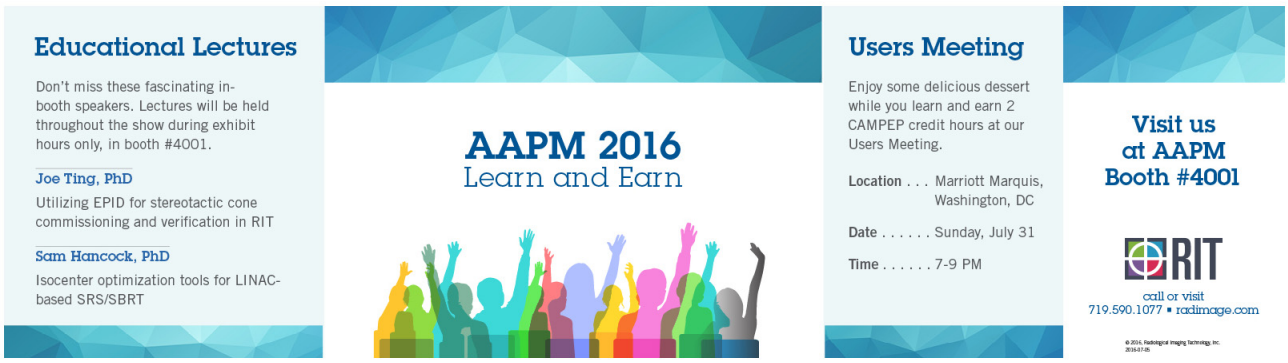
Medical Physics
The International Journal of Medical Physics Research and Practice

Effects of quantum noise and binocular summation on dose requirements in stereoradiography

Andrew D. A. Maidment, Predrag R. Bakic, and Michael Albert

Citation: Medical Physics **30**, 3061 (2003); doi: 10.1118/1.1621869
View online: http://dx.doi.org/10.1118/1.1621869
View Table of Contents: http://scitation.aip.org/content/aapm/journal/medphys/30/12?ver=pdfcov
Published by the American Association of Physicists in Medicine



Effects of quantum noise and binocular summation on dose requirements in stereoradiography

Andrew D. A. Maidment,[a] Predrag R. Bakic, and Michael Albert
Department of Radiology, Hospital of the University of Pennsylvania, 3400 Spruce Street, Philadelphia, Pennsylvania 19104

(Received 23 September 2002; accepted for publication 5 September 2003; published 18 November 2003)

In the case of a quantum-noise limited detector, signal detection theory suggests that stereoradiographic images can be acquired with one half of the per-image dose needed for a standard radiographic projection, as information from the two stereo images can be combined. Previously, film–screen stereoradiography has been performed using the same per-image dose as in projection radiography, i.e., doubling the total dose. In this paper, the assumption of a possible decrease in dose for stereoradiography was tested by a series of contrast-detail experiments, using phantom images acquired over a range of exposures. The number of visible details, the effective reduction of the dose, and the effective decrease in the threshold signal-to-noise ratio were determined using human observers under several display and viewing conditions. These results were averaged over five observers and compared with multiple readings by a single observer and with the results of an additional observer with limited stereoscopic acuity. Experimental results show that the total dose needed to produce a stereoradiographic image pair is approximately 1.1 times the dose needed for a single projection in standard radiography, indicating that under these conditions the human visual system demonstrates almost ideal binocular summation. [DOI: 10.1118/1.1621869]

Key words: digital radiography, stereoradiographic technique, dose, observer study, human perception

I. INTRODUCTION

A significant limitation of projection radiography is that relevant findings are often obscured or mimicked by the x-ray shadows of other anatomical structures (summation artifacts). By comparison, stereoradiography allows superimposed structures to be spatially separated, reducing the confounding effects of overlap. Stereoradiography was first proposed by Thomson in 1896[1] and continued to be used for certain radiographic procedures until the 1980s. The first recorded use of stereomammography occurred in 1930.[2] The advantages and disadvantages of stereoradiographic techniques are discussed in several radiographic textbooks and articles, e.g., Curry *et al.*[3] Improvements in image quality of conventional film–screen radiographic systems combined with the added effort, cost, and patient dose led to a decline in the use of stereoradiography and stereomammography.[3] Development of digital x-ray detectors and softcopy reading has resulted in a renewal of interest in stereoradiography as a potentially useful method of reducing summation artifacts.

The issue of dose requirements for stereoradiography is generally either not mentioned[4–6] in textbooks or it is suggested that a stereoradiographic technique requires twice the exposure of a single projection technique.[7] For modern quantum-limited imaging systems, exposure constraints set fundamental limits on the detectability of small, low contrast objects. For conventional single-projection viewing, the performance of human observers has been found to be closely related to these fundamental limitations, as illustrated by the classic work of Rose.[8] There is, however, scant literature on the detectability of objects when two projections, each subject independently to quantum mottle, are viewed stereoscopically.

Human observers perform certain visual tasks more efficaciously when fusing information from both eyes.[9,10] For example, threshold measurements of sinusoidal patterns have shown a decrease in threshold (i.e., an increase in sensitivity) of a factor of $\sqrt{2}$ when viewing is performed binocularly as opposed to monocularly.[11] This increase in sensitivity is precisely as expected for an ideal observer combining the signal from each eye. The quantum mottle in x-ray projections is different from the sources of noise in these experiments; in a radiograph, a given realization of the noise is permanently recorded at the time of acquisition and it is presented to each eye as a fixed pattern. For example, we note that flat random-noise fields presented stereoscopically cause lustre[12] (i.e., the image appears to shimmer) due to an attempt by the visual system to stereoscopically fuse the random bright and dark variations.

Thus, the question of threshold detection of objects acquired stereoradiographically is different from detection in other stereoscopic settings. To address this difference and examine the role of dose in stereoscopy we performed an observer study. Observer studies of stereoradiography have been reported in the literature previously. Kundel *et al.*[10] compared stereoscopic acuity and effects of monocular depth cues under direct vision and in stereofluoroscopy, while Goodsitt, Chan, and Hadjiiski[13] studied depth perception in stereomammography. Berkson *et al.*[6] compared the number

of false negatives/positives in mono- and stereoscopic chest radiographs and Hsu *et al.*[14] studied detection of simulated abnormalities in stereomammography.

To investigate the detectability of small and low-contrast objects under single-projection and stereoradiographic conditions we performed an experiment in which human observers were asked to determine the limits of visibility of objects in a contrast–detail (C–d) phantom, thus allowing the psychophysical thresholds[15] for the two conditions to be compared. Stereoscopic conditions consisted of each eye being presented with an independently acquired radiographic projection. Single-projection or ''monoscopic'' conditions consisted of both eyes viewing the same radiographic projection. From the point of view of signal-detection theory, the detectability of objects under stereoscopic conditions should be equivalent to the detectability of such objects in a single projection acquired at an exposure equal to the sum of the exposures used to acquire the stereoscopic views.

The ability of the human visual system to combine information from both eyes in such a manner is referred to in the psychophysical literature as ''binocular summation.'' While binocular summation has been demonstrated to occur under a variety of conditions, little attention has been paid to conditions relevant to stereoradiography in which each eye is presented with a pattern of quantum mottle which is fixed at the time of radiographic acquisition but which is independent of the pattern of mottle presented to the other eye. Fixed but interocullarly uncorrelated noise has been investigated, but for the detection of sinusoidal patterns[16] (with only one observer) or such patterns windowed by a Gaussian.[17] While in principle the image of any object can be expressed in term of Fourier components, the human visual system shows significant nonlinearities, so experiments using simple objects of finite spatial extent are clearly suggested. In the work being reported here, we have chosen to address this issue under the simplifying circumstance that all of the objects in the C–d phantom lie in a single plane that coincides with both the imaging plane and the face of the monitor with which the images were presented to the observer. This obviates the need to address the issue of the varying abilities of observers to fuse images containing objects in different planes.

