2-AFC Observer Study of Shape and Contrast Discrimination in Digital Stereomammography

Andrew D. A. Maidment, Todd Karasick, Predrag R. Bakic, Michael Albert

University of Pennsylvania, Department of Radiology

3400 Spruce Street, Philadelphia, PA 19104

ABSTRACT

We continue to evaluate fundamental factors that affect the ability of human observers in digital stereomammography. A 2-alternative forced choice (2-AFC) observer study for discrimination of simulated objects in the presence of x-ray quantum noise was performed. In our previous contrast-detail and 2-AFC studies investigating the detection of simulated lesions, we observed that at the same total dose observers perform similarly for stereoscopic and monoscopic imaging. The current experiments were designed to investigate discrimination tasks. Three or four observers attended a series of sessions, each consisting of 300-400 image pairs. We sequentially evaluated discrimination of images based on object shape and contrast. In each trial, two images were presented, each containing a small disk of known size and position, but which differ in terms of blurring or contrast to background. The observers indicated the image containing the disk with greater blurring or higher contrast. The experiments were repeated for 3 or 4 different values of signal-to-noise ratio (SNR), and for 3 different diameters. The fraction of correct responses was computed for each test condition. Detection performance was compared in terms of the linear fit of d' as a function of SNR. Preliminary results again confirmed the advantage of stereoscopy. For the discrimination of blurred objects, the ratio of d'(SNR) averaged over all conditions took values in the range of 1.22-1.73 for the three observers (average 1.45), compared to the theoretically expected value of 1.41. No advantage was seen for discrimination of contrast (average 1.02) It appears that suppression of quantum noise in stereoscopically viewed simulated images by the human visual system enables advantages in discrimination of small lesions with different shape. It is possible, therefore, to match the dose of a stereo pair to the dose required for a single mammogram.

Keywords: Digital Mammography, Stereomammography, 2-AFC Observer Study, Observer Performance

1. INTRODUCTION

Digital mammography has long been touted in terms of the increased signal-to-noise ratio (SNR) of the resultant images.[1,2,3,4] Most digital mammography imaging systems today have at least a 2-fold improvement in the detective quantum efficiency (DQE) as compared to state-of-the-art screen-film imaging systems.[2,5,6] Studies based on phantom images (e.g., contrast-detail studies) demonstrate a corresponding increase in object detectability, which is consistent with the increased DQE and hence image SNR.[2,7] Yet clinical trials to date have not shown a significant improvement in specificity, sensitivity, accuracy or area under the ROC curve.[8,9] One reason for this discrepancy can be attributed to non-quantum "noise" sources; the most substantial of which is anatomic noise due to the breast parenchymal pattern. Thus, it would appear that one of the main remaining limitations to the mammographic detection of breast cancer is not the dose used to produce an image, but rather the superposition of anatomic structures in the breast. This superposition can obscure breast lesions leading to false negatives or it can result in false lesions due to the summation of non-adjacent tissue leading to false positives.

To overcome these false negatives and false positives, and hence to improve observer performance, requires that images be acquired in which the conspicuity of the breast parenchyma is reduced. There are a number of options to achieve this goal, including contrast enhanced mammography[10,11,12,13,14] and imaging the breast in 3-dimensions. There are numerous 3D radiographic methods, including digital stereomammography, digital tomosynthesis and computed tomography of the breast.[15,16,17,18,19,20,21] What has been lacking to date is a method of inter-comparing the performance of these methods in terms of observer performance and radiation dose. Such an evaluation either requires repeated exposure of many patients, or careful phantom studies or simulations. We have attempted to compare the performance

of these 3D imaging methods using both phantom and simulation studies.[18,19,20] To date, we have only compared monoscopic and stereoscopic digital radiography and mammography.

Stereomammography is a procedure in which two mammograms are acquired sequentially while the breast is held in constant compression. The two mammograms are acquired with the x-ray tube in slightly different positions. Typically, the x-ray tube is rotated or translated laterally by ±3° from the normal. By viewing the films in a geometry which recreates the acquisition geometry, it is possible to view the structures in the breast in 3D. In film-based stereomammography, the total dose is twice that of monoscopic mammography, as the characteristic curve (H&D curve) of the film dictates the dose. In digital stereomammography, the total dose for stereomammography should be comparable to that of monoscopic digital mammography, because in theory the quantum noise in the two images should be uncorrelated and hence the total SNR should depend on the total dose to the breast, not the number of images.

