

Use of a Human Visual System Model to Predict Observer Performance with CRT vs LCD Display of Images

Elizabeth A. Krupinski PhD,¹ Jeffrey Johnson PhD,² Hans Roehrig PhD,¹
John Nafziger PhD,³ Jiahua Fan MS,¹ and Jeffery Lubin PhD³

This Project evaluated a human visual system model (JNDmetrix) based on just noticeable difference (JND) and frequency-channel vision-modeling principles to assess whether a Cathode ray tube (CRT) or a liquid crystal display (LCD) monochrome display monitor would yield better observer performance in radiographic interpretation. Key physical characteristics, such as veiling glare and modulation transfer function (MTF) of the CRT and LCD were measured. Regions of interest from mammographic images with masses of different contrast levels were shown once on each display to six radiologists using a counterbalanced presentation order. The images were analyzed using the JNDmetrix model. Performance as measured by receiver operating characteristic (ROC) analysis was significantly better overall on the LCD display ($P = 0.0120$). The JNDmetrix model predicted the result ($P = 0.0046$) and correlation between human and computer observers was high (r^2 (quadratic) = 0.997). The results suggest that observer performance with LCD displays is superior to CRT viewing, at least for on-axis viewing.

KEY WORDS: Vision model, observer performance, digital display

ALTHOUGH PACS (picture archiving and communication systems) and teleradiology have changed the practice of radiology quite dramatically over the past number of years, there are still many changes taking place in the digital reading room.¹⁻³ One aspect that seems to evolve continuously is the display.^{4,5} Although early investigations tended to focus on how many monitors were required in the digital reading environment and what resolution was required,⁵ a critical issue today is whether flat-panel liquid crystal displays (LCDs) are more suitable for primary diag-

nostic reading than the more traditional CRT (cathode ray tube) display.^{6,7} In many respects, the LCD has performance characteristics that exceed those of the CRT.^{8,9} The MTF (modulation transfer function) of an LCD is essentially isotropic, whereas the MTF of a CRT is non-isotropic. At the highest digitally addressable spatial frequency along the CRT scanline in the horizontal direction, the value is usually only 10% to 20%. In the vertical direction the MTF at twice the line frequency is closer to 30% to 40%.^{8,9} The result is that, for the CRT, more than half the contrast modulation is lost at the highest spatial frequencies, but this degradation does not occur with the LCD. Practically speaking, this could mean degradation of high-frequency lesions in radiographic images such as microcalcifications in mammograms, resulting in reduced detection performance.

Veiling glare, defined as the diffuse spreading of light or scattering within various parts of a display device,^{9,10} is also significantly different for the LCD compared to the CRT. In the CRT monitor, the thick faceplate of the vacuum bulb, back-scattering of electrons and light leakage

¹From the Department of Radiology, University of Arizona, 1609 N. Warren Ave., Tucson, Arizona.

²From the Siemens Corporate Research, 775 College Rd. East, Princeton, New Jersey.

³From the Sarnoff Corporation, 201 Washington Rd., Princeton, New Jersey.

Correspondence to: Elizabeth A. Krupinski PhD, tel: 520-626-4498; fax: 520-626-4376; e-mail: krupinski@radiology.arizona.edu

Copyright © 2004 SCAR (Society for Computer Applications in Radiology)

Online publication 24 August 2004

doi: 10.1007/s10278-004-1016-4

through aluminum-layer non-uniformities, all contribute to veiling glare. This diffuse spreading of light results in a significant degradation of the maximum contrast capabilities of the CRT monitor. The decreased contrast capabilities will especially affect the detection of low-contrast objects in images such as masses in mammograms. LCDs (liquid crystal displays), on the other hand, tend to have much less (almost negligible in fact) veiling glare than CRTs because there is no need for the vacuum barrier glass faceplate. The protective layer for the LCD TFTs (thin-film transistors) may result in some veiling glare, but not nearly as much as with the CRT display. The question that arises is whether the relatively low veiling glare and isotropic MTF of an LCD display affects observer performance less than the generally higher veiling glare and non-isotropic MTF of a CRT display. The goal of the present investigation was to compare performance on an LCD to a CRT display using both human and model observers.