II. MATERIALS AND METHODS

A. Image acquisition

Radiographic images of an RMI-180 mammographic C–d phantom (Gammex RMI, Middleton, WI) were used in the observer study. The Lucite™ phantom consists of 90 disc-shaped objects aligned in nine columns. Adjacent objects within the same column differ in diameter by a factor of $\sqrt{2}$, with a range of 7.07–0.312 mm. The thickness of the objects in adjacent columns also differs by a factor of $\sqrt{2}$, with a range of 1.0–0.062 mm. This phantom was chosen based upon the size and contrast of the objects. The phantom projections were acquired with a prototype amorphous-selenium DirectRay flat panel digital x-ray detector (Hologic/Direct Radiography, Newark, DE),[18] without a grid to avoid Moire artifacts. Observers were presented with a subregion of 650 $\times$810 pixels of dimension 0.45 mm, magnified approximately threefold from the original resolution of 139 μm detector elements (dels), which included all the phantom details.

Imaging was performed at 60 kVp and at six milliampere-seconds (mAs) stations in the range of 2–100 mAs, using a general radiography Bi-150 30/50 x-ray tube (Siemens, Munich, Germany) with tungsten target, a measured HVL of 1.34 mm Al (at 60 kVp), and a Heliophos 5S generator (Siemens, Munich, Germany). Additional filtration of 6 cm of Lucite was used to simulate tissue and appropriately harden the beam. The block of Lucite was positioned near the x-ray focus in order to reduce scatter and provide for uniform beam filtration. The phantom was imaged in contact with the detector (112 cm from the focal spot) and was placed within a wide Lucite frame of equal thickness, providing uniform scatter throughout the whole phantom area. At each mAs station, we acquired five images of the phantom, ten bright field images (same mAs, but with the phantom and frame removed), and ten dark field images (x rays off). The bright and dark fields were used for correcting the gain and offset variations of the individual dels. Ten images of each were chosen to reduce the effect of the x-ray quantum noise and detector noise on the corrections.

B. Image processing

The raw acquired phantom images were corrected to compensate for the effects of detector nonuniformity and adjusted for the nonlinearity of human contrast sensitivity.

1. Detector nonuniformity

A standard procedure for reducing the pixel variations due to the gain and offset variations of dels is to apply a dark and bright field correction. Corrected pixel values were computed as

$$(I_{i,j})_{\mathrm{corr}} = K\frac{I_{i,j} - E\{D_{i,j}\}}{E\{B_{i,j}\} - E\{D_{i,j}\}} + \bar{I}_{\mathrm{new}}, \tag{1}$$

where I and I_{corr} are the original and corrected pixel values, respectively; $E\{D\}$ and $E\{B\}$ are the averages of the dark and the bright fields, respectively; and K and I_{new} are parameters transforming the range (contrast) and the mean image pixel values. Subscripts i and j denote the position of the pixel in the image array.

The existence of malfunctioning dels, whose corresponding pixel values are not proportional to the incident x-ray flux and which differ significantly from their neighboring pixels, was observed and the corresponding pixel values corrected. Recently, several researchers have investigated this problem. Aach and Metzler[19] proposed image deconvolution using the spectral analysis of the defect image and Tang *et al.*[20] used wavelet analysis for identification and interpolation of flat panel images used for cone beam CT.

In our case, the average value and the variance of all the dark fields were computed for each del, and the same procedure was repeated for all of the bright fields. First, dels for which the digital value was constant across all of the dark-

and bright-field images were identified as nonfunctioning. The pixels corresponding to the nonfunctioning dels were replaced by the average of their immediate neighbors. Second, we computed the spatial variance within 3×3 pixel neighborhoods of the phantom images, and averaged the spatial variance values over a large neighborhood, of size 5×5 pixels, and two oriented small neighborhoods, of size 3×5 and 5×3 pixels. The averaged variance values were then tested against the following criterion:

$$\frac{\langle\sigma\rangle_{\mathrm{sm}}}{\langle\sigma\rangle_{\mathrm{lg}}} > C, \tag{2}$$

where $\langle\sigma\rangle_{\mathrm{lg}}$ is the spatial average of the values of the pixel standard deviations computed over the 5×5 neighborhood, $\langle\sigma\rangle_{\mathrm{sm}}$ is the larger of the averages computed over the 3×5 and 5×3 neighborhoods, and C is the criterion value. Pixels satisfying the criterion were replaced by the average of their immediate neighbors. Many of the affected dels were grouped along several vertical and horizontal lines. The value of parameter C that would correct all of the pixels along these lines would also produce an undesirably high number of replaced pixels randomly distributed over the image. Replacing too many correct pixels might change the statistics of the noise and affect the results of the study. Therefore, we identified the lines of malfunctioning dels in images and directly applied the correction, which affected fewer than 0.3% of pixels. For the rest of the image, the value of $C=2.1$ was used, which resulted in fewer than 3% of pixels being replaced.

2. Nonlinear human contrast sensitivity

Human contrast sensitivity is a nonlinear function of luminance. Relative changes in luminance are more easily noticed in bright image areas than in dark areas.[21] Perceptual linearization is an image transformation that adjusts the brightness so that equal changes in the quantity being displayed (here, x-ray fluence at the detector surface) will be equally perceived. The contrast sensitivity of the human visual system is approximated by Barten's model.[22] The model was derived for a standard target of a 2 deg×2 deg square filled with a horizontal or vertical sinusoidal modulation of 4 cycles/deg, placed in a uniform background of the mean target luminance.[21] The threshold modulation at which the target becomes just visible to the average observer defines the just-noticeable difference (JND), at the luminance value of the background. The interpolation of the luminance levels corresponding to 1023 JNDs is given by the grayscale standard display function:[21]

$$\log_{10} L(N_\Delta) = \frac{\Sigma_{i=0}^{4} a_i (\ln N_\Delta)^i}{1+\Sigma_{i=1}^{5} b_i (\ln N_\Delta)^i}, \tag{3}$$

where N_Δ is the JND index (i.e., the number of just-noticeable differences) and L is the corresponding luminance value. Coefficients a_i and b_i of the rational polynomial interpolation are given in the DICOM standard.[21] We have incorporated the grayscale standard display function into the transformation from the flat-panel output 12 bit pixel value to the monitor input 8 bit digital-driver value. This transformation consists of three parts.[23] This decomposition offers flexibility to perform observer studies using different monitors or at different overall brightness or contrast levels.

In the first step, the monitor characteristic, relating luminance levels to the 8 bit digital driver levels, was interpolated from photometric measurements made by a TEK Lumacolor J17 photometer (Tektronix, Beaverton, OR). The luminance values were measured in the center and at the periphery of the monitor. The luminance values corresponding to the monitor input 8 bit digital-driver levels of 1 and 255 were equal to 1.5 and 95.3 Cd/m^2, respectively. The average values at 24 digital driver levels were used for interpolation. A fourth-order polynomial fit was used for interpolation with maximum error of 1.68 digital driver levels (0.6%). Second, the transformation from the luminance values to the number of JNDs was approximated by the grayscale standard display function, given by Eq. (3). Third, the linear transformation from 12 bit flat panel output to the JND index was used to adjust the common overall brightness and contrast of all the phantom images included in the study. The overall brightness was selected in a small preliminary observer study as the one which gave the most consistent increase among four observers in the number of details seen stereo- versus monoscopically. Based on that study, the midlevel between the largest, highest contrast element and its annular neighborhood was transformed to a JND index of 350, corresponding to a digital driver level of 90. Little variation in performance was noted regardless of the average brightness used.

We have used two strategies for displaying the phantom images. In the first approach, we kept the contrast between the first phantom detail and its background fixed for all the images displayed in the study. The average brightness in the first detail and the average brightness around the first detail were mapped to JND indices 328 and 372, corresponding to digital driver levels of 73 and 110, respectively. In the second approach, the images were modified to have fixed variance of the background noise. Each 12 bit pixel value was multiplied by the ratio of the standard deviation of the background noise desired for display to the standard deviation of the background noise calculated over the region around the first detail. The JND indices were chosen to ensure visualization of ±3 s.d. about the mean. The average background was mapped to JND index 400 and the mapping was scaled to give a standard deviation of 50 JNDs in the background. A discussion of the effects of the display strategy on the results of the experiment is given in Sec. IV. In both strategies, the linear transformation from the flat-panel output to the JND index was calculated independently for each image to compensate for the small deviations in actual exposure from the exposure expected from the nominal mAs. An example of a processed phantom image is shown in Fig. 1.

