Journal of Digital Imaging

A Perceptual Evaluation of JPEG 2000 Image Compression for Digital Mammography: Contrast-Detail Characteristics

Sankararaman Suryanarayanan, Andrew Karellas Ph.D., Srinivasan Vedantham, Sandra M. Waldrop, and Carl J. D'Orsi

In this investigation the effect of JPEG 2000 compression on the contrast-detail (CD) characteristics of digital mammography images was studied using an alternative forced choice (AFC) technique. Images of a contrast-detail phantom, acquired using a clinical full-field digital mammography system, were compressed using a commercially available software product (JPEG 2000). Data compression was achieved at ratios of 1:1, 10:1, 20:1, and 30:1 and the images were reviewed by seven observers on a high-resolution display. Psychophysical detection characteristics were first computed by fitting perception data using a maximum-likelihood technique from which CD curves were derived at 50%, 62.5%, and 75% threshold levels. Statistical analysis indicated no significant difference in the perception of mean disk thickness up to 20:1 compression except for disk diameter of 1 mm. All other compression combinations exhibited significant degradation in CD characteristics.

KEY WORDS: Digital mammography, JPEG 2000, image compression, contrast-detail, breast cancer

DIGITAL MAMMOGRAPHY HAS EVOLVED into a viable and clinically applied technique for breast cancer detection.[1,2] Technological developments in digital mammography are the direct result of recent advances in solid-state flat-panel detector technology and computational methods.[2] Large-area high-resolution X-ray detectors are likely to generate critical data storage and management needs for clinical image data. Furthermore, the evolution of limited-angle tomography[3,4] and computed tomography for digital mammography[5,6] would demand enormous storage and data transmission needs, making data compression a critical factor to the successful clinical implementation of these burgeoning technologies. In addition, the growth in teleradiology and picture archiving and communications systems (PACS) is likely to create a need for data in a form that is suitable for transmission across networks. In recent years, a variety of compression techniques has been investigated for medical image compression to enable compact data storage and transmission without significant loss in diagnostic information content.[7-10] Recently, the JPEG 2000 compression standard was introduced and is actively being pursued as a viable technique for medical image compression.[11-14] This compression standard utilizes the discrete wavelet transform (DWT)[15] and differs from the discrete cosine transform (DCT)-based JPEG compression technique. The DWT technique used in JPEG 2000 is a full-frame transform that can be applied to an entire image. The standards set for JPEG 2000 provides a number of favorable features such as lossy to lossless compression within a single bit stream, improved compression efficiency, multiple-image

From the Department of Radiology, Emory University School of Medicine, 1364 Clifton Rd. NE, Atlanta, GA, 30322, USA.

Certain aspects of this research were presented as a poster at the SCAR 2003 conference in Boston, MA, and was awarded the second place in the "poster awards" category by the SCAR committee.

Correspondence to: Andrew Karellas Ph.D., tel.: (404)-712-2411/5809; fax: (404)-712-5813; e-mail: akarell@emory.edu

Online publication 18 February 2004
doi: 10.1007/s10278-003-1728-x

resolution, progressive decoding, region of interest coding, and error tolerance.[9,12] The wavelet transform in JPEG 2000 decomposes the image data based on scale or resolution. By using recursive low-pass and high-pass digital filter techniques, images with lower resolution (sub-bands) are generated.[9,12] Compression can be achieved by suitably quantizing and coding the sub-bands as permitted by the JPEG 2000 standard. The flexible ordering of bit streams in JPEG 2000 enables multiresolution decoding, file size selection, signal-to-noise ratio, progressive decoding, and visual weighting of sub-band coefficients.[9,12] From a visual standpoint, the overlapping basis functions in JPEG 2000 tend to produce smooth artifacts, unlike the blocklike artifacts produced by JPEG.[12] However, the impact of compression on diagnostic information content is still an active area of research.