We have demonstrated the utility of model observers to predict observer performance in a number of previous experiments aimed at defining the optimal display characteristics for viewing radiographic images in the digital environment.¹¹⁻¹³ The eventual goal of investigating the validity of such models for predicting human observer performance is to eliminate, or at least reduce, the number of receiver operating characteristic (ROC) studies conducted. ROC studies are the gold standard for measuring observer performance in radiology, but they are time-consuming and require a significant amount of effort on the part of both the investigator and observer. If we could use model observers to at least narrow the relevant number parameters to study, we could reduce this experimental burden on observers. Additionally, model observers may help us understand better how the human visual system encodes and processes information during the diagnostic interpretation process.

MATERIALS AND METHODS

Display Characterization

Two display parameters have been identified as being significantly different for LCDs and CRTs—MTF and veiling glare. For our purposes, MTF was measured via the

contrast transfer function (CTF), also known as the square wave response.¹⁴ We measured horizontal and vertical CTFs using rectangular fields filled with square-wave contrast patterns. The raw digital image data represent optical density scaled for 12 bits per pixel. Because display functions of softcopy systems are generally expressed as luminance versus digital output, the digital data were converted to luminance values. We measure veiling glare by displaying black disks of varying diameter in the center of the CRT or LCD against a uniform background of maximum luminance. We then use a spot photometer to measure the luminance in the black disk relative to the luminance in the surround and plot it as a function of the diameter of the black disk. We form the veiling glare ratio, V_G as:

$$V_G = (L_b - L_r)/L_s,$$

where L_b is the luminance at the center of the black disk, L_r is the detector dark level, and L_s is the luminance in the surround. We then plot it as a function of the diameter of the black disk.

JNDmetrix Human Visual System Model

The architecture of the JNDmetrix model (JND = just noticeable difference; Sarnoff Corporation, Princeton, NJ)^{15,16} begins with two paired input images and finishes with a JND map showing the magnitude and spatial location of visible differences between the input images. In the first optics stage, the input images are convolved with a function approximating the point spread by the optics of the eye. Image sampling by the retinal cone mosaic is simulated in the second stage by a Gaussian convolution and point sampling sequence of operations. Next, the raw luminance image is converted to units of local contrast, and then decomposed into a Laplacian pyramid. The result is seven frequency bandpass levels. Each pyramid level is convolved with eight pairs of spatially oriented filters with bandwidths derived from psychophysical data. After oriented filtering, the pairs of filter output images are squared and summed, yielding a phase-independent energy response. This mimics a widely proposed transformation in the mammalian visual cortex from a linear response among simple cells to an energy response among complex cells. The phase independence from this operation has some useful properties; e.g., it makes the model less sensitive to the exact position of edges, a property exhibited in human psychophysical performance as well.

At the next “transducer” stage, the energy measure for each pyramid level is normalized by a value approximating the square of the grating contrast detection threshold for that pyramid level and local luminance; then the normalized energy measure is transformed by a sigmoid non-linearity to reproduce the dipper shape of visual contrast discrimination functions. This takes into account nonlinear masking effects that make features less detectable on a nonuniform background. To account for characteristics of human foveal sensitivity, the model includes a pooling stage that averages transducer outputs over a small neighborhood by convolving with a disc-shaped kernel. After this, the model output for each spatial position is essentially an m -dimensional

vector, where m is the number of pyramid levels times the number of orientations. The distance between the vectors for the two inputs is computed in the "distance metric." A spatial map of JND values results, representing the degree of discriminability between the two images. This JND map can be reduced to a single, aggregate value by Minkowski normalization (Qnorm). The map derived is useful because it shows the magnitude and the position of noticeable differences between the two images. Also, all image differences are rendered in the same units, and thus they are quantitatively comparable.