Initially, we preformed phantom studies using a contrast-detail (C-d) phantom, with acquisition under carefully controlled conditions. This first experiment was designed to provide some initial experience with comparing observer performance when viewing either monoscopic or stereoscopic digital radiographs. To simplify this initial experiment, we imaged the phantom with zero-parallax and on a uniform background. The desire was to see whether the human observer could combine the quantum noise patterns in the 2 projections in a stereoscopic pair, and hence observe the stereoscopic image with an increased SNR as compared to either projection image. If the human observer could ideally sum these image data, a $\sqrt{2}$ improvement in SNR would be expected. We observed a SNR increase of 1.33±0.04. Conversely, one could claim that on the basis of quantum noise alone, the dose for a stereomammography image pair need only be 10% greater than the dose for monoscopic mammography.[18,19]

The C-d phantom used in this initial experiment (RMI-180, Gammex-RMI, Middleton, WI) was imaged on a uniform background. Imaging the phantom on a non-uniform background would have required an order of magnitude more images to perform the experiment. If tomosynthetic images of the phantom were to be produced or computed tomography performed, then an additional one to two orders of magnitude more source data would need to be acquired. It was clear that further experiments in this direction would need to be performed in simulation. This realization led to our next published work,[20] in which the detection of lesions in monoscopic and stereoscopic mammography was compared using a 2-alternative forced-choice (2-AFC) methodology. This 2-AFC detection experiment again considered objects in a uniform background and with zero parallax. Yet, the experiment yielded results which did not completely agree with the previous C-d phantom experiments. We attributed the observed difference to the low SNR of the detected objects, as compared to the C-d experiments, which led to observers using different scoring criteria.

The C-d experiment images consisted of arrays of objects of differing diameter and subject contrast. The observers were trained to inspect the objects from largest to smallest and from greatest to least contrast. The objects were inspected for general roundness ("whether or not more than 50% of the edge was visible"), size ("whether or not more than 50% of the object was missing"), and the expected position in the detail array.[18] The goal was to prevent misinterpretation of the clustered background noise as phantom details. The 2-AFC detection experiment images consisted of two 256x256 regions with Gaussian noise, in which one region had a 10, 14 or 20 pixel radius disk superimposed. The observer needed to determine, to the best of their ability, which region contained the disk. We quickly learned that no benefit was achieved with stereoradiography without the addition of high contrast fiducial markers (similar to Toto Circles).[22] We reasoned that this observation was due to the low contrast of the disks. Observers in the 2-AFC detection experiment reported using overall brightness, and differential brightness between the centers of the 2 regions as the main criteria for detection. As it was not guaranteed that each projection had equally conspicuous disks, the eye would fuse random bright regions in the two projections. However, the addition of fiducial markers caused the eye to fuse these markers and then more reliably fuse the disks. For comparison, the SNR for the limiting objects in the C-d experiments was between 4 and 6, while the SNR for the limiting objects in the 2-AFC detection experiment was between 1 and 3.

As a result of these observations, we reasoned that to reproduce the experimental conditions of the original C-d experiment, and to better simulate mammographic imaging, we would need to perform 2-AFC experiments with objects which were readily distinguishable from the background. This realization has led to the experiments presented in this publication. In this paper, we consider 2-AFC discrimination experiments in which two regions are presented, each with a disk-like object, but were the objects differ either in their edge definition or their contrast to the background.

2. METHODOLOGY

This paper describes a series of 2-AFC discrimination experiments designed to ascertain the ability of the human observer to combine image information presented independently to each eye. The general methodology follows that presented previously.[20] The observer is presented images using custom designed software that allowed rapid and simple observer responses to stimuli. No manipulation of the images was allowed, and observer viewing distance was roughly maintained at 1m. The same computer and display technology was used as before. [20]

The computer display was perceptually-linearized and calibrated, per DICOM Part 14, using a Tektronics J17 photometer.[23] The calibrated display input values were then scaled in terms of the standard deviation (SD) of image noise calculated per pixel. This calibration is presented in Figure 1. Throughout the remainder of the paper, the contrast of objects will be specified in terms of signed SD.
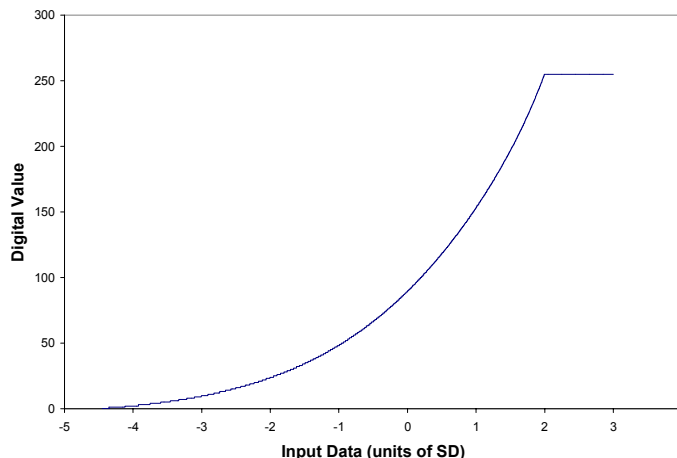
The 2-AFC discrimination experiments were designed to allow the observer to discriminate between two disk-like objects, one in each of two 256x256 regions. Each object was centered in its region. The objects were identical except for one of two properties, either the edge sharpness or the contrast to the background. One such image pair is illustrated in Figure 2 (top). The display software would present an image pair and wait for user input. Left/Right buttons of a trackball were used to select the image containing the disk with greater unsharpness or with greater contrast to the background. The correct response was then shown (Figure 2 – bottom), and a corresponding tone was produced. Then, the screen was blanked momentarily to minimize residual image effects, and the procedure was repeated. The order of the image presentation is discussed below.