In physical evaluations of imaging systems, observer-independent metrics, such as the modulation transfer function (MTF) and detective quantum efficiency (DQE), are commonly used. Although such metrics describe the physical characteristics, they do not provide a description of the visual or perceptual information content of the images. On the other hand, imaging characteristics have been analyzed by investigators using contrast-detail (CD) methodologies[16-19] that provide useful insights into the image quality aspects of an imaging system. Contrast-detail performance is also a widely used quality control tool to assess clinical imaging systems. An important factor that motivates CD analysis is that it encompasses the observer or the "end user" as part of the imaging chain, which is critical if an imaging system is either used or is intended for clinical imaging.[19]

In this investigation the impact of JPEG 2000 image compression in mammography was studied using a phantom-based methodology. The CD characteristics of uncompressed and JPEG 2000 compressed images were used as a metric to compare the impact of image compression. This study was designed based on an alternative forced choice (AFC) paradigm with human observers evaluating pre-and postcompression images. Psychophysical detection data from the observers were fitted to a mathematical model from which the CD characteristics were obtained.

MATERIALS AND METHODS

Digital Mammography System Description

A clinical full-field digital mammography system (Senographe 2000D, GE Medical Systems, Milwaukee, WI) was used to acquire the images for this study. The system comprises a columnar cesium iodide (CsI) scintillator coupled to an amorphous silicon (a-Si:H) photodiode array with a pixel pitch of 100 μm providing a field of view of approximately 19 $\times$ 23 cm. The system is capable of operating in three different autoexposure modes that can be selected based on the preference for either lower dose or higher contrast. The digital mammography system stores images in two formats, raw and processed.

Contrast-Detail Phantom

A commercially available contrast-detail phantom (CDMAM, Nuclear Associates, NY) was used as the test object in this study. The phantom consists of a thin aluminum base that contains circular gold disks that are logarithmically sized from 0.10 to 3.2 mm in diameter and from 0.05 to 1.6 μm in thicknesss.[20] A later version of this phantom has been described in Veldkamp et al.[20] The disks are arranged in a matrix of squares such that within each square one disk is centrally placed and an additional disk is randomly placed at one of the four corners. Within each square, both the central and corner disks have the same diameter and thickness. However, along a row of squares, the disk thickness is constant while logarithmically varying in diameter, and along a column, the diameter remains constant while the thickness varies logarithmically. For the purpose of this study, additional acrylic was added to bring the total thickness of the phantom to 4.5 cm. The main advantage of this phantom is the presence of the randomly spaced corner disk in each square that facilitates alternative forced choice experiments.

Image Acquisition, Compression, and Observer Study

Images of the CDMAM phantom were acquired in the "contrast-auto" mode for better image contrast. A total of 10 images were acquired and the processed images were selected for this study as radiologists view processed images clinically. The images were compressed using commercially available JPEG 2000 image compression software (JPEG 2000, Aware Inc., Bedford, MA) at compression ratios of 1:1, 10:1, 20:1, and 30:1 such that there were 10 images at each compression level thereby yielding a total 40 images. The portion of the phantom that contained columns corresponding to disk diameters 1.0, 0.8, and 0.63 mm was extracted from each image and displayed to the observers (Fig 1).