C. Observer study

1. Selection and display of images

Six medical physicists participated in the study. Each observer was presented with 60 images, consisting of 30 for

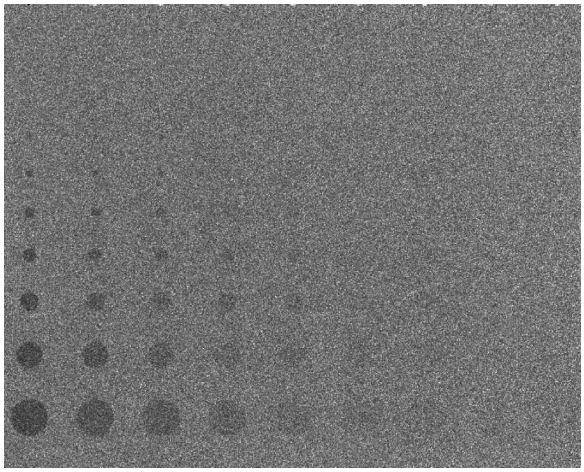


FIG. 1. An example of a phantom image used in the study. The image was acquired with an exposure of 10mAs and postprocessed to compensate for the effects of detector nonuniformity and to adjust for the nonlinearity of human contrast sensitivity. The image is displayed with the opposite polarity to film.

monoscopic viewing and 30 for stereoscopic viewing. The monoscopic images were displayed so that both the left and right eye saw the same image, while for stereoscopy, the left and right eye images were different. Each of the observers saw the same set of monoscopic images (there were five images for each of six exposure levels). A set of five out of twenty possible stereo image pairs for each mAs was selected randomly for each observer.

The two sequences of 30 images, one for mono- and the other for stereoscopic viewing, were separately randomly permuted and then interleaved into a single sequence of the form:

$$\ldots M_n S_n M_{n+1} S_{n+1} \ldots,$$

where M_i and S_i represent mono and stereo images, respectively. Next, the positions of the mono and stereo images, in each mono–stereo pair, were randomly changed in order to break the repetition of mono following stereo, and to decrease possible observer bias.

2. Test of stereoscopic acuity

The stereoscopic acuity of the observers participating in the study was tested using a standard clinical test, the RANDOT Stereotest (Stereo Optical Company, Chicago, IL). The test was administered to identify individuals lacking stereoscopic vision as a crude correlate of a test of binocular summation. The test consists of image pairs with or without angular disparity. The images in a pair are orthogonally polarized and the stereo effect is observed using appropriately polarized glasses. The images are divided into three groups with different content (geometric figures, animals, and triplets of circles) designed to assess the stereoscopic acuity in a wide range of patients' age. The images in each of the groups gradually decrease in angular disparity and roughly cover the range of 20–500 arcsec when viewed from 40 cm distance. Stereoscopic acuity is measured by identifying the image, and the corresponding angular disparity, beyond which the observer cannot distinguish between the objects with and without disparity. Table I shows the minimum observable angular disparity for the six observers. One individual (observer 6 in Table I) was eliminated from discussion below, based upon poor stereoscopic vision.

TABLE I. Results of the RANDOT stereo test for six observers. Tabulated are the values of the minimum observable angle of stereoscopic disparity in arcseconds.

Observer	$\Delta\alpha_{\min}$ (arcsec)
1	25
2	70
3	30
4	25
5	25
6	200

3. Viewing conditions and scoring

The C–d experiments were performed in a darkened room, with a black monitor background. The distance between the observer's eyes and the monitor was approximately 1 m, so that the smallest objects had an apparent diameter of approximately 3 arc min. A 21 in. color monitor was used in the study (GDM-5410, Sun Microsystems, Santa Clara, CA) with a Creator 3D 960×680, 112 Hz graphics card (Sun Microsystems, Santa Clara, CA), and a stereoscopic goggles system (StereoGraphics, San Rafael, CA). A transmitter, attached to the monitor, controlled the opening and closing of the liquid crystal shutters in the goggles synchronously with the two alternating images on the monitor. Therefore, each eye was presented with the corresponding image at a repetition frequency of 56 Hz. The goggles were used for both mono and stereo experiments to provide a comparable brightness level in both experiments.

Several training sessions for image scoring were organized for observers in order to establish a uniform decision criterion. 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. The goal was to prevent misinterpretation of the clustered background noise as phantom details. A graphical user interface was developed allowing an observer to identify the smallest detail seen in a column by clicking over the detail position.

We have analyzed the relationship between the inter- and intraobserver variability by having one observer perform the C–d experiment five times with the same image order as presented to each of the other observers. Thus, we could compare the average C–d performance of a group of observers with a single round of experiments, and the average performance of a single observer after repeating the study multiple times. For further comparison, we have also repeated the C–d experiments for all the observers without the use of

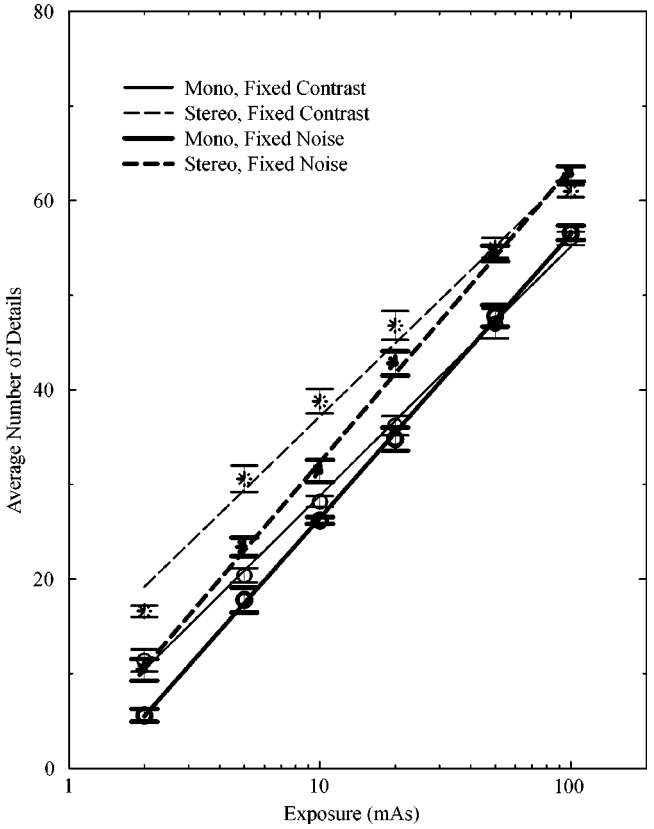


FIG. 2. Number of details seen in the C–d phantom for two image display modes: fixed contrast (thin) and fixed noise (bold). Shown are the average number of details (symbols), the standard error (bars), and the linear fits (lines), for mono- (solid) and stereoscopic (dashed) viewing.

the stereoscopic goggles. Due to the 112 Hz switching between the two stereoscopic images on the monitor, the observers were, effectively, presented with an average of the left and right eye image. Therefore, we could compare the performance of observers attempting to utilize information from disparate images with the performance of observers looking at the combined images.

III. RESULTS

The observers took, on average, 39 min (s.d.=2.8 min) to score a set of 60 combined mono and stereo images. Figure 2 shows the number of details seen mono- and stereoscopically by a single observer, averaged over all images with the same mAs. Figure 3 shows the difference between the number of details for stereo and mono images, for all the observers. When averaging over different observers, we have included the results of all but one of the observers. This observer reported uncorrected vision in one eye, corresponding to low stereo acuity, and demonstrated a large value of the minimum observable angular disparity in the stereo vision RANDOT test (Table I, observer 6). Furthermore, the difference between the number of details seen stereo- versus monoscopically, for the observer in question (Fig. 3, points labeled by triangles), was significantly smaller compared to the others.

