Image compression in digital mammography: Effects on computerized detection of subtle microcalcifications

Heang-Ping Chan,[a] Shih-Chung B. Lo,[b] Loren T. Niklason,[c] Debra M. Ikeda,[d] and Kwok Leung Lam[e]
Department of Radiology, University of Michigan, Ann Arbor, Michigan 48109

(Received 13 June 1995; resubmitted 14 March 1996; accepted for publication 29 May 1996)

Our previous receiver operating characteristic (ROC) study indicated that the detection accuracy of microcalcifications by radiologists is significantly reduced if mammograms are digitized at 0.1 mm×0.1 mm. Our recent study also showed that detection accuracy by computer decreases as the pixel size increases from 0.035 mm×0.035 mm. It is evident that very large matrix sizes have to be used for digitizing mammograms in order to preserve the information in the image. Efficient compression techniques will be needed to facilitate communication and archiving of digital mammograms. In this study, we evaluated two compression techniques: full frame discrete cosine transform (DCT) with entropy coding and Laplacian pyramid hierarchical coding (LPHC). The dependence of their efficiency on the compression parameters was investigated. The techniques were compared in terms of the trade-off between the bit rate and the detection accuracy of subtle microcalcifications by an automated detection algorithm. The mean-square errors in the reconstructed images were determined and the visual quality of the error images was examined. It was found that with the LPHC method, the highest compression ratio achieved without a significant degradation in the detectability was 3.6:1. The full frame DCT method with entropy coding provided a higher compression efficiency of 9.6:1 at comparable detection accuracy. The mean-square errors did not correlate with the detection accuracy of the microcalcifications. This study demonstrated the importance of determining the quality of the decompressed images by the specific requirements of the task for which the decompressed images are to be used. Further investigation is needed for selection of optimal compression technique for digital mammograms. © *1996 American Association of Physicists in Medicine.*

Key words: mammography, digital, microcalcifications, image compression, computer-aided diagnosis

I. INTRODUCTION

X-ray mammography is the most effective method in detection of early breast cancers.[1] Because of the stringent requirements for imaging of subtle lesions, the image recording systems for mammography have to provide very high spatial resolution and high contrast sensitivity. At present, screen-film systems specially designed for mammography are the only recording medium that can provide the image quality needed. However, because of the advancement in digital imaging technology, digital mammography is becoming a realistic goal. Digital mammography offers the advantages of electronic transmission, consultation, and archiving as well as image enhancement and computer-aided diagnosis.[2] These will potentially make mammography more widely accessible, reduce the cost, and improve the diagnostic accuracy of mammography.

Our previous receiver operating characteristic (ROC) study indicated that the detection accuracy of microcalcifications by radiologists is significantly reduced if mammograms are digitized at 0.1 mm×0.1 mm.[3] Our recent study also showed that detection accuracy by computer decreases as the pixel size increases from 0.035 mm×0.035 mm.[4] It is evident that very high resolution digitization has to be used for mammograms in order to preserve the information in the image. A 18 cm×24 cm mammogram digitized at 0.05 mm ×0.05 mm results in a matrix size of about 4000×5000. A four-view study thus will provide 160 megabytes of digital data. The transmission and archiving of such a large amount of data is therefore one of the important considerations in implementation of digital mammography. An efficient data compression scheme that can reduce the amount of data without degradation of the image quality for human and machine interpretation will alleviate these problems.

Much effort has been devoted to evaluate compression methods for radiological images. Most studies so far applied to digital chest radiography[5–10] because it is the most commonly performed procedure in medical imaging, and some direct digital imaging systems for chest radiography are already available. Recently several investigators have extended their studies to general radiological images, including computed tomography (CT), and magnetic resonance (MR) images based on DCT[11–13] and wavelet-type decomposition methods.[14–16] Some preliminary studies have also been performed for digitized mammographic images.[17–19]

In this study, we explored some of the issues involved in compression of mammographic images for applications in computerized detection of microcalcifications.[20] Because primary digital mammography systems are not yet available,