As with previous experiments, a uniform background and zero parallax were used. Disks with sizes of 10, 14 or 20 pixels radius were used, again as before. The disks for the edge sharpness or "blurring" experiments were created by convolving each disk with a square blurring kernel, as given in Table 1. The contrast of the disks was determined experimentally in a series of 2-AFC flyer experiments to achieve a probability correct of

Figure 1: Relationship between input data (in units of per pixel standard deviation of the noise), and monitor digital driver level.
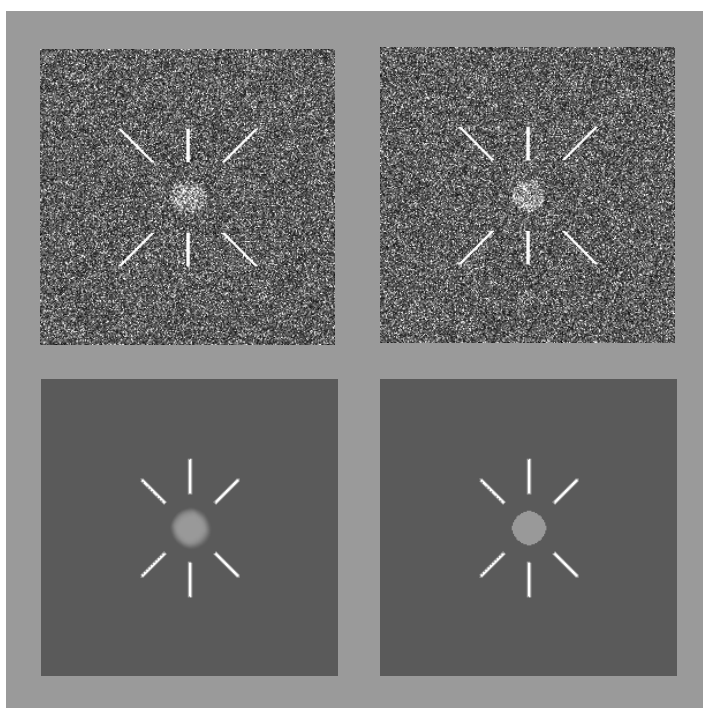

Figure 2: Example of a test condition (radius = 14, blur kernel 7x7) with contrast enhanced to allow review in print. The upper half of the figure is indicative of how the image pairs appear on the screen. The lower half is illustrative of the response feedback screen.

between 0.65 and 0.85 for each blurring kernel and disk radius. The contrasts used in the final observer experiments are given in Table 2. Three levels of probability correct were used, corresponding approximately to low (L=0.65), medium (M=0.75) and high (H=0.85) probabilities of detection.

Radius	Blurring Kernels			
10	3x3	5x5		
14	3x3	5x5	7x7	
20	3x3	5x5	7x7	9x9

Table 1: The blurring kernels used for each sized object.

Radius	P_{corr}	Contrast			
		3x3	5x5	7x7	9x9
10	L	0.60	0.50	-	-
	M	0.80	0.60		
	H	1.00	0.70		
14	L	0.60	0.40	0.30	-
	M	0.70	0.45	0.35	
	H	0.80	0.50	0.40	
20	L	0.45	0.30	0.20	0.16
	M	0.71	0.38	0.25	0.30
	H	0.87	0.54	0.35	0.35

Table 2: The contrast used for presentation of each sized object as a function of the probability correct (L, M, and H P_{corr} values), and each blurring kernel.

The contrast discrimination experiments were performed with two disks of slightly different contrast. One disk was presented with a reference contrast (or pedestal) of 1.0. The second disk was presented with a contrast, $1+\Delta$, that differed by a small amount, Δ, from that of the pedestal. For the three disk sizes, contrasts corresponding to Δ=0.06, 0.09, 0.12 and 0.15 were used. As with the blurring experiments, these values were obtained from initial flyer experiments by a single observer.