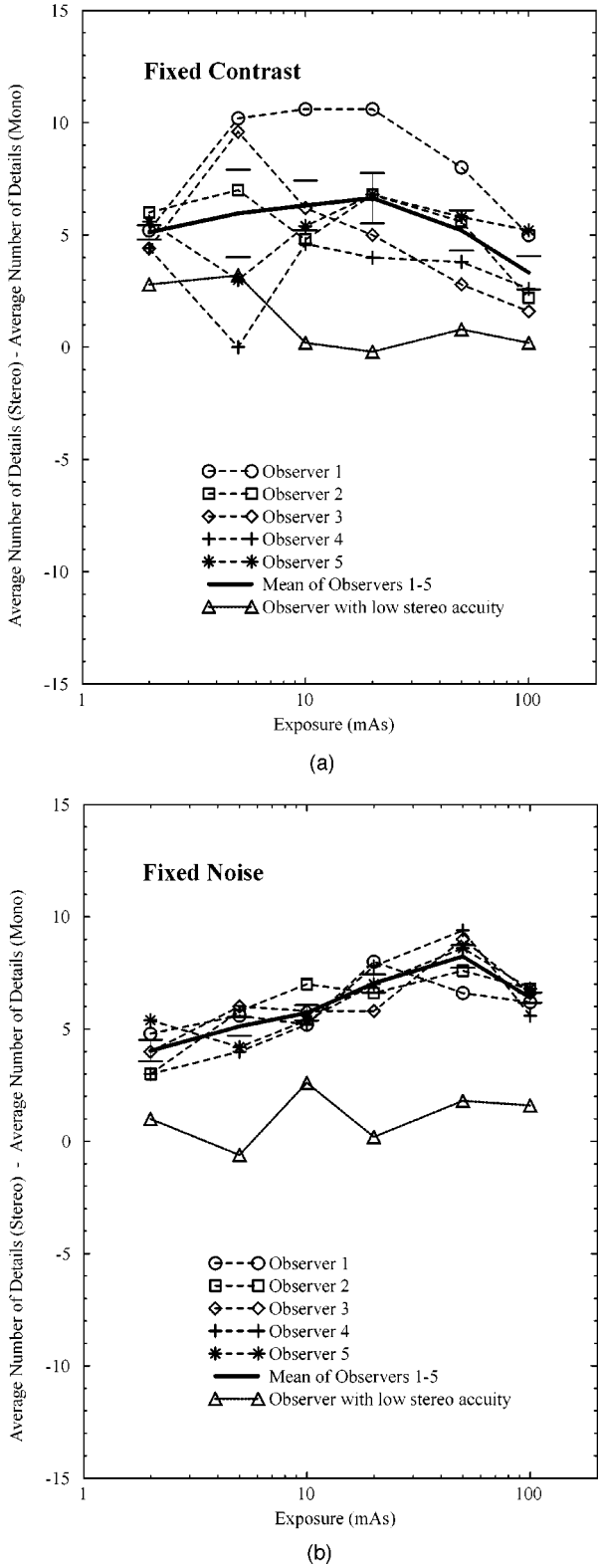


FIG. 3. Comparison between mono- and stereoscopic viewing. (a) Difference in the number of objects seen by stereo- vs monoscopy, for each observer, for images displayed with fixed contrast. (b) Difference in the number of objects seen by stereo- vs monoscopy, for each observer, for images displayed with fixed noise.

The detection task in this experiment consisted of simple low-contrast objects obscured by noise. The noise characteristics of our detector can be well modeled by white noise.[18] Thus, according to Rose's model,[8] objects should appear to

be visible when their signal-to-noise ratio exceeds a threshold value SNR_T (Rose's "k"). For simple objects in white noise, the signal-to-noise ratio is given by

$$\mathrm{SNR} = td\sqrt{KX}, \tag{4}$$

where X is the exposure, t is thickness of the object, d is the diameter of the object, and K is a proportionality constant. For single-projection viewing, one thus expects

$$\mathrm{SNR}_M = td\sqrt{KX_M}, \tag{5}$$

where a subscript M has been added to indicate that this corresponds to single-projection or "monoscopic" projection. For stereo, an observer would ideally combine signals from both views with complete efficiency, giving

$$\mathrm{SNR}_{S_{ideal}} = td\sqrt{KX_S} = td\sqrt{2KX_M} = \sqrt{2}\,\mathrm{SNR}_M, \tag{6}$$

where $X_S = 2X_M$ represents the total dose for two projections. However, as the human visual system may not be completely efficient in binocular summation of the two views, we write

$$\mathrm{SNR}_S = td\sqrt{KX_{\mathrm{S\,eff}}}, \tag{7}$$

where $X_{\mathrm{S\,eff}}$ is the effective exposure that would result in a single projection image in which observers should report the same number of objects as for the stereoscopic pair.

A. Analysis of C–d curves

Under single projection or "monoscopic" conditions, the expectation is that the objects at the limit of visibility should correspond to a fixed threshold, and thus

$$\log_{10} t + \log_{10} d = \log_{10} \mathrm{SNR}_T - \tfrac{1}{2}\log_{10} K - \tfrac{1}{2}\log_{10} X, \tag{8}$$

gives a line on a log–log plot of thickness t versus diameter d. For stereoscopic conditions,

$$\log_{10} t + \log_{10} d = \log_{10} \mathrm{SNR}_T - \tfrac{1}{2}\log_{10} K - \tfrac{1}{2}\log_{10} X_{\mathrm{S\,eff}}. \tag{9}$$

As both lines have slope -1 on a log–log plot, the vertical displacement between these lines is independent of the value of the horizontal coordinate, thus

$$\log_{10} d_M - \log_{10} d_S = \log_{10}\frac{d_M}{d_S} = \log_{10}\sqrt{\frac{X_{\mathrm{S\,eff}}}{X_M}}, \tag{10}$$

where d_M and d_S are the diameters at the limits of visibility under monoscopic and stereoscopic conditions for a fixed thickness and a given per-projection exposure.

For each observer we calculated their C–d curves by averaging the size of the smallest visible detail in each column, over a set of images, i.e., for the same exposure. Computing the average detail size is problematic when no details are visible in a given column for one or more images, since the contribution of those images to the average is ambiguous. To avoid this problem, we chose the median diameter for generating C–d curves. Figure 4 shows the average C–d curves for a single observer (the same observer as in Fig. 2) and Fig. 5 shows the linear approximation of the C–d curves from Fig. 4. See also Fig. 5.

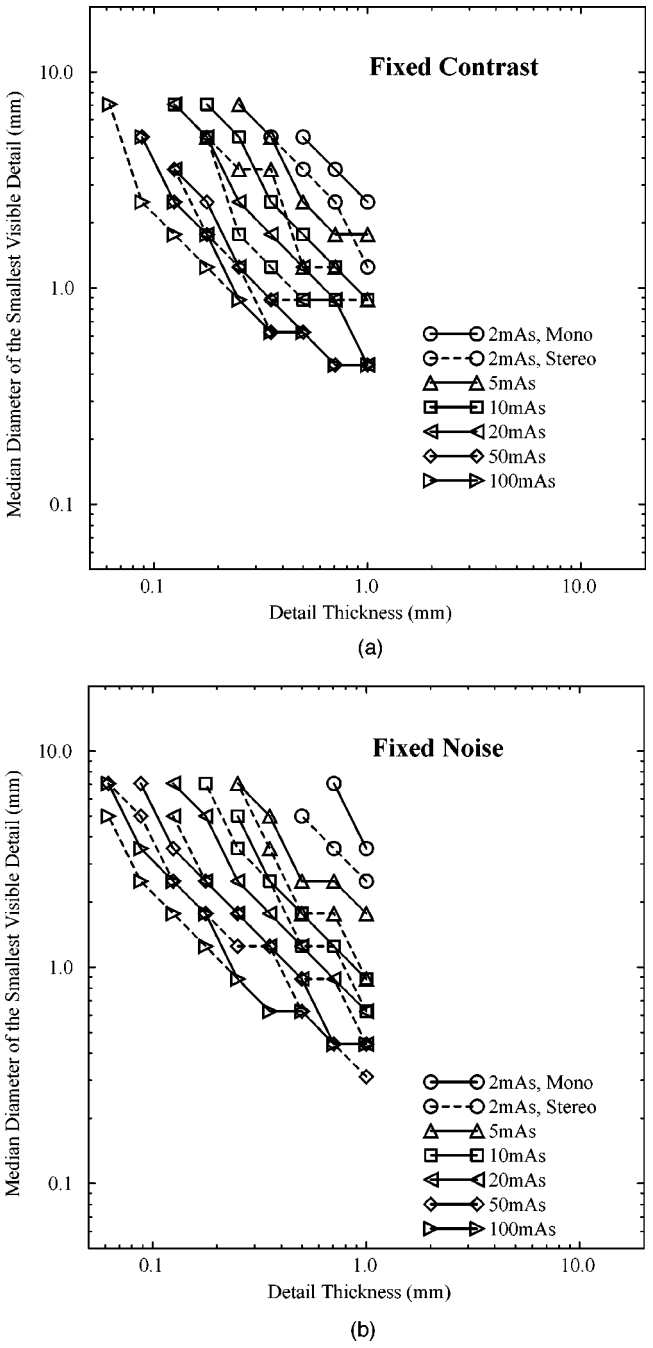


FIG. 4. C–d curves for a single observer (observer 1), for monoscopic (solid) and stereoscopic (dashed) viewing, and two display conditions: (a) with fixed contrast and (b) with fixed noise.

