Digitization requirements in mammography: Effects on computer-aided detection of microcalcifications

Heang-Ping Chan, Loren T. Niklason,[a] Debra M. Ikeda,[b] Kwok Leung Lam,[c] and Dorit D. Adler
Department of Radiology, University of Michigan, Ann Arbor, Michigan 48109-0326

(Received 5 July 1993; accepted for publication 18 April 1994)

We have developed a computerized method for detection of microcalcifications on digitized mammograms. The program has achieved an accuracy that can detect subtle microcalcifications which may potentially be missed by radiologists. In this study, we evaluated the dependence of the detection accuracy on the pixel size and pixel depth of the digitized mammograms. The mammograms were digitized with a laser film scanner at a pixel size of 0.035 mm $\times$0.035 mm and 12-bit gray levels. Digitization with larger pixel sizes or fewer number of bits was simulated by averaging adjacent pixels or by eliminating the least significant bits, respectively. The SNR enhancement filter and the signal-extraction criteria in the computer program were adjusted to maximize the accuracy of signal detection for each pixel size. The overall detection accuracy was compared using the free response receiver operating characteristic curves. The results indicate that the detection accuracy decreases significantly as the pixel size increases from 0.035 mm $\times$ 0.035 mm to 0.07 mm $\times$ 0.07 mm ($P<0.007$) and from 0.07 mm $\times$ 0.07 mm to 0.105 mm $\times$ 0.105 mm ($P<0.002$). The detection accuracy is essentially independent of pixel depth from 12 to 9 bits and decreases significantly ($P<0.003$) from 9 to 8 bits; a rapid decrease is observed as the pixel depth decreases further from 8 to 7 bits ($P<0.03$) or from 7 to 6 bits ($P<0.02$). Although the results are expected to depend on the noise and spatial resolution properties of the x-ray screen-film system and the film digitizer, as well as on the image processing methods, this study demonstrates that very high spatial resolution (a pixel size of 0.035 mm $\times$ 0.035 mm or smaller) will be required for digital mammographic systems in order to capture the very subtle microcalcifications detectable on screen-film mammograms.

Key words: mammography, digital, computer-aided diagnosis, microcalcifications, spatial resolution, gray-level resolution

I. INTRODUCTION

X-ray mammography has been proven to be the most effective method in detection of minimal breast cancers and this improves the chance of survival for breast cancer patients.[1–3] However, it has also been reported that radiologists detect only about 70%–90% of the breast cancers that are visible on retrospective review of the mammograms.[2,4–9] In our previous receiver operating characteristic (ROC) study, we found that radiologists' detection accuracy of mammographic microcalcifications improved significantly ($P<0.001$) when a computer-aided diagnosis (CAD) scheme assisted in searching the images for clusters.[10] CAD thus has the potential to further improve the efficacy of mammography for detection of breast cancers.

Currently, screen-film mammography is the only imaging modality which can provide the high contrast and spatial resolutions required for detection of subtle clustered microcalcifications. It is important to preserve this information when the image, recorded on film or on other recording devices, is digitized and processed for human or computer detection. In a previous ROC study, we found that the detection of subtle microcalcifications by radiologists is adversely affected when mammograms are digitized at a pixel size of 0.1 mm $\times$ 0.1 mm.[11] However, the digitization requirements such as pixel size or pixel depth for mammographic imaging are still unknown, primarily because high-resolution digital systems with pixel sizes smaller than 0.1 mm $\times$ 0.1 mm are not commonly available.

A number of research groups have been developing computerized methods for detection of microcalcifications on digitized mammographs.[12–25] The images were generally obtained from digitizing mammograms recorded with conventional screen-film systems by using laser scanners, video cameras, or charged-coupled device (CCD) cameras at various resolutions. Since the detectability of microcalcifications depends not only on the subtlety of the signals on the original mammograms but also on the quality of the digitized images, it is important to evaluate the dependence of their detectability on the spatial and gray-level resolutions of the digitization process.

In this study, we evaluated the dependence of the detection accuracy of microcalcifications on pixel size and pixel depth of the digitized mammograms.[26] The mammograms were digitized with a high-resolution laser film scanner. The results will provide information on the digitization requirements of mammograms in CAD applications.

II. MATERIALS AND METHODS

A. Acquisition of digital mammograms

In this study, we investigated the effects of digitization pixel size and pixel depth on computerized detection of microcalcifications. Because microcalcifications are the most

 0094-2405/94/21(7)/1203/9/$1.20

subtle feature on mammograms that requires both high spatial resolution and high contrast resolution for detection, the resolution requirements for microcalcifications would represent an upper limit for digitization of mammograms. Twenty-five mammograms were selected from patient files by experienced radiologists (D.M.I., D.D.A.). The mammograms were obtained with a dedicated mammographic system with a 0.3-mm focal spot, molybdenum (Mo) anode, and 0.03-mm Mo filter, and a 5:1 reciprocating grid. Kodak Min R/MRE mammographic screen/film system using extended cycle processing was employed as the image receptor. Each mammogram contained a cluster of subtle microcalcifications which was subsequently removed by surgical biopsy and verified by specimen radiographs. The mammograms were digitized with a laser film scanner (LUMISYS DS-1000) at a pixel size of 0.035 mm × 0.035 mm and 12-bit gray levels. The laser scanner was calibrated such that the optical density unit (OD) was linearly related to pixel values in the range 0.1–2.8 OD at 0.001 OD/pixel value. The slope of the calibration curve outside this range decreased slightly. To facilitate image processing and display, the pixel values were linearly converted such that low optical density was represented by high pixel values. Mammograms digitized at large pixel sizes or fewer number of bits were digitally simulated by averaging adjacent pixels or truncating the least significant bits, respectively, rather than obtained by redigitization with the laser scanner.

