conversion factors for the estimation of mean glandular breast dose,” Phys. Med. Biol. 35(9), 1211–1219 (1990).

20 I. S. Kyprianou, A. Badano, B. D. Gallas, and K. J. Myers, “Singualr value description of a digital radiographic detector: Theory and measurements,” Med. Phys. 35(10), 4744–4756 (2008).

21 J. A. Brezovich and G. T. Barnes, “A new type of grid,” Med. Phys. 4(5), 451–453 (1977).

22 International Standard – Medical electrical equipment: Characteristics of digital x-ray imaging devices,” IEC 62220-1 (International Electrotechnical Commission, Geneva, Switzerland, 2003), p. 18.

23 A. E. Burgess, “Comparison of receiver operating characteristic and forced choice observer performance measurement methods,” Med. Phys. 22(5), 643–655 (1995).

24 K. C. Young, A. Alsager, J. M. Oduko, H. Bosmans, B. Verbrugge, T. Geertse, and R. V. Engen, “Evaluation of software for reading images of the CDMAM test object to assess digital mammography systems,” Proc. SPIE 6913, 69131C (2008).

25 K. C. Young, J. J. H. Cook, J. M. Oduko, and H. Bosmans, “Comparison of software and human observers in reading images of the CDMAM test object to assess digital mammography systems,” Proc. SPIE 6142, 614206 (2006).

26 S. Rivetti, N. Lanconelli, R. Campanini, M. Bertolini, G. Borasi, A. Nitrosi, C. Danielli, L. Angelini, and S. Maggi, “Comparison of different commercial FFDM units by means of physical characterization and contrast-detail analysis,” Med. Phys. 33(11), 4198–4209 (2006).

27 M. J. Tapiovaara and R. Wagner, “SNR and DQE analysis of broad spectrum X-ray imaging,” Phys. Med. Biol. 30(6), 519–529 (1985).

28 A. Badano, I. S. Kyprianou, M. Freed, R. J. Jennings, and J. Sempau, “Effect of oblique x-ray incidence in flat-panel computed tomography of the breast,” IEEE Trans. Med. Imaging 28(4), 696–709 (2009).

29 A. Badano, I. S. Kyprianou, R. J. Jennings, and J. Sempau, “Anisotropic imaging performance in breast tomosynthesis,” Med. Phys. 34(11), 4076–4091 (2007).

30 A. Badano, I. S. Kyprianou, and J. Sempau, “Anisotropic imaging performance in indirect x-ray imaging detectors,” Med. Phys. 33(8), 2698–2713 (2006).

31 A. E. Burgess, “The Rose model, revisited,” J. Opt. Soc. Am. 16(3), 633–646 (1999).

32 H. H. Barrett and K. J. Myers, Foundations of Image Science (Wiley InterScience, New Jersey, 2004).