3. RESULTS

3.1. Discrimination on the basis of Edge Sharpness

The results presented for the edge sharpness (or blurring) discrimination experiments are for 3 observers. Each observer attended approximately 20 sessions, where each session consisted of approximately 8 sets of 50 observations. A total of 54 test conditions were viewed; namely, stereo/mono; 3 disk diameters; 3 contrasts; and 2, 3 or 4 blur kernels. Thus for each test condition 150 observations were performed per observer, for a total of 8,100 observations per observer. The experiments were set so that stereoscopic and monoscopic images were read sequentially in sets of 50 for each test condition. The order of the test conditions was randomized for each observer.

The results of the 2-AFC experiments on edge sharpness discrimination are shown in Figure 3. Shown are the combined results for all three observers. The data presented consist of d' values calculated from

$$d' = 2\,\mathrm{erf}^{-1}(2P_{Corr} - 1)$$

measured for each test condition. The d' values for the monoscopic and stereoscopic images for each test condition are presented parametrically. A linear least-squares fit of $d'_s = k\, d'_m$ yielded a value of $k=1.45\pm0.06$. The Pearson correlation coefficient of the fit was 0.17. These data are consistent with the assumption that $d'_s = \sqrt{2}\, d'_m$.

The dose requirement for stereoscopy can be calculated using equation 6 from reference (20), where $k=k_S/k_M$, yielding $X_{2S}/X_{1M} = 0.95\pm0.04$. This result would support the conclusion that in the absence of detector noise, the dose required to produce a stereoradiographic image pair is equal to the dose required to produce a monoscopic radiograph with equal detectability of objects.

The results of the observer study did not differ significantly between observers or between the various simulated objects. The value of k for different observers varied between 1.22 and 1.73. The responses of the three observers, as a function of object radius and blurring kernel, are given in Table 3. With the exception of the 3x3 blur kernel for the 20 pixel radius disk, the results do not vary significantly. The 3x3, radius 20 images were universally considered to be "very difficult" by the readers.

The performance of a single observer is shown in Figure 4 for the various disk sizes and blur kernels. In the figure, d' is plotted as a function of the engineering SNR,

$$SNR^2 = \frac{\sum_{i,j}\left(I_{i,j} - I'_{i,j}\right)^2}{\sigma^2_{pix}}$$

of the discrimination tasks, where I and I' are the pixel values in the two source images (with and without blurring), and σ^2_{pix} is the per pixel variance. The points are connected by lines to facilitate reading the graph. Assuming that the data can be evaluated together, we calculated the observer efficiency by fitting the data in aggregate with a linear least-squares fit forced through the origin. The slope is 0.27 ± 0.01, resulting in a very low observer efficiency[24] of 7%. Assessing the same data by size and kernel resulted in slopes of between 0.22 and 0.35. The greatest single outlier was the "20_3x3" data (i.e., radius 20,
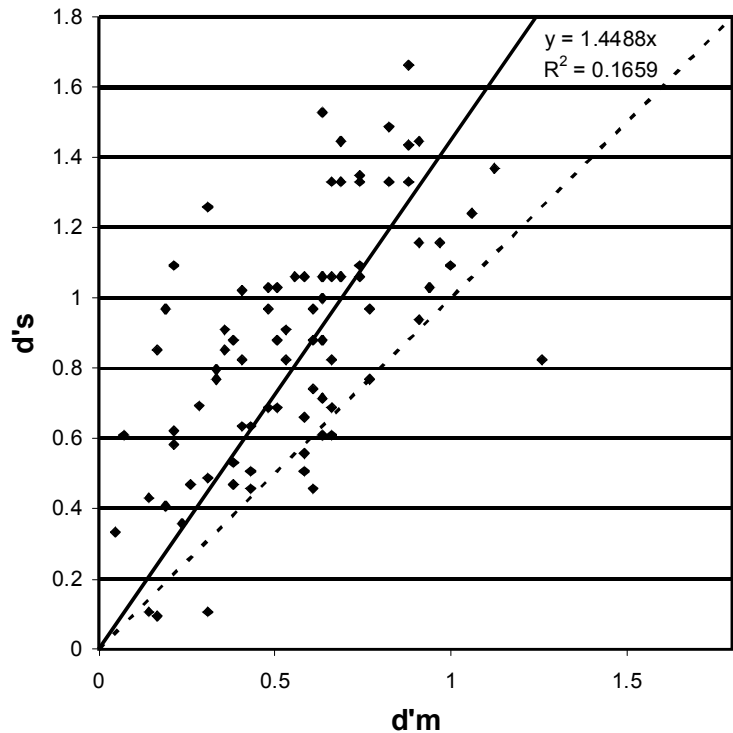