Because of the large memory space and the long computer time required for processing the entire mammogram, a region of interest (ROI) with 840×1680 pixels (at a pixel size of 0.035 mm) which contained the cluster of microcalcifications was selected for this study. The ROI corresponded to an area of 17.29 cm^2 on a mammogram. The ROI was displayed on a workstation at full resolution and the coordinates of the individual microcalcifications in the ROI were identified by comparison with the original film mammogram using a magnifier. These locations were used as the "true" locations for scoring correct detection of microcalcifications on this image by the computer. No attempt was made to compare these true locations with specimen radiographs because of the very different appearance and relative spatial locations of microcalcifications on the mammogram and on the specimen radiograph due to the different projection and image quality. A total of 293 microcalcifications were identified in the 25 ROI images.

B. Computerized detection of microcalcifications

Our current computerized detection program consists of three major steps: preprocessing, segmentation, and classification. In the preprocessing step, the signal-to-noise ratio (SNR) of the microcalcifications is enhanced with a difference-image technique.[10,12,13] A signal-enhancement filter is used to enhance the signal and smooth the random noise. A signal-suppression filter is used to remove the signal and again smooth the random noise. The two images are then subtracted to produce a difference image in which the low-frequency structured background is removed and the high-frequency noise is suppressed. In this study, both the signal-enhancement filter and the signal-suppression filter are linear filters so that their difference is used as a single bandpass filter, referred to as the SNR enhancement filter in the discussion below. For segmentation, an automated locally adaptive thresholding procedure is devised.[12,13] The histogram of the difference image is determined and regions of pixels greater than a threshold pixel value are determined by global thresholding in combination with a region growing technique. The procedure is performed iteratively until the number of regions obtained falls within a range which is defined by inputting a maximum number and a minimum number. At each of these potential signal sites, a local region of $n \times n$ pixels, centered at the centroid of the potential signal, is used to estimate the local root-mean-square (rms) noise. In order to reduce the error caused by the potential signal, the central $l \times m$ pixels which may contain a signal are excluded from this estimation. A local gray-level threshold is calculated as the product of the local rms value and an input SNR threshold, which is a global value applied to the entire breast image. A second region growing is then performed to determine the connected pixels above the local gray-level threshold. The centroid location, size, maximum contrast, SNR, and shape of each signal are extracted at this step. Some of these features are used for detection as described in this paper, while others are used for further analysis such as classification of malignant and benign microcalcifications described elsewhere.[27]

For signal and nonsignal discrimination, three criteria are used in this study. A size threshold is set to exclude signals below a certain size which are likely to be noise. A contrast threshold is also set to exclude potential signals which have a contrast higher than an input number of standard deviations above the average contrast of all potential signals found with local thresholding. This criterion excludes the very high-contrast signals that are likely to be large benign calcifications and artifacts. A regional clustering procedure is then applied to the remaining signals; a signal is called positive if the number of signals found within a neighborhood of an input diameter around that signal is greater than an input threshold number. The remaining signals which are not found to be in the neighborhood of any potential clusters will be considered isolated noise points and excluded. This clustering criterion is useful for reducing false positives because true microcalcifications of clinical interest always appear in clusters on mammograms. The clustering criterion is therefore used for detection of either individual microcalcifications or clusters.

After screening with the three signal-extraction criteria, the individual signals that pass the tests are compared with the "truth" file of the input image and the number of true-positive (TP) signals and the false-positive (FP) signals are scored. A detected signal is considered to be TP if its spatial location is within a distance of about 0.35 mm (e.g., 10 pixels at a pixel size of 0.035 mm) from the location of a microcalcification in the truth file. Once a microcalcification in the truth file is identified with a detected signal, it will not be used again for comparison with other detected signals. A detected signal that cannot be identified with any microcalcifications in the truth file will be considered to be an FP. A microcalcification in the truth file that cannot be identified

with any detected signals will be a false negative (FN). The fraction of TP and the number of FP signals detected on each image for a given input SNR threshold are then averaged over the entire set of images used and the mean TP and mean number of FP per image for that input SNR threshold are determined. In this study, because the input image is an ROI on a mammogram, the mean number of FP per "image" (ROI) is converted (divided by 17.29 cm^2) to the mean number of FP per cm^2 which will then be relatively independent of input image size. The trade-off between the TP and FP detection rates of the computer program when the input SNR threshold varies is evaluated by free-response receiver operating characteristic (FROC) analysis.[28] A low SNR threshold corresponds to a lax criterion with a large number of FP signals. A high SNR threshold corresponds to a stringent criterion with a small number of FP signals and a loss in TP signals. The detection accuracy of the computer program under different conditions can then be assessed by comparison of the FROC curves.

C. Optimization of parameters

The same computerized detection procedure was applied to images of various pixel sizes and bit depths. For each pixel size, we tested various combinations of input parameters until the maximum detection curve was obtained. The filter parameters determined the maximum detection accuracy that could be achieved as discussed in Sec. III. The filter kernel was therefore carefully chosen to obtain the highest detection at each pixel size.

The effect of bit depth was evaluated at the smallest pixel size, which was found to provide the highest detection accuracy as described below. Because of the fixed spatial resolution, the same filter kernel found to be optimal in the pixel size study was used for images of various bit depths. However, the SNR threshold and contrast threshold were adjusted to accommodate the decreasing contrast sensitivity at decreasing bit depths.

III. RESULTS

The optimization procedures performed at each pixel size are demonstrated with results for the 0.035-mm pixel size. Three examples of the SNR enhancement filters studied are shown in Fig. 1. Spatial filtering is performed in spatial domain as a convolution process. However, the optical transfer functions (OTFs) of the filters are plotted so that their bandpass characteristic can be compared. The filters suppress the low-frequency structured background as well as the high-frequency noise. Note that the spatial-frequency contents of a digital image above the Nyquist frequency are also modulated by the filter OTFs. The FROC curves obtained from these filters are shown in Fig. 2. There is a noticeable change in detection accuracy as the filter OTF varies. It appears that the frequency band peaked at about 1.8 cycles/mm is important to provide differentiation between signal and background structures (see also Fig. 5). However, the bandwidth affects the overall detection accuracy. Filter 1 provides the highest detection among the filters studied for the images of 0.035-mm pixel size. If higher-frequency components are enhanced such as that obtained with filter 2, there is a trade-off in the SNR due to increased high-frequency noises, resulting in detection of a larger number of FPs. On the other hand, if high frequencies are oversuppressed such as that obtained with filter 3, the loss of the important high-frequency signal information precludes detection of the very subtle microcalcifications. The sensitivity with filter 3 cannot reach the same level as that with filter 1 even at low SNR thresholds that allow a large number of FPs.