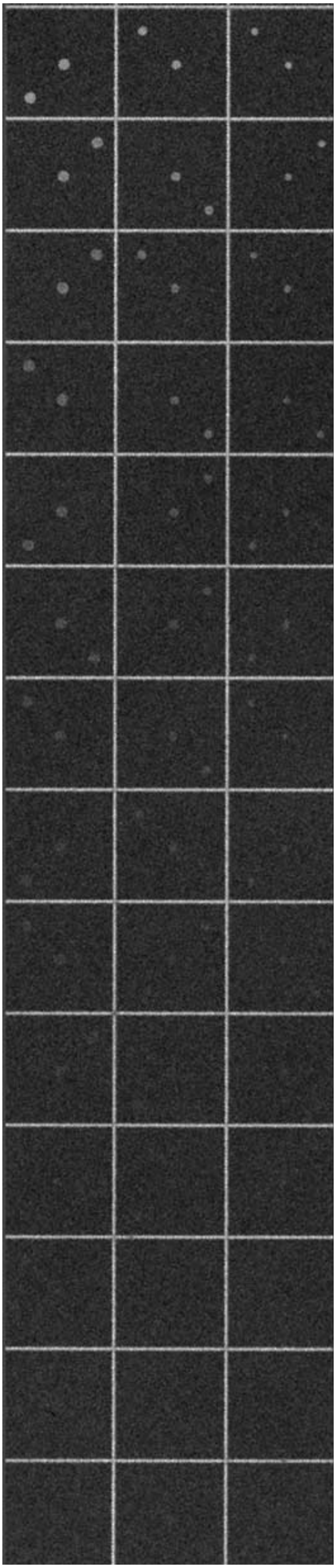

Fig 1. Digitally cropped portion of the CDMAM phantom that was used in this study. Acquired images of the phantom were rotated approximately 45° in order to vertically align the columns with respect to the observer.

An image display program was developed using Interactive Data Language (IDL 5.5, Research Systems Inc, Boulder, CO) that displayed the images to the observers in random order. Prior to the actual observation session, observers were trained and familiarized with the process. The images were displayed on a high-resolution gray-scale monitor (M21PCF1RE, Clinton Medical, Loves Park, IL) in a darkened room. Observers were not allowed to adjust the window or level of the images throughout the study. A total of seven observers participated in this study and each observer independently reviewed the images. Marking sheets (without the actual disk locations) that resembled Figure 1 were provided and the observers were required to identify the location on the corner disk in each square. Since this was a forced-choice study, observers were asked to estimate the probable location of disks in instances where the disks were not perceivable.

Data Analysis

Based on observer responses for each disk diameter, "percentage correct" or "proportion correct" values were computed for various disk thicknesses for each observer and fitted using a maximum-likelihood technique. To analyze the "proportion correct" detection data, we used a signal detection model that hypothesizes a continuous decision variable internal to the observer with Gaussian probability density functions for the presence or absence of the disk.[21-23] The distance between the means of these two overlapping distributions can be represented as $d' = u\Delta C$, where u is the slope parameter that needs to be determined and ΔC is the disk contrast (perceived disk thickness). As described by Ohara et al.,[23] one can then relate the probability of correct choice, $p\,(d')$, to the slope parameter u as [Eq. (A15) of ref. 23]

$$p(d') = p(u\Delta C) = \int_{t=-\infty}^{\infty} [\phi(t)]^{M-1} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-(u\Delta C - t)^2}{2}\right) dt, \quad (1)$$

where $\phi(t)$ is the cumulative Gaussian distribution and the slope parameter u is estimated using a maximum-likelihood algorithm developed using MATLAB (Version 6, The MathWorks Inc., Natick, MA).[21-23] A comprehensive description of the maximum-likelihood technique that was implemented in this investigation has been provided by Ohara et al.[23] In this study we had K trials defined as $K = NL$, where N was the number of repetitions at L disk thickness levels. We used the derivation of Aufrichtig[21] to compute the variance of u for all K trials as

$$\sigma_u^2 = \frac{1}{\sum_{k=1}^{NL} 1/\sigma_{u_\kappa}^2} = \left(\sum_{k=1}^{NL} \frac{\left(\frac{dp_k}{\cdot} dd_k\right) \Delta C_k^2}{p_k(1-p_k)} \right)^{-1}.$$

From the estimated mean value and variance of u, Eq. (1) can be used to generate percent correct detection curves, $p(d')$ substituting the computed value of u, the value of M (4 in our case), and the contrast or disk thickness levels.