Figure 3: Observer performance plotted parametrically for stereoscopic and monoscopic rendition of images. Data are shown for 3 observers.
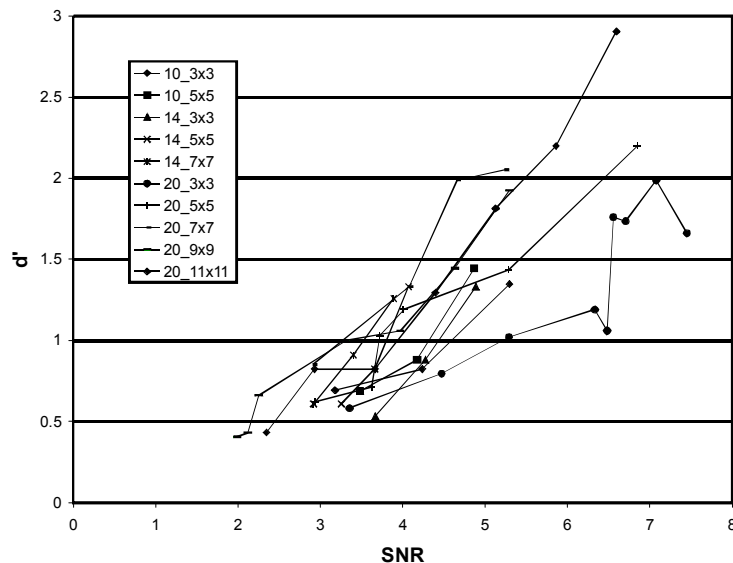

Figure 4: Observer performance for a single observer for different disk radii and blur kernels presented as a function of SNR. Data are combined from flyer experiments and final experiments. The data for specific kernel/radius combinations are connected by lines for convenience.

3x3 kernel), as noted below in Table 3. Examining figure 4, it would appear that the data do not trend towards the origin. An unconstrained linear least-squares fit results in an slope of 0.53±0.10, supporting this assertion.

Radius	⟨d'$_s$⟩/ ⟨d'$_m$⟩			
	3x3	5x5	7x7	9x9
10	1.60	1.55		
14	1.44	1.35	1.49	
20	1.09	1.55	1.45	1.65

Table 3: Observer performance (k) averaged over 4 observers, and presented as a function of the disk size and blur kernel.

3.2. Discrimination on the Basis of Differential Contrast

The results presented here for the contrast discrimination experiments are for 4 observers. Each observer attended 18 sessions, where each session consisted of 8 sets of 50 observations. A total of 24 test conditions were viewed; namely, stereo/mono; 3 disk diameters; and 4 values of Δ. Thus, for each test condition, 300 observations were performed per observer, for a total of 7,200 observations per observer. At each of the 18 sessions, the observer viewed 400 images with a fixed disk diameter. The images were presented in a randomized order with regard to contrast increment (Δ) and mono/stereo rendition. The order of the sessions was randomized for each observer.

The results of the 2-AFC experiments on contrast discrimination are shown in Figure 5. Shown are the combined results for all four observers. The data presented consist of d' values calculated as above, measured for each test condition. The d' values for the monoscopic and stereoscopic images for each test condition are presented parametrically. A linear least-squares fit of $d'_s = k\, d'_m$ yielded a value of k=1.02±0.01. The Pearson correlation coefficient of the fit was 0.75. These data are consistent with the assumption that $d'_s = d'_m$. Thus, it would appear that there was no benefit to viewing stereoradiographic images in this experiment. Possible reasons for this result are discussed in Section 4.
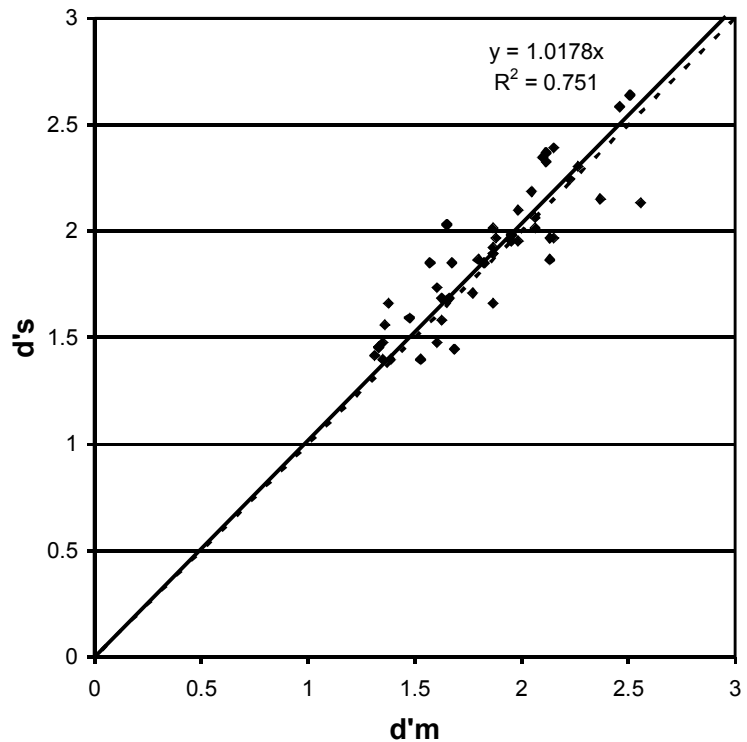

Figure 5: Observer performance plotted parametrically for stereoscopic and monoscopic rendition of images. Data are shown for 4 observers.