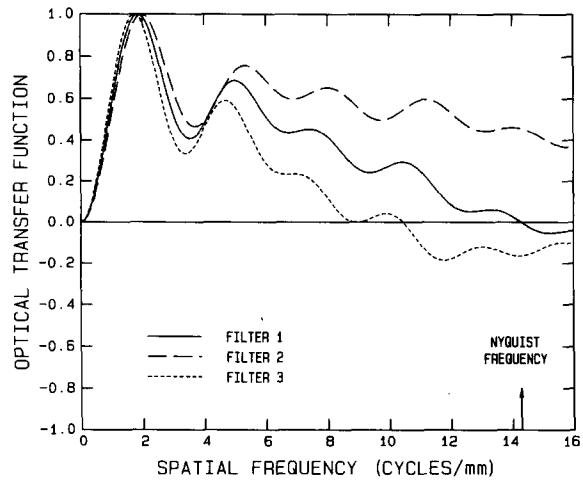

FIG. 1. Examples of SNR enhancement filters tested for processing the images at a pixel size of 0.035 mm × 0.035 mm. The signal-enhancement filter kernel of filter 1 is 2×2 pixels, of filter 2 is 1×1 pixel, and of filter 3 is 3×3 pixels.

Filter 1 was chosen as our SNR enhancement filter. The dependence of the detection accuracy on the size threshold is shown in Fig. 3. It can be seen that for a given number of FPs the TP rate increases as the size threshold increases from 4 to 8 pixels. When the size threshold increases beyond 8 pixels, the FROC curve at larger FP rates begins to fall due to the loss in small-size microcalcifications. If the size range

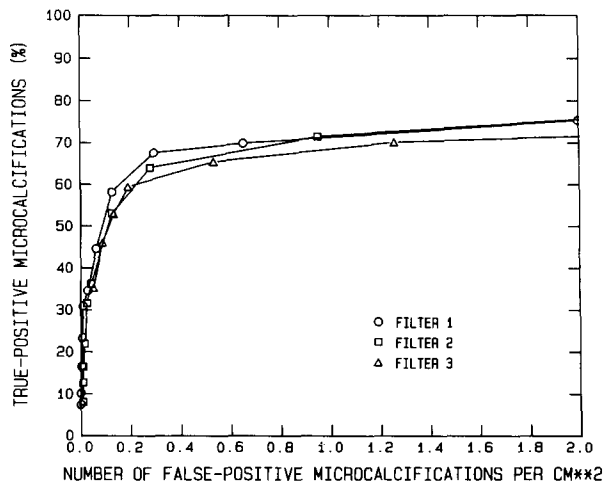

FIG. 2. FROC curves obtained by detection on the images of 0.035 mm × 0.035 mm pixel size processed by the three filters shown in Fig. 1. The data points on the FROC curves in this study were obtained by varying the input SNR threshold in the local gray-level thresholding process.

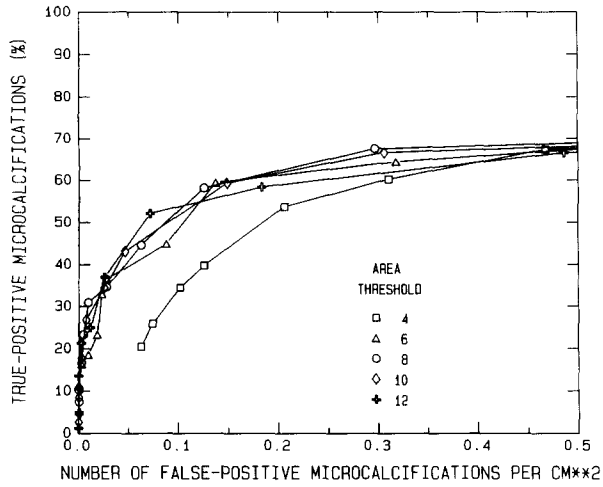

FIG. 3. Dependence of detection accuracy on the size threshold of microcalcifications, chosen as an input parameter in the computer program. The pixel size of the digitized images is 0.035 mm × 0.035 mm.

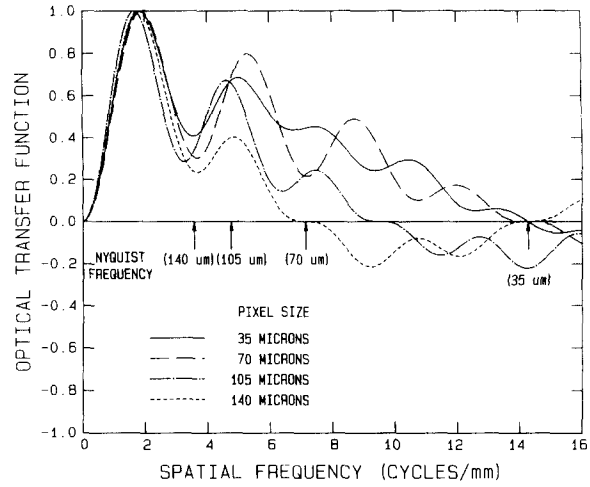

FIG. 5. SNR enhancement filters for processing images of different pixel sizes used in this study. The Nyquist frequency of each pixel size is indicated. The signal-enhancement filter kernels for the four pixel sizes 0.035, 0.070, 0.105, and 0.140 mm are 2×2 pixels, 1×1 pixel, 1×1 pixel, and 1×1 pixel, respectively.

is kept constant, the detection accuracy depends on the minimum number of signals required to form a cluster in the clustering procedure as shown in Fig. 4. When the criteria is stringent, i.e., more signals are required to be in the neighborhood, the number of FP signals appears to decrease because scattered FP signals are less likely to pass the test. However, when the criteria becomes too stringent, the very subtle cluster for which relatively fewer individual microcalcifications can be detected may be excluded due to insufficient number of detected signals in the neighborhood. For the set of test images used in this study, the FROC curves are found to be highest if a minimum of four detected signals are required in the clustering criterion.