Images and Observer Study

A series of 40 breast images with malignant masses and 40 images with benign masses were downloaded from the database for Screening Mammography.¹⁷ These base images were used to generate a larger series of images. 512×512 regions of interest around each mass were extracted from the original images. The masses were removed via digital processing to generate signal-absent versions of the images. Lower contrast (75%, 50%, and 25%) versions of the masses were created with weighted superposition of the signal-present and signal-absent image versions. The final test set therefore contained five versions of each of the 80 base images (100% or original contrast, 75%, 50%, 25%, 0% or signal absent), for a total of 400 images. The images were then down sampled to 256×265 for input as region-of-interest images into the visual system model. To capture the effects of exactly what the human observer sees on the monitors, we capture each image with a high-performance CCD camera. The camera enables linear oversampling in which $4 \times 4 = 16$ CCD pixels are sampled for each monitor pixel. This generates high-fidelity images that cannot be differentiated visually from those displayed on the monitor.

The set of 400 images was shown to six radiologists, once on a high performance LCD (Dome Ci5, Planar Systems, Beaverton, OR) and once on a high performance CRT (Siemens SMM 210200P with a P45 phosphor; Siemens Health Services, Erlangen, Germany) monitor using a counterbalanced presentation design. Each presentation used a different image randomization order. A minimum of 2 weeks passed between viewing sessions. Ambient room lights were turned off. The monitors were set to the same maximum luminance (500 cd/m^2 , 0.8 cd/m^2 minimum), were the same resolution (2048×2560), and were both calibrated to the DICOM-14 Gray Scale Display Function Standard. Viewers were seated 27 cm from the display. The images were displayed in the center of each monitor and were viewed on-axis (since the LCD has the potential to change the appearance of images when viewed off-axis). The observers' task was to examine each image and decide if a mass was present or absent. They then rated their confidence in that decision on a 6-point scale. No image processing (e.g., window/level adjustment) was permitted. Each viewing session lasted approximately one hour. The rating data were analyzed using the Multi-Reader Multi-Case Receiver Operating Technique of Dorfman, Berbaum and Metz.¹⁸ Area under the curve (Az) values were generated for

each observer, and the overall differences between conditions were analyzed statistically using analysis of variance (ANOVA) techniques with appropriate post-hoc tests.

RESULTS

The MTF and veiling glare measurements of the CRT and LCD monitors are shown in Figures 1 and 2, respectively. It can be seen that the veiling glare for the LCD monitor is much lower than that for the CRT monitor, and that the MTF is essentially the same in both the horizontal and vertical directions for the LCD but quite disparate for the CRT.

The human observer performance results (ROC Az) are shown in Figure 3 for each of the lesion contrast levels used. Overall, performance with the LCD was significantly higher than with the CRT when evaluated with a t -test for paired observations ($t = 2.727$, $df = 23$, $P = 0.0120$). In every contrast condition, performance was higher with the LCD and reached statistical significance with the 75% and 50% contrast masses ($P < 0.05$). All six readers performed higher at all contrast levels with the LCD than with the CRT. The VDM (Fig 4) predicted the same pattern of results ($t = 2.854$, $df = 638$, $P = 0.0046$) for all lesion contrast levels, with the LCD yielding better discrimination performance than the CRT. The correlation (Fig 5) between human Az and VDM JND results was high (r^2 quadratic = 0.997).

DISCUSSION

The results of this study demonstrate that performance with images displayed on an LCD monitor with lower veiling glare and isotropic MTF is significantly better than on a CRT monitor with higher veiling glare and non-isotropic MTF. The results apply at all lesion contrast levels, although it was only at the 75% and 50% contrast levels that statistical significance was reached for the human observers. The human and model results closely parallel each other and demonstrate essentially the same pattern of results. In terms of the model, this provides us with further evidence that the JNDmetrix model is a very suitable tool for helping to predict observer performance in radiologic image interpretation tasks as a