Contrast-Detail Characteristics

The CD characteristics were obtained at three different detection threshold levels: 50%, 62.5%, and 75%. For each observer and diameter, the disk thickness corresponding to the desired threshold point in the detection curve, for example, 50%, was noted (Fig 2).[17] We used perceived disk thickness as a measure of contrast, ΔC in this study. In cases where a corresponding value was not available at the desired threshold level, linear interpolation between adjacent data values was performed to estimate the corresponding disk thickness. Contrast-detail characteristics were derived for all diameters, observers, compression ratios, and threshold levels. Finally, for each compression ratio, the corresponding CD curves from individual observers were averaged to obtain an average CD characteristic at a given compression ratio.

Statistical Analyses

An analysis of variance of disk thickness was done with a means model using SAS Proc Mixed software (version 8, SAS Inc., Cary, NC). Compression ratio, threshold, and disk diameter were fixed effects, and all possible interaction terms were included in the statistical model. A compound symmetry variance–covariance form in observer measurements was assumed for disk thickness and robust estimates of the standard errors of parameters were used to do statistical tests and construct 95% confidence intervals. Statistical comparisons of mean disk thickness were limited to the six pairwise comparisons between compression ratios at each level of threshold and disk diameter. A Bonferroni adjustment ($P < 0.0083$) was used for the six pairwise comparisons. The reported P values are two-sided.

RESULTS

The "proportion correct" detection characteristics for two different observers for a 0.63-mm-diameter disk at 1:1 and 30:1 compression ratios are shown in Figures 2a and b, respectively. The maximum-likelihood model-generated detection curves fit well to the experimental data. Overall, for the detection task that was studied, 84 detection curves were obtained for all disk diameters, compression ratios, and observers (3 diameters × 4 compression ratios × 7 observers). The mixed effects linear model used to analyze disk thickness provided the flexibility of modeling the means and the variances. Statistically significant differences for all fixed

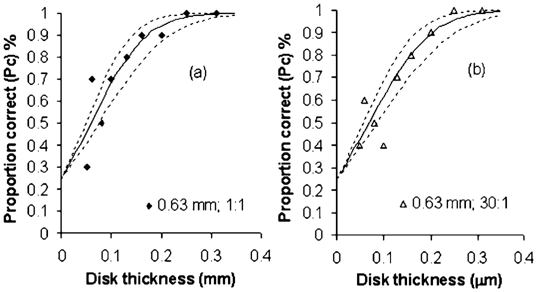

Fig 2. Proportion correct detection characteristics for a 0.63-mm-diameter disk for two different observers at (a) 1:1 and (b) 30:1 compression. The smooth bold line indicates the model fitted response and the dotted lines represent error bands. The markers are raw experimental data points based for these observers. The fine dotted horizontal and vertical lines indicate perceived disk thickness corresponding to a desired threshold level.

effects and interaction terms were identified ($P < 0.0001$ for each factor and each interaction term). Since the three-way interaction effect for compression ratio, threshold, and disk diameter was significant, statistical comparisons focused on comparing mean thickness between compression ratios at each level of threshold and disk diameter (nine levels). The estimate of within-observer variance for disk thickness was approximately equal to the between-observer variance suggesting that the variability between repeated measurements by the same observer and the variability between measurements from different observers were similar.

The mean disk thickness computed from the responses of seven observers at each of three threshold levels and each of three disk diameters for the four compression ratios is shown in Figure 3. The vertical bars are the 95% confidence interval for each estimate of mean disk thickness. Generally, the sizes of the mean disk thickness differences between the four compression ratios at each threshold and at each disk diameter appear small and most of the 95% confidence intervals overlap. However, a few of the statistical tests indicated consistent differences among disk thickness at a given diameter across threshold levels. For example, statistically significant differences in mean disk thickness at a disk diameter of 1.0 mm at each threshold level were seen when comparing compression ratio 1:1 to 10:1 ($P < 0.0001$), 1:1 to 20:1 ($P < 0.0001$), and 1:1 to 30:1 ($P \leq 0.0009$). The