 0094-2405/96/23(8)/1325/12/$10.00

digital mammograms in this study were obtained by digitization of screen-film mammograms. We selected two image compression techniques, the Laplacian pyramid hierarchical coding (LPHC)[21] and the discrete cosine transform (DCT), with full frame entropy coding (FFEC)[5,6] for processing of mammograms. The LPHC technique was chosen because of its similarity to the difference-image technique that we used for enhancement of microcalcifications in the automated detection algorithm. The DCT technique was commonly used for an irreversible image compression and FFEC was developed to eliminate the block artifacts and improve compression efficiency. The DCT-FFEC technique was further implemented using bit plane splitting to improve the preservation of detailed information.[7] We compared the compression efficiency of these techniques for digitized mammograms. The fidelity of the information in the reconstructed image was evaluated by the detectability of the microcalcifications by an automated computer program. The results were compared with the mean square error (MSE), which was a commonly used indicator of information loss in image compression.

II. MATERIALS AND METHODS

A. Data set of digital mammograms

Twenty-five mammograms were selected from patient files from the Department of Radiology at the University of Michigan. All mammograms were acquired with American College of Radiology accredited machines and recorded with Kodak Min R/Min $R-E$ screen-film systems. Each mammogram contained a cluster of subtle microcalcifications, the presence of which had been verified by biopsy. The mammograms were digitized with a high-resolution laser scanner at a pixel size of 35 μm$\times$35 μm and 12-bit gray levels. The digitizer logarithmically amplified the transmitted light through the film before digitization. The scanner was calibrated such that the gray levels were linearly proportional to optical density (OD) in the range of about 0.1–2.8 OD. The optical density range of the scanner was 0–3.5 OD.

Because of the computational requirement for processing the entire breast image that could be greater than 4000$\times$5000 pixels, we manually extracted an ROI of 1024$\times$1024 pixels, which contained the cluster of microcalcifications from each digitized image. The ROIs were used as input images in the compression and detection studies. To establish a "truth" file for the microcalcifications, the coordinate of each individual microcalcification in an ROI was identified manually with a cursor on a display workstation. The locations of the microcalcifications were verified by visually compared with those on the film mammograms using a magnifier. The coordinates were stored in a "truth" file and used for scoring the detection accuracy by the automated procedure, as discussed below. The total number of microcalcifications in the 25 ROIs was 293.

B. Laplacian pyramid hierarchical coding (LPHC)

The LPHC is a noncausal image coding method that decomposes an image into a low-pass image and a sequence of sub-band images, each of which is reduced in spatial resolution by a factor of 2, thereby forming a pyramidal hierarchical structure.[21] The LPHC technique implemented for this study is described in Appendix A. This compression method is developed for progressive image transmission. The low resolution version (the top level in the Gaussian pyramid) of the image is transmitted first to provide an early impression of the image content, progressively higher resolution images are subsequently transmitted to provide greater details. Transmission can be terminated as soon as sufficient image information is received. If the low level images are not needed and therefore not transmitted, the number of bits per image in the transmission is greatly reduced. Furthermore, the image size of the top level Gaussian pyramid image is small and the entropy of the Laplacian pyramid images is low because of the removal of the pixel-to-pixel correlation, the coding of the decomposed images can be more efficient than that of the original image. Further image compression can be achieved by reducing the quantization levels of the pixel values of the Laplacian pyramid images. For the purpose of this study, we will investigate the effects of the image reconstruction levels and the quantization levels of the Laplacian pyramid images on detection accuracy by the computer program.

C. Discrete cosine transform-full frame entropy coding (DCT-FFEC)

Block-DCT techniques are commonly used for compression of continuous-tone digital images. DCT can effectively localize most of the image information (energy) in a small area in the spatial frequency domain. However, the division of the image into small blocks for DCT often introduces blocky artifacts to the reconstructed images when high compression ratios are desired.

The full frame DCT technique transforms the entire image in one block. It not only eliminates the blocky artifacts, but also provides the advantage that the large-size DCT can localize the image information in a relatively smaller bandwidth than the small-size DCT. The coefficients in the full frame DCT matrix can be quantized with linear or nonlinear methods and then encoded by various coding techniques. For example, with a full frame bit allocation (FFBA) technique,[5] a bit-allocation table based on the characteristics of the transformed image and the desired compression ratio is produced. The table indicates the number of bits designated for a specific coefficient or groups of coefficients. The quantized coefficients are then packed into the bit space indicated in the table.