The filter OTFs for the SNR enhancement filters chosen for processing images at various pixel sizes are shown in Fig. 5 and the threshold and clustering criteria are listed in Table I. The dependence of the detection accuracy for individual microcalcifications on pixel size is shown in Fig. 6. It can be seen that the detection accuracy decreases as the pixel size increases from 0.035 to 0.14 mm except in the range when the TP and FP detection rates are very low and the curves become overlapping. The curves are plotted for a clustering criterion of a minimum of four detected signals in a cluster. Similar results are obtained with the criterion of a minimum of three detected signals in a cluster except that the TP rate for each curve will be 1% or 2% lower. Although the difference in the detection curves between different pixel sizes amounts to only a few percent, the systematic difference indicates that some of the very subtle microcalcifications may be missed if a pixel size greater than 0.035 mm is used for digitization. Because the very subtle microcalcifications are most likely to be missed by human observers, a CAD system should detect these microcalcifications accurately in order to be clinically useful.

We attempted to apply the FROCFIT curve fitting program[29] to our FROC data and thus to estimate the area under the AFROC (alternative FROC) curve. However, the resulting fits were very poor, indicating that the assumptions made in the development of the FROCFIT program might not be satisfied in our situations. We also attempted to use

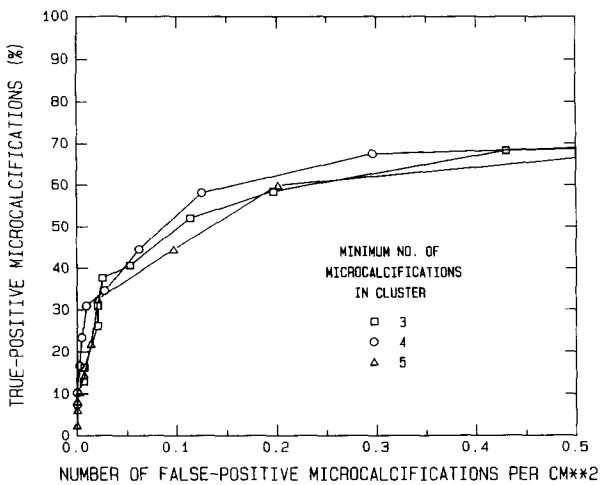

FIG. 4. Dependence of detection accuracy on the minimum number of signals within a 1-cm-diam neighborhood for each detected signal. The number is chosen as an input parameter in the computer program. The pixel size of the digitized images is 0.035 mm × 0.035 mm.

TABLE I. Input parameters used for detection of microcalcifications on mammograms of various pixel sizes.

Pixel size (mm)	Local thresholding region size (pixels)	Upper and lower bounds on size of microcalcification (pixels)	Cluster diameter (pixels)	Minimum No. of microcalcifications in cluster
0.035	151	8, 400	301	4
0.070	75	2, 200	151	4
0.105	51	1, 80	101	4
0.140	37	1, 60	75	4

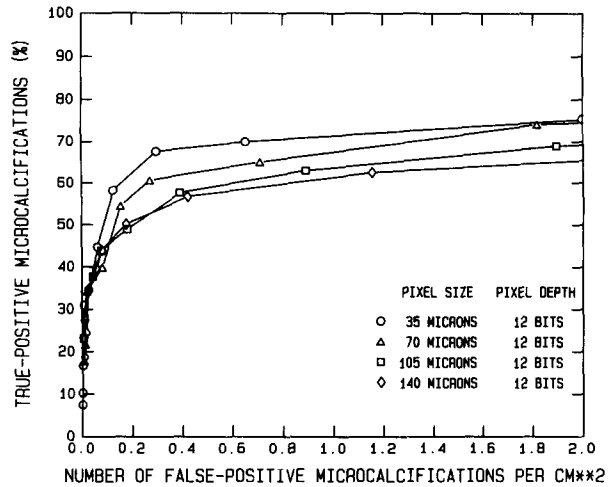

FIG. 6. Dependence of detection accuracy on pixel size. The bit depth of all images is 12 bit.

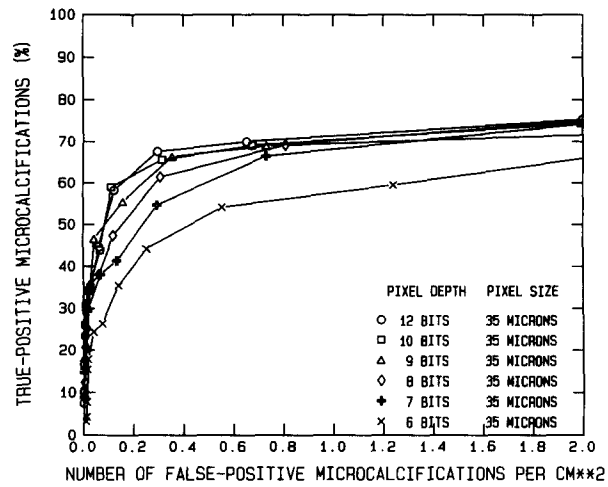

FIG. 7. Dependence of detection accuracy on pixel depth. The pixel size of all images is 0.035 mm × 0.035 mm.

AFROC analysis[30] for our detection data and then applied the CLABROC program[31,32] to the pair of correlated AFROC curves. This approach was again found to be unsuitable because the individual FP signals detected in an image were not independent in this detection task. Before an appropriate model and statistical analysis could be developed, we applied a paired t test to estimate the statistical significance of the differences between the detection accuracy obtained from each pair of conditions. The number of TP signals detected under the first condition for an image at an SNR threshold that yielded a given mean number of FP signals was paired with the corresponding TP signals detected under the second condition for the same image at an SNR threshold that yielded a similar mean FP. The t test was performed for TP pairs in the range of 0.1–0.8 FP signals per cm^2. This range of FPs was chosen because it is likely to be a typical range in which a CAD program will be operated; the ranges of very high FPs or low TPs will not be used as operating points in practical situations. The results of the paired t test indicated that the difference in the TP signals between pixel sizes of 0.035 and 0.07 mm was statistically significant at $P<0.007$, and between pixel sizes of 0.07 and 0.105 mm was statistically significant at $P<0.002$. The difference between pixel sizes of 0.105 and 0.14 mm fell short of statistical significance with $P=0.794$.