Figure 6(a) shows the summation efficiency $R_X = X_{\mathrm{S\,eff}}/(2X_M)$, calculated using the fitted C–d curves, as per Eqs. (8) and (9). The summation efficiency is the fraction of the total exposure of a stereo image pair ($2X_M$) at which a single projection image would need to be acquired and still result in the same number of objects visible as was observed when two projections, each taken at exposure X_M, were viewed stereoscopically. If binocular summation was perfectly efficient under radiographic conditions, the ratio R_X would be equal to 100%. R_X is shown averaged over the five observers, with standard errors, for both display conditions

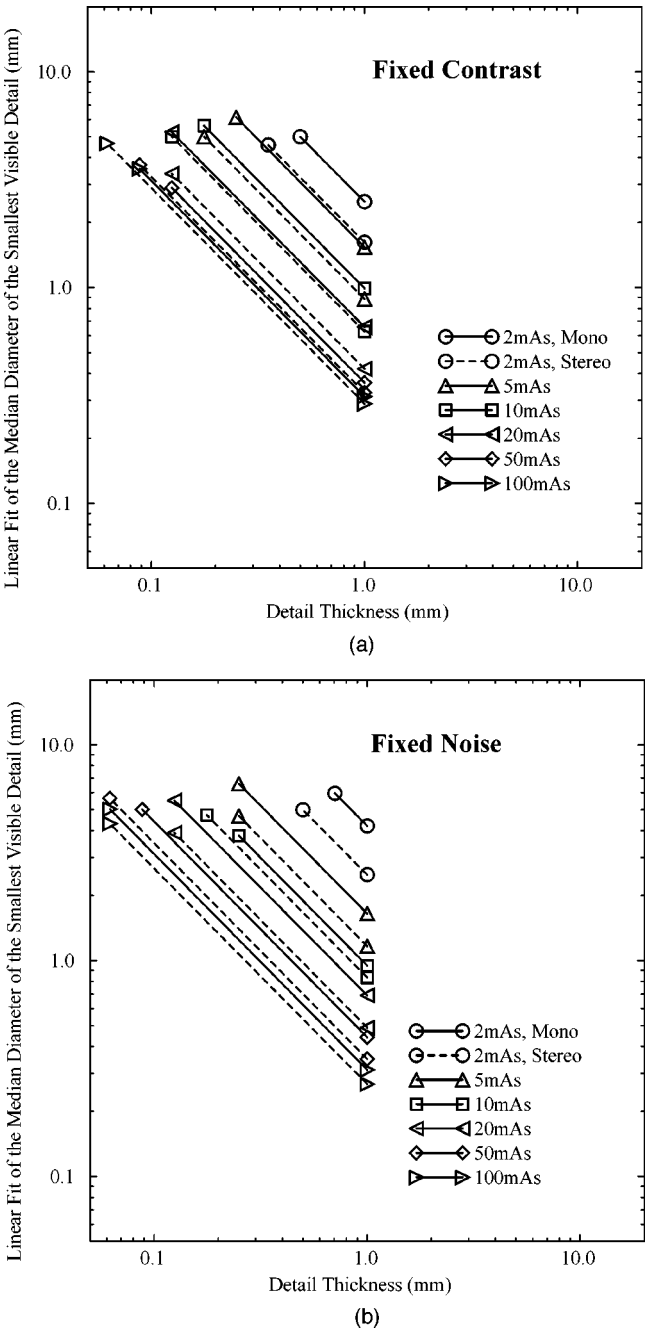


FIG. 5. Linear approximation of C–d curves for a single observer (observer 1), for monoscopic (solid) and stereoscopic (dashed) viewing, and two display conditions: (a) with fixed contrast and (b) with fixed noise.

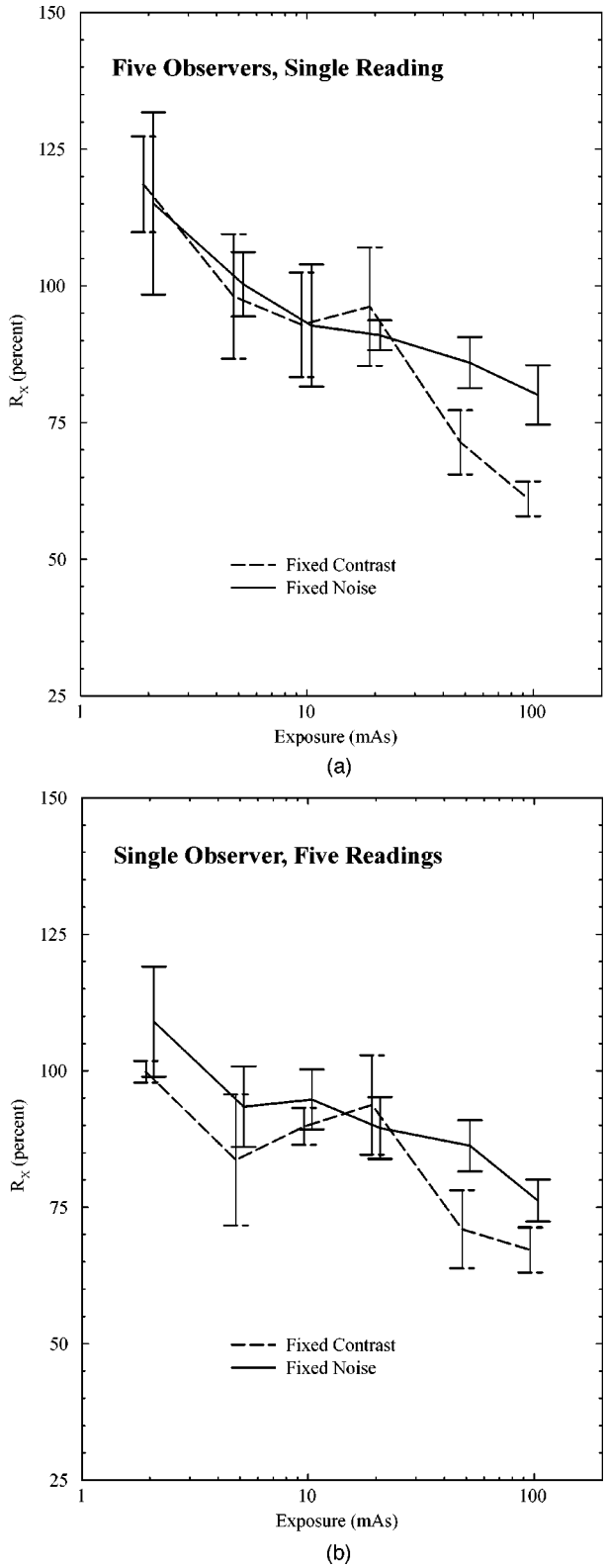


FIG. 6. Summation efficiency, $R_X=[X_M/X_S]/2=[(d_M/d_S)^2]/2$, as a function of exposure, for two display conditions, and averaged: (a) over five observers and (b) over five readings by a single observer (observer 3). Bars represent the range of ± 1 s.d. error.

and for each per-projection exposure level. Figure 6(b) presents the average of R_X for five readings performed by a single observer and both display conditions. For clarity of presentation, the results obtained with different display conditions are slightly horizontally displaced on the graph.