More recently, an entropy coding scheme that does not require a bit allocation table was developed for full frame DCT. For chest radiographs and CT images, the FFEC method was found to be more efficient than FFBA, in that it could produce a lower degree of MSE at a given compression ratio or increased the compression efficiency with the

same MSE.[6] These studies also indicated that a bit splitting-remapping method was useful in preventing errors in encoding for the most significant bits and to lessen edge artifacts caused by compression and decompression.[7] We therefore investigated the FFEC technique with and without bit splitting remapping for image compression in computer-aided diagnosis (CAD) applications. A description of the FFEC technique implemented for this study is given in Appendix B.

D. Mean-square error (MSE)

A commonly used indicator of information loss in a compression–decompression scheme is the MSE between the original image, $g(x,y)$, and the decompressed image, $g_r(x,y)$, as given by

$$\text{MSE}=\frac{1}{N_p}\sum_{\forall x}\sum_{\forall y}[g(x,y)-g_r(x,y)]^2, \quad (1)$$

where N_p is the total number of pixels in the image. The MSE is a global measurement of distortion of the image by a lossy compression technique. The use of MSE as an indicator of information loss in our computerized detection of microcalcifications on compressed–decompressed mammograms was evaluated in this study.

E. Computerized detection of microcalcifications

We have described our CAD algorithm for detection of microcalcifications in detail previously.[4,20,22,23] Briefly, there are three major steps in the algorithm: preprocessing, segmentation, and classification. In the preprocessing step, the input digital mammogram is processed with a signal-enhancement filter and a signal-suppression filter. The difference of these two filtered images results in an image in which the structured background is suppressed and the signal-to-noise ratio (SNR) of the microcalcifications is enhanced. This is also referred to as a difference-image technique. In the segmentation step, the program determines the gray level histogram of the processed image within the breast region. A gray level thresholding technique is used to locate potential signal sites above a global threshold. The threshold is changed iteratively until the number of sites obtained falls within the chosen input maximum and minimum numbers. At each potential site, a locally adaptive gray level thresholding technique in combination with region growing is performed to segment the connected pixels above a local threshold, which is calculated as the product of the local root-mean-square (rms) noise and an input SNR threshold. The characteristics of a segmented signal such as the size, contrast, SNR, and its location, are determined.

In the classification step, the computer program performs three tests to distinguish signals from noise or artifacts. A lower bound is imposed on the size to exclude signals below a certain size, which are likely to be noise, and an upper bound is set to exclude signals greater than a certain size, which are likely to be large benign calcifications. A contrast upper bound is also set to exclude potential signals that 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 artifacts and large benign calcifications. A regional clustering procedure is then applied to the remaining signals; a signal is kept if the number of signals found within a neighborhood of a chosen input diameter around that signal is greater than an input minimum number. The remaining signals that are not found to be in the neighborhood of any potential clusters will be considered isolated noise points or isolated calcifications and excluded. This clustering criterion is useful for reducing false positives because true microcalcifications of clinical interest always appear in clusters on mammograms. The specific parameters used in each step have been described previously.[4]

In this study, a signal-enhancement filter of 2×2 kernel of constant weights and a signal-suppression filter, which was a box-rim filter with a 20×20 kernel of constant weights around the rim and a 12×12 central area of 0 weights,[4] were used for preprocessing of the decompressed images obtained with the DCT-FFEC techniques. The sum of the weights was normalized to unity in each of the filters. For the images compressed with the LPHC technique, we made use of the Laplacian pyramid images to directly generate the difference image. The decoded Laplacian pyramid images at all levels were expanded to the original image size, summed together, and convolved with the 5×5 kernel, $w(m,n)$, defined in Appendix A. The resulting bandpass image was used as the difference image. This was equivalent to using the decompressed image as the signal-enhanced image and the Nth-level Gaussian pyramid image expanded N times to the original image size as the signal-suppressed image, and convolving the difference between the two images with the 5×5 kernel.