To study the dependence of the detection accuracy on pixel depth of digitization the pixel size was fixed at 0.035 mm and the number of bits in a pixel was reduced from 12 to 6 bits by eliminating the least significant bits. The optimal filter found for the test images at a pixel size of 0.035 mm was used for SNR enhancement. Figure 7 shows a comparison of the FROC curves for the pixel depths from 12 to 6 bits when the minimum size of signals detected was set at 8 pixels and the minimum number of signals in the clustering criterion was set to be four as listed in Table I. The detectability of microcalcifications is essentially independent of pixel depth from 12 to 9 bits, with $P=0.35$ between 12 and 10 bits, $P=0.87$ between 10 and 9 bits, and $P=0.77$ between 12 and 9 bits. The detectability decreases significantly ($P<0.003$) as the pixel depth reduces from 9 to 8 bits. The detectability continues to decrease significantly when the pixel depth reduces from 8 to 7 bits ($P<0.03$) and from 7 to 6 bits ($P<0.02$). Similar trends are observed if the minimum size of signals was chosen to be 4–10 pixels and the minimum number of signals in the clustering criterion was set to be three or four. Therefore, the relative dependence of detection accuracy on pixel depth appears to be independent of the parameters chosen in the detection program.

IV. DISCUSSION

In this study, we compared the detectability of individual microcalcifications rather than microcalcification clusters. Since a cluster can be detected even if some of the individual microcalcifications in the cluster are missed, the detectability of clusters is not a sensitive indicator for evaluation of the effects of pixel size and pixel depth. For the set of test images used in this study, similar, but less obvious, trends as those shown in Figs. 6 and 7 were observed if detection of clusters instead of individual microcalcifications were analyzed. The decrease in the TP detection rate of the individual microcalcifications has two implications. First, if a cluster is composed mainly of subtle microcalcifications, the cluster may be missed if the image is digitized at a large pixel size or fewer number of bits per pixel. Second, if a CAD program is to be used in the future for determination of interval change in the number of microcalcifications in a cluster, the increase in the number of subtle microcalcifications may not be detected.

The effects of digitization parameters on detection accuracy depend strongly on the test images used. If the test images contain mainly obvious microcalcifications, it can be expected that the differences in the TP rates among different pixel sizes and pixel depths will decrease. Conversely, if the microcalcifications are more subtle than those used in our study, the dependence of their detectability on the digitization parameters will be stronger. The significance of the results of this study lies not in the absolute detection rates or

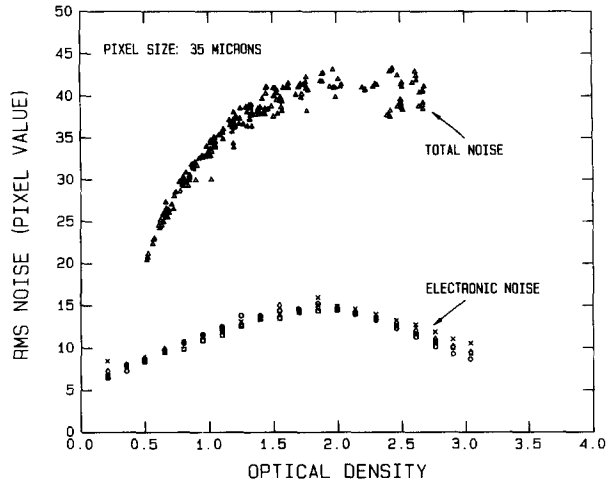

FIG. 8. The rms noise of the digitized images. The triangles indicate the measured total noise including quantum noise, screen-film noise of a Kodak Min R/MRE system, electronic noise, and digitization noise. The circles, squares, and diamonds indicate three measurements of the electronic noise of the laser scanner using a two-scan method as described in the text; the crosses indicate electronic noise measured with a one-scan method.

the amount of their differences. The fact that the detectability of some microcalcifications that are visible on mammograms will be degraded if a pixel size greater than 0.035 mm is used for digitization indicates that very high spatial resolution is required for digital mammograms in CAD application. Due to the limited resolution of our digitizer, we cannot study the dependence of the detectability on pixel size below 0.035 mm. Although computer vision is different from human vision and our results may not be directly applicable to digital mammograms for human observers, careful study of spatial resolution requirements below a pixel size of 0.050 mm for digital mammography systems is warranted.

The effect of pixel depth on detection accuracy depends on the noise of the imaging system. In our case, the noise includes quantum noise, screen-film system noise, digitization noise, and electronic noise of the digitizer. To estimate the total noise, we made noise film samples using a Kodak Min R/MRE mammographic screen/film system. The films were exposed under a 4-cm-thick Lucite phantom with a Mo/Mo spectrum at 28 kVp, which was similar to the techniques used in our department for mammograms of a patient with an average breast size. The optical densities of the films were adjusted by varying the mAs so that samples from 0.5 to 2.7 OD were obtained. The films were processed with extended cycle and were digitized at a pixel size of 0.035 mm and a pixel depth of 12 bits. The rms noises of the digitized images were then calculated in regions of 60×150 pixels, with the long dimension perpendicular to the anode-cathode direction. The size of the regions was chosen to be small enough so that the film density within a region was practically uniform. The rms noise was plotted as a function of the OD of each region as shown in Fig. 8. It can be seen that the total rms noise increases with optical density from about 20 to 42 pixel values (i.e., from 0.02 to 0.042 OD).

To estimate digitization noise and electronic noise, we used a Kodak neutral density step tablet which had density steps from about 0.1 to 3.1 OD. The graininess of the density steps was negligibly low. The tablet was digitized with the laser film digitizer with a pixel size of 0.035 mm and 12 bits. The rms noise at each density step was calculated within a 150×150 pixel region. The measurement was repeated several times on different days. To ensure that the rms variation was caused mainly by noise rather than by nonuniformity in density over the step, we also made two consecutive digitizations without moving the tablet. Since the digitizer is a flat-bed model with a moving glass stage and the starting point of digitization can be reproduced within a fraction of a pixel, the two consecutive digitized tablet images could be registered accurately. By subtracting the two images we obtained a difference image in which the background nonuniformity was removed. We again calculated the rms noise at each density step. The rms noise of a single scan was obtained by dividing the rms noise of the subtracted image by a factor of $\sqrt{2}$. The digitization and electronic noises were estimated to be about 8–15 pixel values from either a single-scan or a two-scan method as shown in Fig. 8. Because digitization noise was negligibly small at 12 bits, the measured noise was predominantly electronic noise.