The performance of a single observer is shown in Figure 6 for the various disk sizes and values of Δ. In the figure, d' is plotted as a function of the engineering SNR, as with Figure 4. As with Figure 4, the observer efficiency is quite poor. The slope from a linear least-squares fit of d' as a function of SNR is 0.29±0.03, with a Pearson correlation coefficient of 0.92. Thus, the observer efficiency is approximately 8%. There is little dependence upon disk size.

4. DISCUSSION

At first inspection, the results for the edge sharpness and contrast discrimination experiments appear contradictory. To deal with this apparent contradiction, it is necessary to consider the two experiments separately. Let us first consider the edge sharpness discrimination experiment.

The chief result of the edge sharpness discrimination experiment is that k=1.45±0.06. This result is consistent with the assumption that $d'_s = \sqrt{2}\, d'_m$. Conversely, we can conclude that for this observer task the total dose needed for a stereoradiographic image pair is equal to the dose of a single projection radiograph ($X_{2S}/X_{1M} = 0.95\pm0.04$). This result holds in the absence of detector noise, and is based upon the equal detectability of objects.

This result is also consistent with the C-d experiments conducted previously.[18,19] This agreement is likely due, in part, to the similar criteria used in selecting objects in the 2 experiments. The C-d experiments involved the selection of the smallest visible object in an array of such objects, which satisfied shape, size and position requirements. The objects at the threshold of detection had SNR in the range of 5-6. In the case of the edge sharpness 2-AFC experiments, the observers again based their selection upon edge sharpness and area estimation. The lowest SNR for objects relative to the back-ground (as opposed to engineering SNR) was 5.6 (see table 4).
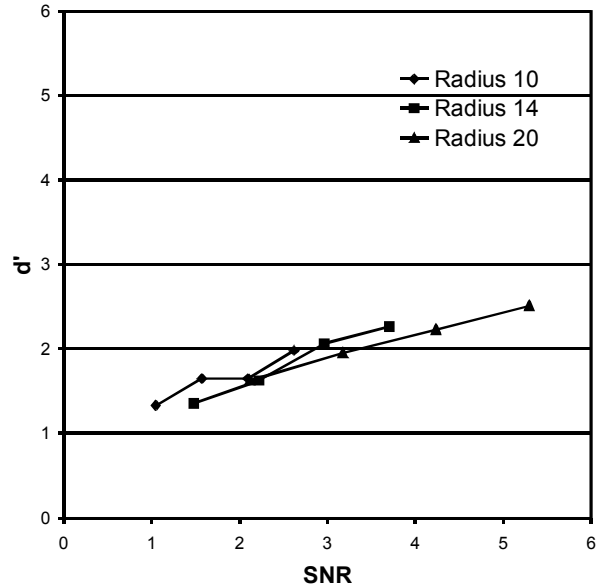

Figure 6: Observer performance for a single observer for different disk radii and incremental contrasts presented as a function of SNR. The data for each radius are connected by lines for convenience.

The contrast discrimination task is quite different. The question being asked of the observer is essentially an estimation task. This is a more challenging task as the eye is not an efficient photometer. A better experimental design might have been to have the two disks in contact (it would be likely that the shape would need to be altered to a square to accommodate this change). However, this experiment is very similar to our previous 2-AFC detection experiments.[20] In the detection experiment, a small contrast difference to the surrounding region was presented. No benefit was seen with the simulation of stereoradiographic images unless high-contrast fiducial marks were added to the image to aid in stereoscopic fusion. Even with the fiducial markers, only a small benefit was observed, corresponding to a value of k≈1.15. The separation of the two disks in the contrast discrimination experiment may have been sufficient to effectively annul any benefit seen stereoscopically.

Radius	P_{corr}	SNR			
		3x3	5x5	7x7	9x9
	L	10.5	8.7		
10	M	14.0	10.5		
	H	17.5	12.2		
	L	14.8	9.9	7.4	
14	M	17.3	11.1	8.6	
	H	19.7	12.3	9.9	
	L	15.9	10.6	7.1	5.6
20	M	25.1	13.4	8.8	10.6
	H	30.7	19.1	12.3	12.3

Table 4: The SNR (relative to the background) of the non-blurred disk for the test conditions of the edge sharpness discrimination experiments.