B. Analysis of threshold SNR

As the experimental C–d curves, such as those shown in Fig. 4, are not quite linear as expected from the Rose model, a second analysis was performed in which thresholds were estimated in a manner that did not directly depend upon the C–d curves having a slope of -1. Figure 7 shows the number of objects with measured SNR_M exceeding a given threshold, SNR_T. SNR_M was measured experimentally from contrast and noise per pixel for objects in each image. In

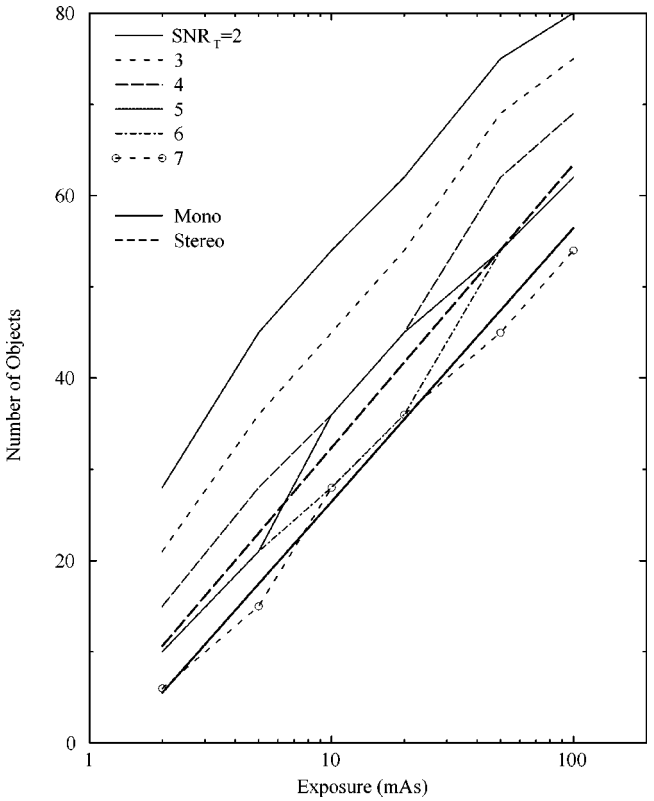


FIG. 7. Number of details seen in the C–d phantom by an ideal observer, with various threshold SNR values (thin). The results of the human observer from Fig. 2 for the fixed noise display and stereoscopic viewing (bold line) are given for comparison.

Rose's model these curves represent the number of visible objects as a function of exposure for various values of the SNR at the threshold of visibility, SNR_T. Thresholds SNR_T of two through seven are shown; the number of objects decreases with increasing threshold. Superimposed on Fig. 7 are the fits to the actual performance of one observer from Fig. 2. For monoscopic viewing, the data for this observer fall between the theoretical curves corresponding to SNR_T thresholds of 5 and 7. This is in approximate agreement with the generally accepted estimates of Rose's k, particularly in light of the relatively stringent requirements for scoring an object as visible, detailed in Sec. II C 3. For an observer capable of combining information from stereoscopic images the signal-to-noise ratio of each object should increase by a factor of $\sqrt{2}$ (i.e., $\mathrm{SNR}_S = \sqrt{2}\,\mathrm{SNR}_M$), so that the object count as a function of per-projection exposure would correspond to a monoscopic signal-to-noise threshold ratio reduced by a factor of $\sqrt{2}$. Consistent with this, the real observer from Fig. 7 does show an increased sensitivity.

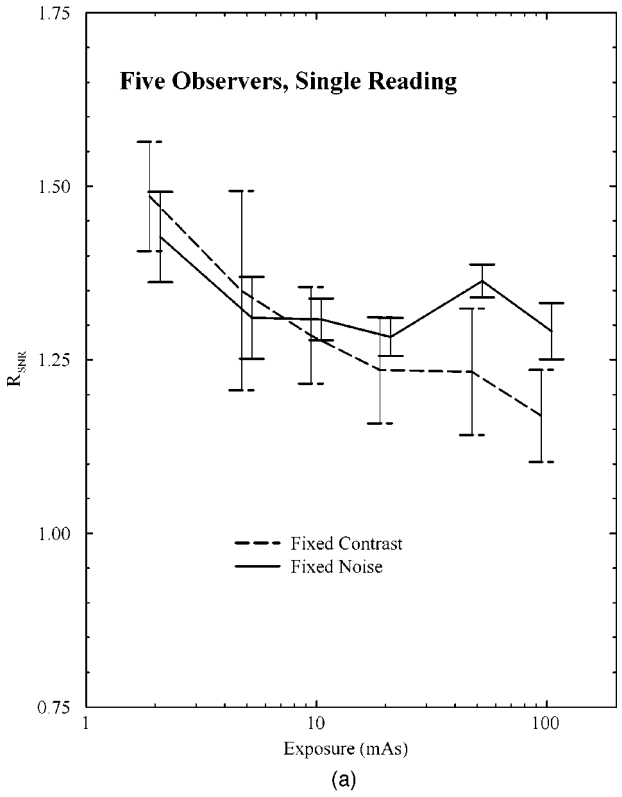
The results were analyzed by identifying a threshold SNR under each viewing condition. For each scoring of the phantom by each observer, the detected details had greater SNR_M than the details not detected, with only a few details near the threshold not following this pattern. Thus, for the scorings of the phantom by each observer, the threshold was calculated as the mean of the smallest SNR_M of the details detected and the largest SNR_M of the details not detected. The averages of these per-scoring values over the five monoscopic viewings give the monoscopic threshold SNR_{T_M} for each viewer at each exposure level, and similarly the five stereoscopic viewings give SNR_{T_S}. Over all scorings by all observers, an average of five objects out of ninety were found to either have SNR above the estimated threshold yet have not been scored as visible or to have SNR below the estimated threshold yet have been scored as visible. A hypothetical observer capable of reducing the quantum noise in the phantom images by using information from both images of the stereoscopic pair would find the signal-to-noise ratio of each disk increased by a factor of $\sqrt{2}$. Thus as both thresholds are rated in terms of the single-projection SNR, SNR_M, one would expect for perfect binocular summation the ratio $R_{\mathrm{SNR}} = \mathrm{SNR}_{T_M}/\mathrm{SNR}_{T_S}$ to take the value $R_{\mathrm{SNR}} = \sqrt{2}$. Figure 8(a) shows the ratio R_{SNR} averaged over the five observers, with standard errors, at each exposure level. Figure 8(b) shows the ratio R_{SNR} for five readings performed by a single observer. Discussion of the experimentally observed results is presented in Sec. IV.

IV. DISCUSSION

Our results indicate that humans can perform binocular summation on pairs of radiographic projections in such a manner as to increase the effective SNR by almost the factor of $\sqrt{2}$ expected for an ideal observer. The results show trends with exposure level and choice of display technique which are not accounted for in detail by this simple signal-detection theoretic model. Factors that contribute to the discrepancy with the prediction of signal-detection theory include the limitations of this model given the complexity of the human visual system, limitations of the hardware used for display of the images, and the possible effects of detector artifacts at higher exposures.

Figure 6, using an analysis based on the experimental C–d curves, shows that when two radiographic projections are viewed, one by each eye, the visibility of objects against background noise corresponds to a total dose of almost 100% of the sum of the doses used for each projection, indicating that the human visual system efficiently combines information from each eye. As the experimental C–d curves, such as those of Fig. 4, show a deviation from linearity not accounted for by our model, an alternate analysis was performed by attempting to estimate the threshold SNR_M of visible objects when viewed in single projections and when viewed stereoscopically. Figure 8 shows that the threshold SNR_M (computed based on the dose for a single projection) is approximately a factor of $\sqrt{2}$ higher for the monoscopic observation of a single projection compared to the stereoscopic observation of two independent projections.

Comparison of the standard error in parts (a) and (b) of Figs. 6 and 8 show that the level of interobserver variation was comparable to the level of intraobserver variation. Some differences were found between fixed contrast and fixed noise presentations. Further, there appears to be a systematic decrease in the observed efficiency of binocular summation with increased dose.