If it was assumed that electronic noise was independent of the image noise (i.e., quantum and screen-film noises), the latter rms noise could be estimated as the square root of the difference of the squared total noise and the squared electronic noise. The image rms noise was estimated to be about 18–39 pixel values (i.e., 0.018–0.039 OD), depending on film density. The image rms noise thus was the dominant noise source in the digital mammograms under the exposure and film processing conditions used.

The result that the detectability of the microcalcifications remained almost constant when the pixel depth was reduced from 12 to 9 bits and decreased only slightly when the pixel depth was further reduced to 8 bits could be explained by the fact that the last 4 bits of information contained mainly noise. If a lower-noise system is developed for recording digital mammograms, it is expected that more bits may be useful in preserving the detectability of the microcalcifications.

Many of the laser scanner systems are designed to operate at an "oversampling" condition, i.e, the laser beam spot is larger than the sampling distance. This design degrades the resolution of the system. However, it also reduces noise and aliasing errors because the aperture response acts as an anti-aliasing filter that suppresses some high-frequency information as well as noise in the image. The sampling distance of our laser scanner at a nominal sampling distance of 0.035 mm was measured many times over a period of more than one year by scanning a calibration strip with marked scales and was found to be 0.034±0.0003 mm. We have not measured the modulation transfer function (MTF) of the beam spot of our laser scanner. The square wave response of the system at a pixel size of 0.035 mm was measured to have a response of 0.5 at about 6 line pairs/mm (lp/mm) and a response of 0.2 at about 10 lp/mm [G. Keyes (private communication)]. If the response is approximated by a Gaussian function, it can be estimated that the full width at half-maximum of an effective beam spot is about 0.07 mm. A

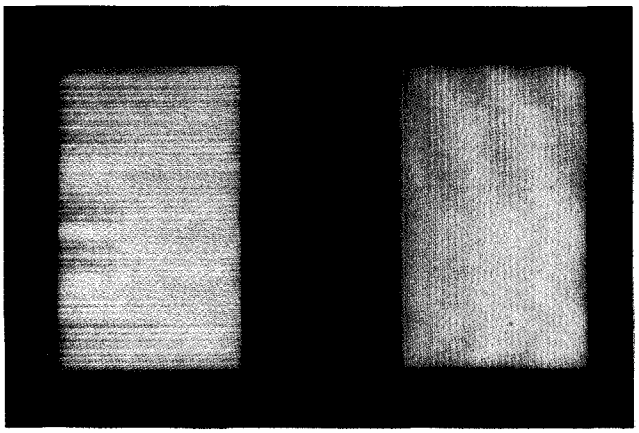

FIG. 9. Digitized images of line-pair patterns of 14.3 lp/mm. The line pattern on the left-hand side is parallel to the scan lines and that on the right-hand side is perpendicular to the scan lines. The pixel size is 0.035 mm × 0.035 mm and the pixel depth is 12 bits. The image was displayed with windowing technique to enhance the contrast.

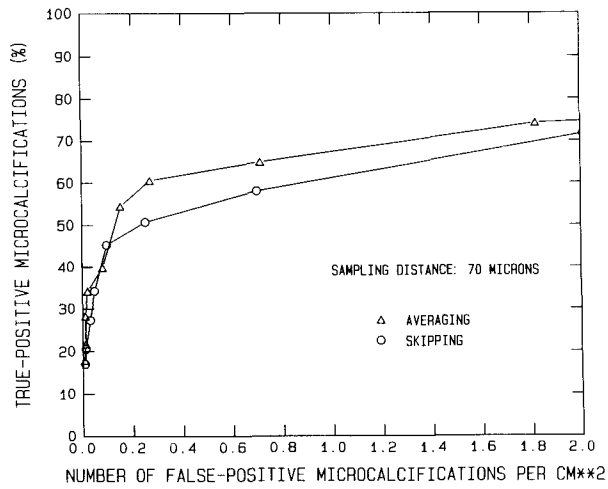

FIG. 10. Dependence of detection accuracy on beam spot size at a sampling distance of 0.07 mm. The images for the upper curve were obtained by averaging every 2×2 neighboring pixels, the images for the lower FROC curve were obtained by skipping every other pixels in the 0.035 mm images.

line-pair resolution pattern of 14.3 lp/mm digitized with the laser scanner at a pixel size of 0.035 mm is shown in Fig. 9. The line pairs can be resolved although some blurred regions due to aliasing can be seen. We have also exposed a Leeds mammographic phantom (1-cm-thick Lucite plate embedded with test objects including line-pair patterns) in contact with a direct MRE film. The phantom image was again digitized with the laser scanner at a pixel size of 0.035 mm. The line-pair patterns parallel to the scan lines can be resolved up to the 16.6 lp/mm group while the line-pair patterns perpendicular to the scan lines can be resolved up to the 14.3 lp/mm group.

According to Figs. 1 and 2, the detection accuracy of microcalcifications on images processed with an SNR enhancement filter which smooths the image with a 2×2 kernel (filter 1) is higher than that on images processed with a filter which does not provide smoothing (filter 2). The comparison shows that smoothing in addition to that by the scanning laser beam spot is needed in order to optimize the SNR of the microcalcifications. If the beam spot were smaller than 0.07 mm a smoothing kernel larger than 2×2 pixels might be required to provide sufficient noise suppression. This implies that the beam spot size of our digitizer may not be the limiting factor in the detection accuracy for the purpose of this study.