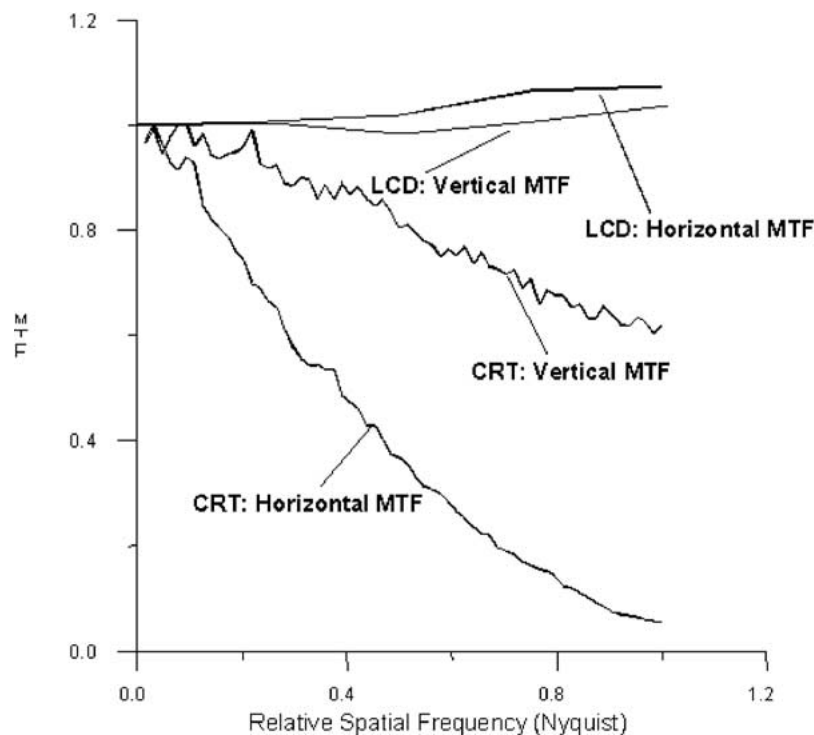


Fig 1. MTF in the vertical and horizontal directions for the cathode ray tube (CRT) and liquid crystal display (LCD) monitors.

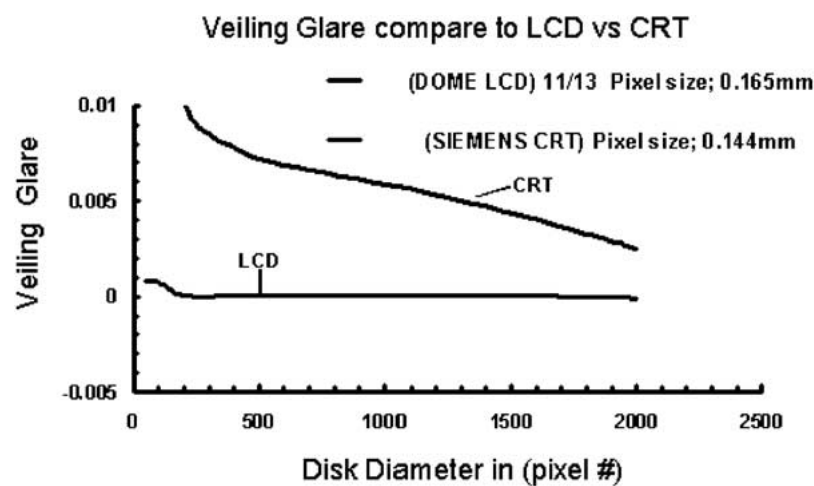


Fig 2. Veiling glare measurements for the CRT and LCD monitors.

function of the manipulation of various viewing parameters.