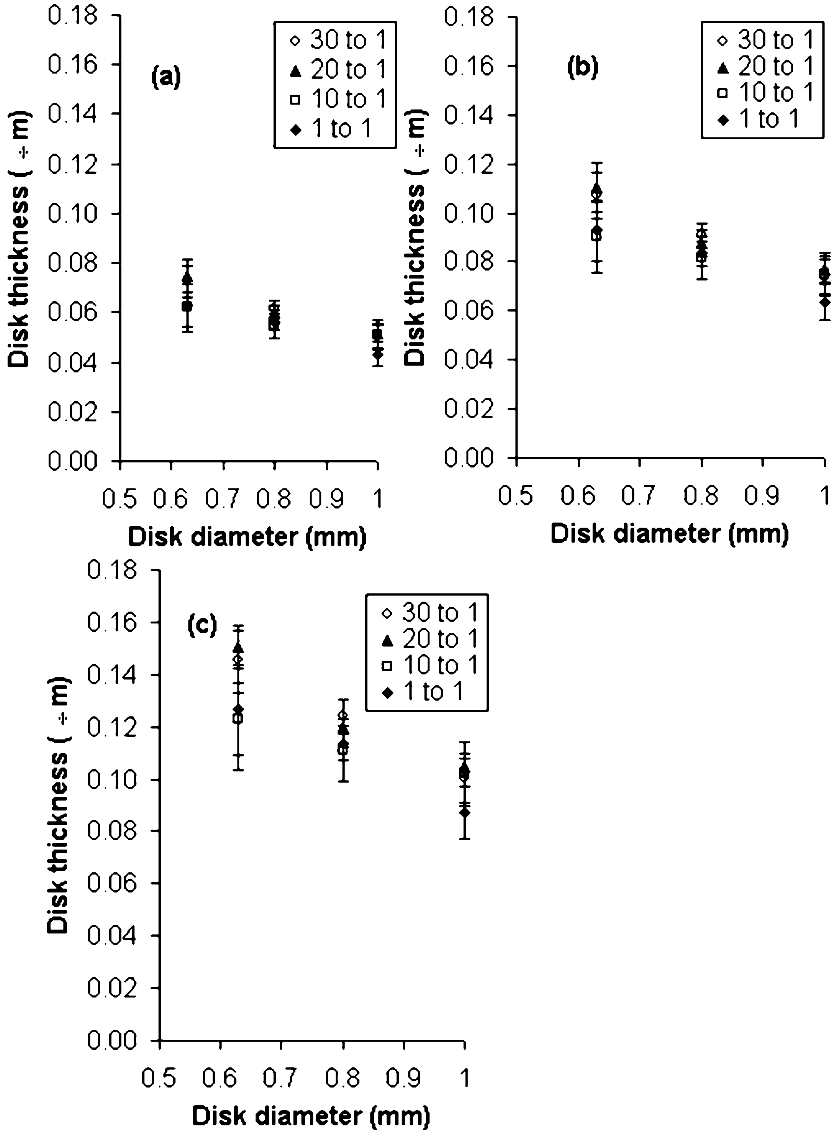

Fig 3. Contrast-detail (CD) characteristics obtained by averaging individual CD characteristics of each of the seven observers at (a) 50%, (b) 62.5, and (c) 75% threshold levels for different compression ratios. The error bars represent 95% confidence intervals.

mean disk thickness at a disk diameter of 1.0 mm was smaller for compression ratio 1:1 compared with 10:1, 20:1, and 30:1. Differences in mean disk thickness were also seen at disk diameter 0.8 mm at each threshold level when comparing disk ratio 10:1 to 30:1 ($P \leq 0.0002$) and at disk diameter 0.63 mm at each threshold level when comparing 1:1 to 30:1 ($P = 0.001$). No statistical differences in the CD characteristics were observed among 1:1, 10:1, and 20:1 compression ratios for diameters 0.8 and 0.63 mm.