To further study the effect of oversampling in comparison with the case that the sampling beam spot is about the same as the sampling distance, we simulated images digitized at a 0.07 mm sampling distance with a 0.07 mm beam spot by skipping every other pixel in both directions of the 0.035 mm images. Detection of the microcalcifications was then performed with the optimal parameters for images of 0.07-mm pixel size as shown in Fig. 5 and Table I. The FROC curve is shown in Fig. 10. Also shown is the FROC curve for detection on the images of 0.07-mm pixel size obtained by averaging four adjacent pixels as described above (Fig. 6), which corresponds to an oversampling situation. It can be seen that images with oversampling provide significantly higher ($P<0.0002$) detection accuracy than those without oversampling. We can also make further comparison between these curves and the FROC curve for images of 0.035-mm pixel size shown in Fig. 6. The latter FROC curve was obtained for images with a 0.07-mm beam spot, a 0.035-mm sampling distance, and a 2×2 signal-enhancement filter. The 2×2 signal-enhancement filter acted like the averaging operation used in simulating the images of 0.07-mm pixel size with oversampling. The only difference between these two sets of images was then the sampling distance. The result indicates that a sampling distance of 0.035 mm, instead of 0.07 mm, improves the detectability significantly ($P<0.007$). The two comparisons demonstrate that proper amount of oversampling will improve SNR by suppressing noise and aliasing errors.

In this study, we digitized images at a pixel size of 0.035 mm and 12 bits and simulated images digitized at larger pixel sizes by averaging adjacent pixels. This approach eliminates the variation in detectability caused by variation in alignment of the digitization matrix with the image or variations in the digitizer such as its contrast calibration and electronic noise. However, images obtained with the simulation approach may differ from those obtained with direct digitization in two aspects. First, in the simulated images, the electronic noise due to the digitizer is reduced by digital averaging, whereas in images obtained with direct digitization, the electronic noise is not smoothed by the OTF of the digitizer aperture (beam spot). Second, for the simulated images, the relationship between the digital and the input analog image spectra in the spatial frequency domain can be described by:

$$\text{digitized image spectrum} = \{\{[\text{image spectrum}][\text{aperture OTF}]_S\} ** [\text{sampling comb function}]_S\} \; [\text{aperture OTF}]_{L'}\} ** [\text{sampling comb function}]_L ,$$

whereas for the images obtained by direct digitization with a large pixel size, the relationship is given by:

digitized image spectrum=
([image spectrum][aperture OTF]$_L$)**[sampling comb function]$_L$,

where the subscript s denotes the small aperture or the small sampling distance, L' the simulated large aperture, L the large aperture or the large sampling distance of direct digitization, and ** denotes the convolution operation. The simulated image will approach the image obtained with direct digitization if the effects of the two factors described above are negligible. This is true for the first factor if the electronic noise is negligible compared with image noise, and for the second factor, if aliasing errors are negligible below the Nyquist frequencies during the sampling steps.

From Fig. 8 and the discussion above, it is found that the total noise is dominated by image noise. The effect of the first factor is therefore likely to be negligible. It is difficult to evaluate the effect of the second factor. The input image spectrum contains high-frequency components because a mammographic imaging system can record information up to 16–20 line pairs/mm. However, since the full width at half-maximum of the beam spot of the laser film digitizer is about 0.07 mm at the nominal size of 0.035 mm as described above, the aliasing errors may be quite low at a sampling distance of 0.035 mm. Aliasing errors in the subsequent sampling will also be limited. The effect on the overall detectability of microcalcifications by the computer caused by the differences in the simulated and directly digitized images would likely be small.

V. CONCLUSIONS

In this study, we found that for computer detection digital mammograms have to be digitized with a pixel size as small as 0.035 mm × 0.035 mm and a pixel depth of at least 9 bits, in order to preserve the detectability of subtle microcalcifications. The detectability of microcalcification clusters may be less sensitive to the spatial and gray-level resolutions than that of individual microcalcifications because a cluster often contains microcalcifications of various sizes, densities, and contrasts. Some of the larger microcalcifications may be detected and the cluster found even if the very subtle ones are missed due to insufficient resolutions. However, it can be expected that the detection rate of some very subtle clusters which contains only subtle microcalcifications may still be affected if mammograms are digitized at lower resolutions. The clinical significance of the miss rate of the very subtle microcalcifications remains to be studied.

Our results are expected to depend on the noise and resolution properties of the image recording and digitization systems, and they may not be directly applicable to images used for human visual detection. However, this study demonstrates that very high spatial resolution and SNR performance or dynamic range will be required for digital mammographic systems. The trade-offs between subtle diagnostic information and technological requirements on acquisition, transmission, archiving, and processing of high-resolution images will have to be carefully considered for such systems.

ACKNOWLEDGMENTS

This work is supported by USPHS Grant No. CA 48129 and a Faculty Research Award (FRA 334) from the American Cancer Society. The authors are grateful to Charles E. Metz, Ph.D., for helpful discussion on FROC and AFROC analyses and for providing the CLABROC programs, to Dev Chakraborty, Ph.D., for providing the FROCFIT program, and to Cynthia Sims-Holmes and Diane Williams for secretarial assistance.

[a]Current address: Radiological Sciences and Technology, Massachusetts General Hospital, Boston, MA 02114.
[b]Current address: Department of Radiology, Stanford University, Stanford, CA 94305-5105.
[c]Department of Radiation Oncology, University of Michigan, Ann Arbor.