There are a few caveats that need to be mentioned regarding this study and our plans for future work. The first is that the study used only mammographic images with masses, so it is not possible to generalize completely to other types of images and lesions. It is also important to remember that we used only small regions of

interest from the original mammograms that eliminated the need for the observers to engage in any extensive searching. It is difficult to say whether performance would be better or worse if we had used the full images and the observers had to search for the lesions. It is highly likely, however, that although overall performance would likely decrease (since they would have to search the entire image and detection errors

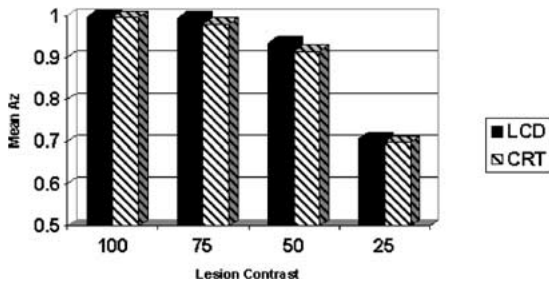


Fig 3. Mean receiver operating characteristic (ROC) Az values for each lesion contrast level for CRT and LCD viewing.

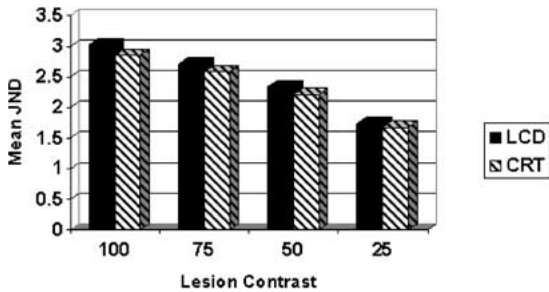


Fig 4. Mean model just noticeable difference (JND) values for each lesion contrast level for CRT and LCD viewing.

would therefore be more likely to occur), the same general pattern of results would be observed.

Another main point is that the present study was not able to determine precisely to what degree MTF and veiling glare independently affected performance. Although we controlled for as many other factors as possible (e.g., luminance, resolution), there are still some intrinsic differences between the CRT and LCD that we cannot control. Therefore, our plans for future work will use only a single monitor (CRT or LCD) and develop image-processing methods to simulate veiling glare changes in the display (i.e., create more and less veiling glare effects compared to the nominal veiling glare of the monitor). In this way we can more precisely observe the effects of veiling glare on performance. We have already examined a method for compensating for MTF degradation in CRT displays and found that compensating for the non-isotropic nature of the CRT MTF does indeed improve observer performance.¹²

It is also important to note that all of the images in this study were displayed at the center

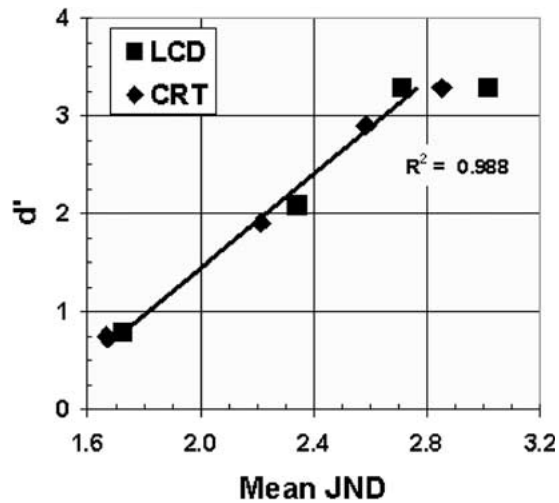


Fig 5. Correlation between human and model observers for each lesion contrast level for CRT and LCD viewing.

of the monitors. Neither the CRT nor the LCD displays are uniform in their characteristics at various locations within the viewing area of the display. For LCDs, this may in fact exacerbate the off-axis viewing problems.^{19,20} Future studies will be done to determine if the viewing angle effects observed differ depending on where the image is displayed on the monitor. Future studies will also use microcalcifications as targets, because they have different physical characteristics and might react differently.

CONCLUSIONS

The present study suggests that LCDs are likely to be suitable for diagnostic viewing of radiographic images in the digital reading environment. Because of their lower veiling glare and isotropic MTF compared to CRT displays, they may provide an even better viewing medium than the more traditional CRT. The effects of off-axis viewing, however, may represent a potential barrier for LCD use in the clinical reading room, but this aspect needs to be further investigated.