These results are also supported by two prior experiments. Hsu *et al.*[25] have previously examined the detection of simulated abnormalities in stereomammograms. Their work differs from our own in that it did not consider quantum noise. Two experiments were performed; one was an "arrangement experiment" in which a particular arrangement of objects in 3D needed to be detected. The other was a "density experiment" in which an object with greater density was to be found. Assuming binormal statistics, the area under the ROC curve, A_Z, is equal to the probability correct in a 2-AFC experiment.[24] Thus, it is possible to estimate that $k=1.42$ for the arrangement experiment. Arguably, Hsu's arrangement experiment is comparable to our edge sharpness experiment, as in both cases the observer is more interested in spatial information than in intensity information. In Hsu's density experiment, $k=1.22$. It can be reasonably argued that this experiment is similar to our contrast discrimination experiment, as estimation of contrast to background is of prime importance to the observer.

In a more recent, clinical example, Getty *et al.*[15] measured the improved detection of masses and calcifications in stereomammograms. If we again assume binormal statistics, we can calculate k=1.24 for calcifications and k=1.05 for masses. Calcifications are relatively high contrast objects in which edge-detail is important, as with our edge sharpness discrimination experiments. Masses are relatively large low-contrast objects, and hence contrast to background becomes the limiting issue. This is more like our contrast discrimination task.

While it is unclear how these observations will ultimately affect the performance of digital mammography, it is interesting to speculate. One can reasonably conclude that stereomammography is more likely to improve the detection of microcalcifications, as the stereoscopic rendition allows the observer to see the 3D spatial inter-relationship of the calcific particles, and the detection of calcifications is clearly limited by quantum statistics and edge information. Conversely, it can reasonably be argued that tomosynthesis will better portray masses, as tomosynthesis can be used to reconstruct the attenuation coefficients of the masses,[26] which is a more sensitive measure of density than projection images (either stereoscopic or tomosynthetic). While this might appear disappointing at first, one should remember that the projection source images for tomosynthetic image reconstruction can easily be rendered as a series of stereoscopic pairs. Thus, stereoscopic and tomosynthetic image rendition of digital mammograms may someday be seen as complimentary.

5. CONCLUSION

The latest in a series of experiments on the benefits of stereoradiography have been presented. These experiments consider the tasks of discriminating between two objects on the basis of edge sharpness and contrast. The edge sharpness experiments provide support for the conclusion that the human observer can efficiently average the signal and noise from each projection image in a stereoradiographic image pair and thereby produce a mental image which has a 1.45-fold improvement in the SNR over either individual image. The contrast experiments showed no such benefit; however, hindsight would indicate that the experimental design could have been improved, and that the question being asked may have been ill-posed.

This project is part of an ongoing effort to understand how effectively the human observer is able to detect and diagnose breast cancer in various radiographic imaging modalities. In the next phase of this research, non-uniform backgrounds and non-zero parallax will be added to the simulation. Non-uniform backgrounds will be provided using our anthropomorphic breast model.[27,28,29] Addition of this model will allow more realistic detection tasks and will allow us to consider tomosynthesis for the first time.

REFERENCES

1   A.D.A. Maidment and M.J. Yaffe. Scanned-slot digital mammography. In Medical Imaging IV: Image Formation, Proceedings of the SPIE **1231**:316-326 (1990).

2   A.D.A. Maidment, M.J. Yaffe, D.B. Plewes, G.E. Mawdsley, I.C. Soutar, and B.G. Starkoski. Imaging performance of a prototype scanned-slot digital mammography system. In: Medical Imaging VII: Physics of Medical Imaging, Proceedings of the SPIE **1896**:93-103 (1993).

3   E.D. Pisano, M.J. Yaffe, B.M. Hemminger, R.E. Hendrick, L.T. Niklason, A.D.A. Maidment, C.M. Kimme-Smith, S.A. Feig, E.A. Sickles, and M.P. Braeuning. Current Status of Full-Field Digital Mammography. *Academic Radiology*, **7**:266-280 (2000).

4   Feig SA. Yaffe MJ. Digital mammography. *Radiographics*. **18**:893-901 (1998).

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

6   E. Samei and M.J. Flynn, An experimental comparison of detector performance for direct and indirect digital radiography systems. *Med. Phys.* **30**:608-622 (2003).

7   E.A. Berns, R.E. Hendrick, and G.R. Cutter, Optimization of technique factors for a silicon diode array full-field digital mammography system and comparison to screen-film mammography with matched average glandular dose. *Med. Phys.* **30**:334-340 (2003).

8   E.B. Cole, E.D. Pisano, E.O. Kistner, K.E. Muller, M.E. Brown, S.A. Feig, R.A. Jong, A.D.A. Maidment, M.J. Staiger, C.M. Kuzmiak, R.I Freimanis, N. Lesko, E.L. Rosen, R. Walsh, M. Williford and M.P. Braeuning. Diagnostic Accuracy of Digital Mammography in Patients with Dense Breasts who underwent Problem-Solving Mammography: Effects of Image Processing and Lesion Type. *Radiology*, **226**:153-160 (2003).