[1]M. Moskowitz, "Breast cancer: Age-specific growth rates and screening strategies," Radiology **161**, 37–41 (1986).
[2]M. Moskowitz, "Benefit and risk," in *Breast Cancer Detection: Mammography and Other Methods in Breast Imaging*, 2nd ed., edited by L. W. Bassett and R. H. Gold (Grune and Stratton, New York, 1987).
[3]H. Seidman, S. K. Gelb, E. Silverberg, N. LaVerda, and J. A. Lubera, "Survival experience in the Breast Cancer Detection Demonstration Project," CA Cancer J Clin. **37**, 258–290 (1987).
[4]J. E. Martin, M. Moskowitz, and J. R. Milbrath, "Breast cancer missed by mammography," Am. J. Radiol. **132**, 737–739 (1979).
[5]L. Kalisher, "Factors influencing false negative rates in xeromammography," Radiology **133**, 297–301 (1979).
[6]C. J. Baines, *et al.* "Sensitivity and specificity of first screen mammography in the Canadian National Breast Screening Study: A preliminary report from five centers," Radiology **160**, 295–298 (1986).
[7]B. J. Hillman *et al.* "Mammogram interpretation by physician assistants," Am. J. Radiol. **149**, 907–911 (1987).
[8]L. W. Bassett, D. H. Bunnell, R. Jahanshahi, R. H. Gold, R. D. Arndt, and J. Linsman, "Breast cancer detection: One versus two views," Radiology **165**, 95–97 (1987).
[9]P. J. Haug, I. M. Tocino, P. D. Clayton, and T. L. Bair, "Automated management of screening and diagnostic mammography," Radiology **164**, 747–752 (1987).
[10]H. P. Chan, K. Doi, C. J. Vyborny, R. A. Schmidt, C. E. Metz, K. L. Lam, T. Ogura, Y. Wu, and H. MacMahon, "Improvement in radiologists' detection of clustered microcalcifications on mammograms. The potential of computer-aided diagnosis," Invest. Radiol. **25**, 1102–1110 (1990).
[11]H. P. Chan, C. J. Vyborny, H. MacMahon, C. E. Metz, K. Doi, and E. A. Sickles, "Digital mammography: ROC studies of the effects of pixel size and unsharp-mask filtering on the detection of subtle microcalcifications," Invest. Radiol. **22**, 581–589 (1987).
[12]H. P. Chan, K. Doi, S. Galhotra, C. J. Vyborny, H. MacMahon, and P. M. Jokich, "Image feature analysis and computer-aided diagnosis in digital radiography. 1. Automated detection of microcalcifications in mammography," Med. Phys. **14**, 538–548 (1987).
[13]H. P. Chan, K. Doi, C. J. Vyborny, K. L. Lam, and R. A. Schmidt, "Computer-aided detection of microcalcifications in mammograms: Methodology and preliminary clinical study," Invest. Radiol. **23**, 664–671 (1988).
[14]B. W. Fam, S. L. Olson, P. F. Winter, and F. J. Scholz, "Algorithm for the detection of fine clustered calcifications on film mammograms," Radiology **169**, 333–337 (1988).
[15]D. H. Davies and D. R. Dance, "Automatic computer detection of clustered calcifications in digital mammograms," Phys. Med. Biol. **35**, 1111–1118 (1990).
[16]L. N. Mascio, J. M. Hernandez, and C. M. Logan,"Automated analysis for microcalcifications in high resolution digital mammograms," Proc. SPIE **1898**, 472–479 (1993).
[17]R. M. Nishikawa *et al.*, "Computer-aided detection and diagnosis of masses and clustered microcalcifications from digital mammograms," Proc. SPIE **1905**, 422–431 (1993).
[18]W. Qian *et al.*, "Tree-structured nonlinear filter and wavelet transform for microcalcification segmentation in mammography," Proc. SPIE **1905**, 509–520 (1993).
[19]D. Brzakovic, P. Brzakovic, and M. Neskovic, "Approach to automated screening of mammograms," Proc. SPIE **1905**, 690–701 (1993).

[20]S. Astley *et al.*, "Automation in mammography: computer vision and human perception," Proc. SPIE **1905**, 716–730 (1993).
[21]I. N. Bankman, W. A. Christens-Barry, D. W. Kim, I. N. Weinberg, O. B. Gatewood, and W. R. Brody, "Automated recognition of microcalcification clusters in mammograms," Proc. SPIE **1905**, 731–739 (1993).
[22]N. Karssemeijer, "Recognition of clustered microcalcifications using a random field model," Proc. SPIE **1905**, 776–786 (1993).
[23]L. Shen, R. M. Rangayyan, and J. E. L. Desautels, "Automatic detection and classification system for calcifications in mammograms," Proc. SPIE **1905**, 799–805 (1993).
[24]A. P. Dhawan, Y. S. Chitre, and M. Moskowitz, "Artificial-neural-network-based classification of mammographic microcalcifications using image structure features," Proc. SPIE **1905**, 820–831 (1993).
[25]K. S. Woods, J. L. Solka, C. E. Priebe, C. C. Doss, K. W. Bowyer, and L. P. Clarke, "Comparative evaluation of pattern recognition techniques for detection of microcalcifications," Proc. SPIE **1905**, 841–852 (1993).
[26]H. P. Chan, L. T. Niklason, D. M. Ikeda, D. D. Adler, L. Lee, and D. Y. Shieu, "Digitization requirements in mammography: Effects on computer-aided diagnosis," World Congress on Medical Physics and Biomedical Engineering, 16th International Conference on Medical and Biological Engineering and 9th International Conference on Medical Physics, Kyoto, Japan, 7–12 July [Med. Biol. Eng. Comp. **29** (Suppl.), 20 (1991)].
[27]H. P. Chan, L. T. Niklason, D. M. Ikeda, and D. D. Adler, "Computer-aided diagnosis in mammography: Detection and characterization of microcalcifications," Med. Phys. **19**, 831 (1992).
[28]P. C. Bunch, J. F. Hamilton, G. K. Sanderson, and A. H. Simmons, "A free response approach to the measurement and characterization of radiographic observer performance," Proc. SPIE **127**, 124–135 (1977).
[29]D. P. Chakraborty, "Maximum likelihood analysis of free-response receiver operating characteristic (FROC) data," Med. Phys. **16**, 561–568 (1989).
[30]D. P. Chakraborty and L. H. L. Winter, "Free-response methodology, Alternate analysis and a new observer-performance experiment," Radiology **174**, 873–881 (1990).
[31]C. E. Metz, P. L. Wang, and H. B. Kronman, "A new approach for testing the significance for differences between ROC curves measured from correlated data," in Information Processing in Medical Imaging, edited by F. Deconinck (Martinus Nijhoff, The Hague, 1984), pp. 432–445.
[32]C. E. Metz, J. H. Shen, and B. A. Herman, "New methods for estimating a binormal ROC curve from continuously distributed test results," Presented at the 1990 Annual Meeting of the American Statistical Association; Anaheim, CA, 7 August 1990 (unpublished).