Acknowledgment

This work was supported in part by grant R01 EB002119-04 from the National Institutes of Health (NIH/NIBIB).

REFERENCES

1. Huang HK: Some historical remarks on picture archiving and communication systems. *Comput Med Imaging Graph* 27:93-99, 2003
2. Sinha U, Bui A, Taira R, et al: A review of medical imaging informatics. *Ann Med Imaging Inform* 980:168-197, 2002
3. Huang HK: Enterprise PACS and image distribution. *Comput Med Imaging Graph* 27:241-253, 2003
4. Bennett WF, Vaswani KK, Mendiola JA, et al: PACS monitors: an evolution of radiologists' viewing techniques. *J Digit Imaging* 15:171-174, 2002
5. Horii SC: Workstations. In: Mehta A, Dreyer KJ, Thrall JH (eds) *PACS: A Guide to the Digital Revolution*. Springer-Verlag, New York, 2002, pp 191-235
6. Badano A: Principles of cathode-ray tube and liquid crystal display devices. In: Samei E, Flynn MJ (eds) *Advances in Digital Radiography: Categorical Course in Diagnostic Radiology Physics 2003 Syllabus*. RSNA, Oak Brook, IL, 2003, pp 91-102
7. Hwang SA, Seo JB, Choi BK, et al: Liquid-crystal display monitors and cathode-ray tube monitors: a comparison of observer performance in the detection of small solitary pulmonary nodules. *Korean J Radiol* 4:153-156, 2003
8. Blume H, Steven PM, Cobb M, et al: Characterization of high-resolution liquid-crystal displays for medical images. *Proc SPIE Med Imag* 4681:271-292, 2002
9. Roehrig H, Krupinski E, Chawla AS, et al: Spatial noise and threshold contrasts in LCD displays. *Proc SPIE Med Imag* 5034:174-186, 2003
10. Badano A, Flynn MJ: Method for measuring veiling glare in high-performance display devices. *Appl Opt* 39:2059-2066, 2000
11. Krupinski EA, Johnson J, Roehrig H, et al: Using a human visual system model to optimize soft-copy mammography display: influence of display phosphor. *Acad Radiol* 10:161-166, 2003
12. Krupinski EA, Johnson J, Roehrig H, et al: Using a human visual system model to optimize soft-copy mammography display: influence of MTF compensation. *Acad Radiol* 10:1030-1035, 2003
13. Johnson JP, Nafziger J, Krupinski EA, et al: Effects of grayscale window/level parameters on breast lesion detectability. *Proc SPIE Med Imag* 5034:462-473, 2003
14. Roehrig H: The monochrome cathode ray tube and its performance. In: Kim Y, Horii SC (eds) *Handbook of Medical Imaging, Volume 3. Display and PACS*. SPIE Press, Bellingham, WA, 2000, pp 155-220
15. Lubin J: A visual system discrimination model for imaging system design and evaluation. In: Peli E (ed) *Visual Models for Target Detection and Recognition*. World Scientific Publishers, River Edge, NJ, 1995, pp 245-283
16. Lubin J: The use of psychophysical data and models in the analysis of display system performance. In: Watson AB ed *Digital Images and Human Vision*. MIT Press, Cambridge MA, 1993, pp 163-178
17. <http://marathon.csee.usf.edu/Mammography/Data-base.html>. Accessed 7-16-03
18. Dorfman DD, Berbaum KS, Metz CE: Receiver operating characteristic rating analysis: generalization to the population of readers and patients with the jackknife method. *Invest Radiol* 27:723-731, 1992
19. Badano A, Flynn MJ, Martin S, et al: Angular dependence of the luminance and contrast in medical monochrome liquid crystal displays. *Med Phys* 30:2602-2613, 2003
20. Badano A, Gallas BD, Myers KJ, et al: Effect of viewing angle on visual detection in liquid crystal displays. *Proc SPIE Med Imag* 5029:474-483, 2003