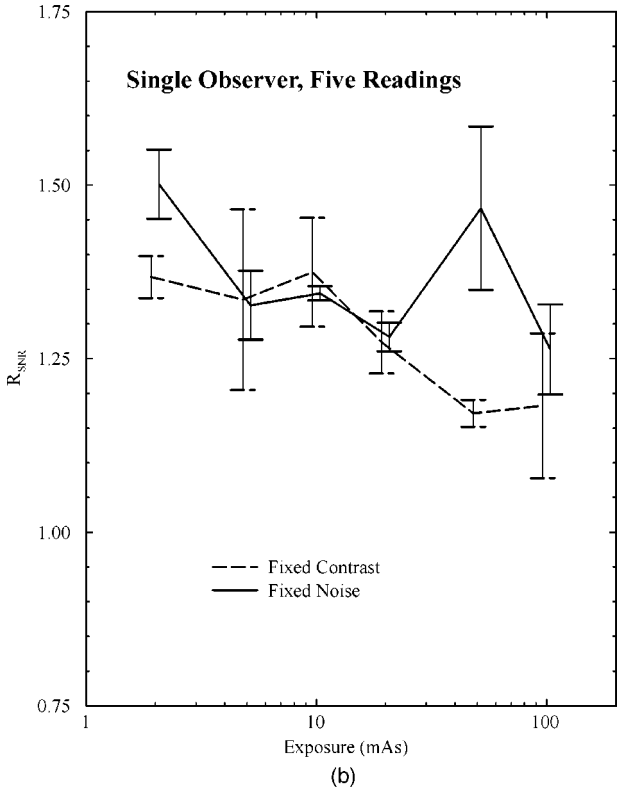


FIG. 8. SNR threshold ratio, $R_{\mathrm{SNR}}=\mathrm{SNR}_{T_M}/\mathrm{SNR}_{T_S}$, as a function of exposure, for two display conditions, and averaged: (a) over five observers and (b) over five readings by a single observer (observer 3). Bars represent the range of ± 1 s.d. error.

A. Effects of display conditions

The transformation from the flat panel 12 bit output to the number of JNDs was implemented in an image-dependent fashion, as described in Sec. II, by keeping fixed (a) the contrast of the largest detail or (b) the standard deviation of the background noise, in all the displayed images. Using the first approach, images acquired with high nominal exposure were displayed such that the background noise range was limited to only a few 8 bit digital driver levels, with the potential for quantization distortion. This distortion may effectively reduce the noise in images at certain exposure values. Although intuitively the noise reduction could be expected to improve the detection of the image objects, it is not obvious how it would affect the advantage of stereoscopic over monoscopic observations. The latter approach, however, changed the contrast of the largest detail in the images with different exposures, which could lead to possible saturation of the display. Moreover, the perceptual linearization is a nonlinear transformation that may affect the perception of the noise patterns. The effects of these issues on the results of the C–d experiment are not obvious.

For all performance measures and both display conditions, we observed better agreement with theory at lower exposures. Values of the summation efficiency R_X (Fig. 6) at the higher exposures showed more benefits from stereoscopy for the images displayed with fixed noise, than for the images with fixed contrast. The SNR ratio (Fig. 8) is near $R_{\mathrm{SNR}}=\sqrt{2}$ at low exposures, as expected for the hypothetical observer combining information from each of the stereo pair of images. At higher exposures, the ratio is less, but still greater than one. More experiments are needed to clarify the observed trend.

The summation efficiency values, averaged over the five observers and all six mAs stations, are $R_X=89.7\%\pm 8.4\%$ for the images displayed with fixed contrast and $R_X=94.2\%\pm 5.0\%$ for the images with fixed noise. The root-mean-square values of the SNR ratios over all six mAs stations, are $R_{\mathrm{SNR}}=1.30\pm 0.09$ for the images displayed with fixed contrast and $R_{\mathrm{SNR}}=1.33\pm 0.04$ for the images with fixed noise.

The total radiographic exposure needed to obtain a stereo pair is equal to the sum of the exposures for each individual radiographic projection. Figures 6 and 8 suggest that the same C–d performance for a single conventional projection taken at exposure X_M, could be achieved stereoscopically with total exposure that is less than twice the exposure for single conventional projections. Taking the ratio of the total exposure required for a stereoradiographic pair to the exposure required for a single-projection image yielding equivalent performance, it follows that

$$\frac{X'_{\mathrm{S\,eff}}}{X_M}=\frac{2X_M}{X_{\mathrm{S\,eff}}}=\frac{1}{R_X}. \tag{11}$$

Thus,

$$X'_{\mathrm{S\,eff}}=\frac{X_M}{R_X}, \tag{12}$$

where $X'_{S\,eff}$ is the total dose required for a stereoradiographic acquisition that would give the same performance on a C–d experiment as a single projection acquired with a dose of X_M, and $X_{S\,eff}$ is our estimate of the single-projection dose which would give the same performance as demonstrated when two images each acquired at a dose of X_M are viewed stereographically. Similarly, $X'_{S\,eff}$ can be calculated from R_{SNR} by

$$X'_{S\,eff}=\frac{X_M}{R^2_{SNR}/2}. \tag{13}$$

Thus, the total exposure required for stereomammography, calculated using Eq. (12), is 1.12 ± 0.13 and 1.06 ± 0.06 times that of a single conventional mammogram, for images displayed with fixed contrast and images displayed with fixed noise, respectively. When using Eq. (13), the total exposure required for stereomammography is 1.19 ± 0.13 and 1.13 ± 0.05 times that of a single conventional mammogram, for images displayed with fixed contrast and images displayed with fixed noise, respectively. Thus, there is a minimal increase in the total exposure and hence the dose required for this detection task in using a stereoscopic technique.

The obtained reduction in the SNR threshold can be compared with the results reported by Hsu *et al.*[14] In that study of stereomammographic detection of simulated objects, the performance was evaluated by a ROC analysis. Reported values of the area under the ROC curve, A_z, were 0.82 (stereo) vs 0.74 (mono) for their arrangement experiment, and 0.75 (stereo) vs 0.71 (mono) for their density experiment. Assuming that A_z is equal to the probability of correct detection, the corresponding reduction in threshold SNR can be computed as the ratio of stereoscopic and monoscopic d' values,[24] yielding $R_{SNR}=1.42$ for the arrangement experiment and $R_{SNR}=1.22$ for the density experiment.

We have performed an additional validation of the contrast-detail experiment results by comparing the performance measures averaged over five observers with the average performance of repeated readings by a single observer [see Figs. 6(b) and 8(b)]. There is little difference in the results of these experiments.

B. Effects of viewing condition

Results of the C–d experiments with and without the stereoscopic goggles are shown in Fig. 9. The error bars correspond to the standard error of the average values ($n=5$). For clarity of presentation, the two data sets are slightly displaced on the graph. The summation efficiency, averaged over all five observers (solid lines) was compared with the performance of observer 6, with low stereo acuity (dashed lines). While the average performance of the five observers with (bold) and without the goggles (thin) did not differ significantly, observer 6 performed much better without the goggles. This was expected as when using the goggles the image shown to the uncompensated eye of the low stereo acuity observer was effectively unused under either the stereoscopic or monoscopic viewing conditions, therefore showing basically no advantage of stereoscopy. Without the goggles, an averaged image was shown to both eyes, and the difference in the background noise for the mono and stereo images became visible, without, however, the advantage of parallax.

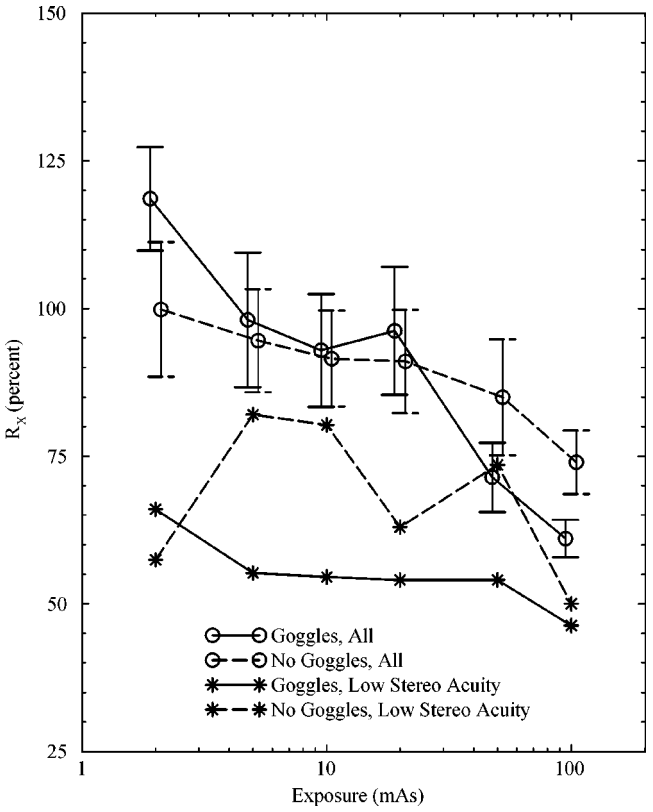


FIG. 9. Comparison of the summation efficiency, R_X, for the C–d experiments with (bold lines) and without (thin lines) stereoscopic goggles. Shown are the results averaged over five observers (solid lines) and for a single observer with low stereoscopic acuity (dashed lines).