DISCUSSION

Psychophysical characterization of imaging systems provides information on the image quality and diagnostic value of a modality. The CD characteristics essentially summarize the information carrying capacity of an imaging system. Variants of the maximum likelihood technique have been used by other investigators[24] in developing perception models. The methodology described in this study provides a means to quantitatively assess perception data and characterize imaging performance. Other automated methods[7,25] can also be used in combination with human observers to analyze system performance.

Based on the results of this study, it appears that JPEG 2000 image compression up to 20:1 does not have a significant impact on the CD characteristics obtained from the test object used in this study for the 0.8- and 0.63-mm-diameter disks. For the conditions investigated in this study, it appears that low-contrast objects with small size (diameter) are inherently difficult to perceive that compression up to a certain level does not alter their perceptibility. However, the difference in the perception of the 1-mm disk at different ratios could be due to the easier perception of the 1-mm disk under no compression conditions relative to disks of smaller diameters. Furthermore, the degree of model fit to the 1-mm detection data may have impacted the CD characteristics. The relative impact of each of the aforementioned effects needs to be investigated further. The significant differences between disk thickness for a specific disk diameter across multiple threshold levels are expected as moving to higher threshold levels because of the underlying nature of the psychometric detection curve. The properties of the image and the compression parameters play a major role in the quality of image compression.[12] Optimization of such parameters can be done using automated techniques[7] or numerical observer models.[17,26] The "task-specific" optimization of image compression parameters was beyond the scope of this study, but it is likely that such optimization techniques could potentially improve the performance of image compression techniques from a perceptual standpoint. Overall, the results of this study suggest that compression ratios up to 20:1 can be applied to images of the type investigated in this study without a significant loss in detection of small targets with the assumption that the difference in the 1-mm disk requires further investigation. However, a more comprehensive analysis is required to optimize compression parameters for task-specific applications. This study demonstrates a CD technique that can be applied effectively to analyze image compression algorithms from a perceptual standpoint.

CONCLUSION

New compression techniques such as JPEG 2000 potentially provide an efficient means for data storage, transport, and management without compromising the diagnostic value of image information within reasonable compression levels. Optimization of compression algorithms with respect to the nature of the image data being compressed is likely to preserve the image quality characteristics at higher compression ratios.

ACKNOWLEDGMENT

This study was supported in part by National Institutes of Health (NIH) grant R01-CA88792 and RO1-EB002123 from the National Cancer Institute (NCI) and National Institute of Biomedical Imaging and Bioengineering (NIBIB), respectively. The contents are solely the responsibility of the authors and do not necessarily represent the official views of the NCI, NIBIB, or the NIH. The authors thank Mr. Mike Dubose and Jeff Stehouwer, Ph.D, for participating as observers in this study. The authors also thank Mr. Kirk A. Easley for providing assistance with statistical analysis.

REFERENCES

1. Lewin JM, Hendrick ER, D'Orsil CJ, et al: Comparison of full-field digital mammography with screen-film mammography for cancer detection: results of 4,945 paired examinations. Radiology 218:873-880, 2001

2. Vedantham S, Karellas A, Suryanarayanan S, et al: Breast imaging using an amorphous silicon-based full-field digital mammography system: stability of a clinical prototype. J Digit Imaging 13:191-199, 2000

3. Niklason LT, Christian BT, Niklason LE, et al: Digital tomosynthesis in breast imaging. Radiology 205:399-406, 1997

4. Suryanarayanan S, Karellas A, Vedantham S, et al: Comparison of tomosynthesis methods used with digital mammography. Acad Radiol 7:1085-1097, 2000

5. Boone JM, Nelson TR, Lindfors K, et al: Dedicated breast CT: radiation dose and image quality evaluation. Radiology 221:657-667, 2001