9   Lewin JM, D'Orsi CJ, Hendrick RE, et al. Clinical comparison of full-field digital mammography and screen-film mammography for detection of breast cancer. *Am J Roentgenol* **179**:671–677 (2002).

10   M. Skarpathiotakis, M J. Yaffe, *et al.*, Development of contrast digital mammography. *Med. Phys.* **29**:2419-2426 (2002).

11   RA Jong, MJ Yaffe, M Skarpathiotakis, et al. Contrast digital mammography: Initial clinical experience. *Radiology*, **228**:842-850 (2003).

12   F.Diekman *et al.*, Use of Iodine-based Contrast Media in Digital Full-field Mammography-Initial Experience, *Fortschr. Röntgenstr* **175**, 342-345 (2003).

13   Lewin JM, Isaacs PK, Vance V, Larke FJ. Dual-energy contrast-enhanced digital subtraction mammography - feasibility. *Radiology*, **229**:261-8 (2003).

14   A.D.A. Maidment and M. Albert. Tissue Discrimination Methods in Mammography. In Advances in Digital Radiography edited by Ehsan Samei, Radiological Society of North America, 189-197 (2003)

15   David J. Getty, Stereoscopic and Biplane Imaging. In Advances in Digital Radiography edited by Ehsan Samei, Radiological Society of North America, 199-209 (2003)

16   Niklason LT, Christian BT, Niklason LE, et al. Digital tomosynthesis in breast imaging. *Radiology* **205**:399–406 (1997).

17   Elizabeth A. Rafferty, Breast Tomosynthesis, In Advances in Digital Radiography edited by Ehsan Samei, Radiological Society of North America, 219-226 (2003)

18   A.D.A. Maidment, P.R. Bakic, and M. Albert. Effects of quantum noise and binocular summation on dose requirements in stereoradiography. *Med. Phys.*, **30**:3061-3071 (2003).

19   P.R. Bakic, M. Albert, and A.D.A. Maidment. Dose Requirements in Stereoradiography. In Physics of Medical Imaging edited by L. Antonuk and M.J. Yaffe, Proceedings of the SPIE, **4682**:126-137 (2002).

20   P.R. Bakic, M. Albert, and A.D.A. Maidment**.** "2-AFC Observer Study of Digital Stereomammography". In Image Perception, Observer Performance and Technology Assessment edited by D.P. Chakraborty and E.A. Krupinski, Proceedings of the SPIE, **5034**:1-9 (2003).

21   J. M. Boone, N. Shah, and T. R. Nelson, A comprehensive analysis of DgNCT coefficients for pendant-geometry cone-beam breast computed tomography. *Med. Phys.* **31**:226-235 (2004).

22   H. Kundel, C. Nodine, L. Toto *et al.,* A circle cue enhances detection of simulated masses on mammographic backgrounds, in *Medical Imaging 1997, Image Perception,* edited by H.L. Kundel, Proceedings of the SPIE, **3032**:81-84, 1997.

23   DICOM Standard PS-3.14-2000: Gray-scale Standard Display Function, National Electrical Manufacturers Association, Roslyn, VA, 2000.

24   A.E. Burgess, Comparison of receiver operating characteristics and forced choice observer performance measurement methods. *Med. Phys.* **22**:643-655 (1995).

25   J. Hsu, D.M. Chalberg, C.F. Babbs, Z. Pizlo, and E. Delp, Preclinical ROC studies of digital stereomammography. *IEEE Trans. Med. Imaging* **14**:318-327 (1995).

26   T. Wu, R.H. Moore, E.A. Rafferty, and D.B. Kopans. Measurement of X-ray attenuation coefficient of breast tissue and tumor in digital tomosynthesis (abstr.). In: RSNA Scientific Assembly and Annual Meeting Program (RSNA, Oak Brook, IL, 2003), 425.

27   P.R. Bakic, M. Albert, D. Brzakovic, and A.D.A. Maidment. Mammogram Synthesis using a 3-D simulation: I. Breast Tissue Model and the exam simulation. *Med. Phys.*, **29**:2131-2139 (2002)

28   P.R. Bakic, M. Albert, D. Brzakovic, and A.D.A. Maidment. Mammogram Synthesis using a 3-D simulation: II. Evaluation of Synthetic Mammogram Texture. *Med. Phys.*, **29**:2140-2151 (2002)

29   P.R. Bakic, M. Albert, D. Brzakovic, and A.D.A. Maidment. Mammogram Synthesis using a three-dimensional simulation: III. Modeling and evaluation of the breast ductal network. *Med. Phys.*, **30**:1914-1925 (2003)