C. Effects of additive detector noise

The images used in this study were acquired radiographically, with a digital x-ray detector. All images, regardless of mAs, were x-ray quantum noise limited, rather than detector noise limited. Thus, in the C–d experiments reported, the limiting factor for detection was the x-ray quantum fluctuations. The experimental conditions resulted in detector exposures in the range of 0.7–36 mR, thus the results reported are likely to be generally applicable to digital radiography and digital mammography. Extensions of this approach can be envisioned to address the issue of the dose requirements for tomosynthesis. However, in that instance, careful evaluation of the dominant noise source in the individual projections, and in the reconstructed data is necessary.

D. Other issues affecting the efficiency of stereoscopy

The following issues are potential sources of concern in our experiment. Since it was, to the best of our knowledge, the first C–d study of stereoscopy in radiography, no reference suggestions were found to address the experimental de-

sign. Future experiments, modified to improve the following conditions, may give a more accurate estimate of the advantages of stereoradiography.

Observers used the stereoscopic goggles while viewing both mono and stereo images, in order to keep the average brightness level comparable and to avoid a bias of having the mono and stereo images directly identified. The goggles reduce the brightness level of the scene for each eye. We have not analyzed the amount of this brightness reduction, or the spatial uniformity. Perceptual linearization was applied without taking into account this brightness reduction, and this might be of concern if the nonlinearity of the human contrast sensitivity was not properly compensated. The reduction in overall brightness when using the goggles presents a confounding issue in the experiments with and without goggles, but as the experiments without goggles were intended chiefly as a check of our experimental design and procedures, we have not pursued this question. Note, also, that structured anatomic noise is not included in this set of test images, but its modeling is considered for future research.[25,26] We have also begun another approach to testing the hypothesized benefits of stereoradiography using a 2-alternative forced-choice experiment.[27]

V. CONCLUSIONS

A series of C–d experiments was performed, testing the hypothesis that the radiation dose required for stereoradiography is one half the dose for a single x-ray image viewed monoscopically, due to the combination of images by the human visual system. For images acquired at a fixed exposure, more objects were detected under stereoscopic than monoscopic viewing conditions. For the lower range of the x-ray exposures tested, where the quantum fluctuations were easily visualized, the increase in the number of details observed was the same as expected for a hypothetical observer combining the two projections to remove noise. With increasing x-ray exposure, the increment between the number of objects detected under stereoscopic and monoscopic conditions decreased, but was always positive. The experimental results indicate that a stereoradiograph can be acquired at a dose approximately 1.1 ± 0.1 times that of a single projection radiograph. The study results are potentially influenced by the brightness reduction of stereoscopic goggles, background noise quantization, and the lack of anatomical noise. More significantly this study has addressed dose requirements to compensate for inefficiencies in binocular summation of the images, leaving questions of dose required for stereographic fusion for further work.

ACKNOWLEDGMENTS

This work was funded by the U.S. Department of Defense, Grant No. DAMD 17-98-1-8169. The authors are grateful to Dan Beideck, Scott Cupp, Natasa Milojkovic, and William Tao Shi for participating in the experiment.

[a]Author to whom correspondence should be addressed; electronic mail: maidment@rad.upenn.edu

[1] E. Thomson, ''Stereoscopic roentgen pictures,'' Electrical Engineer **21**, 256 (1896).
[2] S. L. Warren, ''The roentenologic study of the breast,'' Am. J. Roentgenol. Radium Ther. **24**, 113–124 (1930).
[3] T. S. Curry III, J. E. Dowdey, and R. C. Murry, Jr., *Christiansen's Physics of Diagnostic Radiology*, 4th ed. (Lea & Febiger, Philadelphia, PA, 1990).
[4] C. A. Jacobi and D. Q. Paris, *Textbook of Radiologic Technology*, 6th ed. (C. V. Mosby, St. Louis, MO, 1977).
[5] M. Tortorici, *Concepts in Medical Radiographic Imaging: Circuitry, Exposure, and Quality Control* (Saunders, Philadelphia, PA, 1992).
[6] J. Berkson, C. A. Good, D. T. Carr, and A. J. Bruwer, ''Identification of 'positives' in roentgenographic readings,'' Am. Rev. Respir. Dis. **81**, 660–665 (1960).
[7] R. R. Carlton and A. McKenna Adler, *Principles of Radiographic Imaging: An Art and a Science* (Delmar, Albany, NY, 1992).
[8] A. Rose, *Vision: Human and Electronic* (Plenum, New York, 1973).
[9] R. W. Reading, *Binocular Vision: Foundations and Applications* (Butterworths, Boston, MA, 1983).
[10] H. L. Kundel, P. R. Lynch, L. Peoples, and H. M. Stauffer, ''Evaluation of observer performance using televised stereofluoroscopy,'' Invest. Radiol. **2**, 200–207 (1967).
[11] F. W. Campbell and D. G. Green, ''Monocular versus binocular visual acuity,'' Nature (London) **208**, 191–192 (1965).
[12] C. W. Tyler, ''Cyclopean vision,'' in *Binocular Vision*, edited by D. Regan (CRC, Boca Raton, FL, 1991).
[13] M. M. Goodsitt, H.-P. Chan, and L. Hadjiiski, ''Stereomammography: Evaluation of depth perception using a virtual 3D cursor,'' Med. Phys. **27**, 1305–1310 (2000).
[14] 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).
[15] G. Cohen, ''Contrast-detail analysis of imaging systems: caveats and kudos,'' in *Recent Developments in Digital Imaging*, edited by K. Doi, L. Lanzl, and P.-J. P. Lin (American Institute of Physics, New York, 1985), pp. 141–159.
[16] C. Braccini, G. Gambardella, and G. Suetta, ''A noise masking experiment in grating perception at threshold: The implications on binocular summation,'' Vision Res. **20**, 373–376 (1980).
[17] S. Pardhan and D. Rose, ''Binocular and monocular detection of Gabor patches in two-dimensional noise,'' Perception **28**, 203–215 (1999).
[18] D. L. Lee, L. K. Cheung, L. S. Jeromin, E. F. Palecki, and B. Rodricks, ''Radiographic imaging characteristics of a direct conversion detector using selenium and thin film transistor array,'' Proc. SPIE **3032**, 88–96 (1997).
[19] T. Aach and V. Metzler, ''Defect interpolation in digital radiography—how object-oriented transform coding helps,'' Proc. SPIE **4322**, 824–835 (2001).
[20] X. Tang, R. Ning, R. Yu, and D. Conover, ''Cone beam volume CT image artifacts caused by defective cells in x-ray flat panel imagers and the artifact removal using a wavelet-analysis-based algorithm,'' Med. Phys. **28**, 812–825 (2001).
[21] *DICOM Standard PS-3.14-2000: Gray-scale Standard Display Function* (National Electrical Manufacturers Association, Rosslyn, VA, 2000).
[22] P. G. Barten, ''Physical model for the contrast sensitivity of the human eye,'' Proc. SPIE **1666**, 57–72 (1992).
[23] P. R. Bakic, M. Albert, and A. D. A. Maidment, ''Dose requirements in stereoradiography,'' Proc. SPIE **4682**, 126–137 (2002).
[24] D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics* (Wiley, New York, 1966), Sec. 2.6, p. 47.
[25] F. O. Bochud, C. K. Abbey, and M. P. Eckstein, ''Statistical texture synthesis of mammographic images with clustered lumpy backgrounds,'' Opt. Express **4**, 33–43 (1999).
[26] P. R. Bakic, M. Albert, D. Brzakovic, and A. D. A. Maidment, ''Mammogram synthesis using a 3D simulation. II. Evaluation of synthetic mammogram texture'' Med. Phys. **29**, 2140–2151 (2002).
[27] P. R. Bakic, M. Albert, and A. D. A. Maidment, ''2-AFC observer study of digital stereomammography,'' Proc. SPIE **5034**, 1–9 (2003).