6. Chen B, Ning R: Cone-beam volume CT breast imaging: feasibility study. Med Phys 29:755-770, 2002

7. Chan H, Lo SB, Niklason LT, et al: Image compression in digital mammography: effects on computerized detection of subtle microcalcifications. Med Phys 23:1325-1336, 1996

8. Egashira K, Nakata H, Watanabe H, et al: Clinical evaluation of irreversible data compression for computed radiography of the chest. J Digit Imaging 11:176-181, 1998

9. Foos DH, Muka E, Slone RM, et al: JPEG 2000 compression for medical imagery. Medical Imaging 2000: PACS Design and Evaluation. Eng Clin Issues 3980:85-96, 2000

10. Good WF, Sumkin JH, Ganott M, et al: Detection of masses and clustered microcalcifications on data compressed mammograms: an observer performance study. AJR Am J Roentgerol 175:1573-1576, 2000

11. Information available at http://www.jpeg.org/JPEG2000.html, official homepage of JPEG 2000

12. Jones PW, Rabbani M: JPEG compression in medical imaging. In: Kim Y, Horii SC eds Handbook of Medical Imaging, vol 3: Displays and PACS. SPIE Press, Washington, DC, 2002, pp 221-275

13. Savcenko V, Erickson BJ, Persons KR, et al: An evaluation of JPEG and JPEG 2000 irreversible compression algorithms applied to neurologic computed tomography and magnetic resonance images. Joint Photographic Experts Group. J Digit Imaging 13(2 Suppl 1):183-185, 2000

14. Sung MM, Kim HJ, Yoo SK, et al: Clinical evaluation of compression ratios using JPEG2000 on computed radiography chest images. J Digit Imaging 15:78-83, 2002

15. Harpen MD: An introduction to wavelet theory and application for the radiological physicist. Med Phys 25:1985-1993, 1998

16. Liu H, Fajardo LL, Barrett JR, et al: Contrast-detail detectability analysis: comparison of a digital spot mammography system and an analog screen-film mammography system. Acad Radiol 4:197-203, 1997

17. Suryanarayanan S, Karellas A, Vedantham S, et al: Flat-panel digital mammography system contrast-detail comparison between screen-film radiographs and hard-copy/images. Radiology 225:801-807, 2002

18. Wang J, Anderson J, Lane T, et al: Contrast-detail characteristic evaluations of several display devices. J Digit Imaging 13(2 Suppl l):162-167, 2000

19. Wang J, Langer S: A brief review of human perception factors in digital displays for picture archiving and communications systems. J Digit Imaging 10:158-168, 1997

20. Veldkamp WJH, Thijssen MAO, Karssemeijer N: The value of scatter removal by a grid in full field digital mammography. Med Phys 30:1712-1718, 2003

21. Aufrichtig R: Comparison of low contrast detectability between a digital amorphous silicon and a screen-film based imaging system for thoracic radiography. Med Phys 26:1349-1358, 1999

22. Aufrichtig R, Xue P: Dose efficiency and low-contrast detectability of an amorphous silicon x-ray detector for digital radiography. Phys Med Biol 45:2653-2669, 2000

23. Ohara K, Doi K, Metz CE, et al: Investigation of basic imaging properties in digital radiography. 13. Effect of simple structured noise on the detectability of simulated stenotic lesions. Med Phys 16:14-21, 1989

24. Chakraborty DP: Maximum likelihood analysis of free-response receiver operating characteristic (FROC) data. Med Phys 16:561-568, 1989

25. Chakraborty DP, Eckert MP: Quantitative versus subjective evaluation of mammography accreditation phantom images. Med Phys 22:133-143, 1995

26. Eckstein MP, Abbey CK, Bartoff JL: Model observer optimization of JPEG image compression. SPIE Int Soc Opt Eng Image Percept Performance 3981:106-115, 2000