F. Analysis of detection accuracy

After passing the size, contrast, and the regional clustering criterion, the detected individual microcalcifications would be compared with the "truth" file of the input image. The numbers of true-positive (TP) and false-positive (FP) microcalcifications were scored. A detected signal was scored as a TP microcalcification if it was within 0.35 mm from a true microcalcification in the "truth" file. Once a true microcalcification was matched to a detected microcalcification, it would be eliminated from further matching. Any detected microcalcifications that did not match to a true microcalcification were scored as FPs. The trade-off between the TP and FP detection rates by the computer program was evaluated by the free-response receiver operating characteristic (FROC) analysis[24] by varying the input SNR threshold. A low SNR threshold corresponded to a lax criterion with a large number of FPs. A high SNR threshold corresponded to a stringent criterion with a small number of FPs and a loss in TPs. In this study, the FP rate was expressed as the number of FPs per unit area of the ROI image in order to reduce its dependence on the image size.[4] The information content of the reconstructed images was then evaluated by comparison of the FROC curves, which indicated the detection accuracy of the computer program.

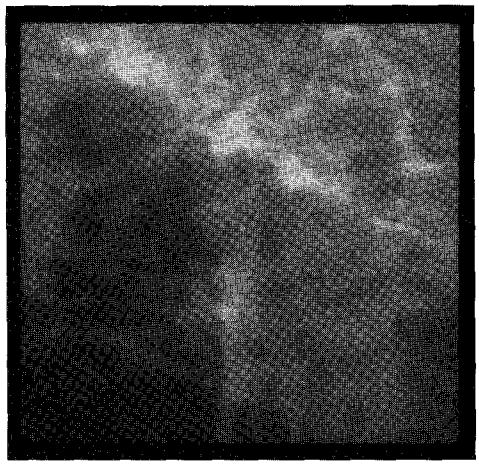


FIG. 1. An original ROI image with a cluster of microcalcifications used in the dataset of this study. A subtle cluster of about ten microcalcifications is located near the center of the ROI.

III. RESULTS

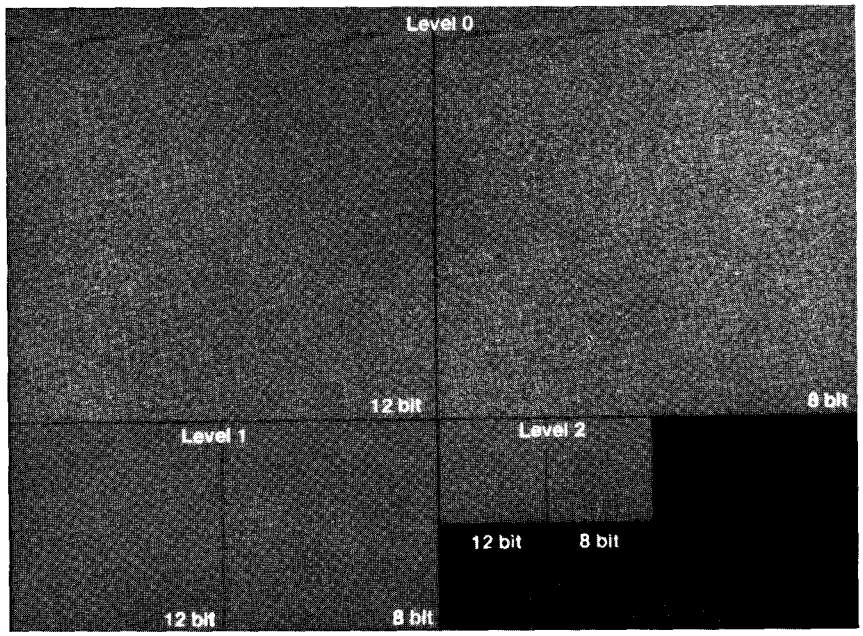
Figure 1 is an example of a 1024×1024-pixel ROI from a mammogram digitized to 12 bits. A subtle cluster of about ten microcalcifications is located near the center of the ROI. Three levels of the Laplacian images of the ROI at 12 bits, as well as the corresponding Laplacian images quantized to eight bits, are shown in Fig. 2(a). It can be seen that the frequency range of the information in the Laplacian images decreased with increasing levels on the pyramid. The gray level histograms of the two level-0 Laplacian images were plotted in Fig. 2(b), which illustrates the low entropy (defined in Appendix A) in a Laplacian pyramid image and the further reduction of entropy by requantization. To demonstrate visually the effect of requantization on image fidelity, we reconstructed the ROI from the three levels of eight-bit Laplacian pyramid images and the top level of the Gaussian pyramid, as shown in the flow diagram of LPHC in Appendix A. The error image between the original 12-bit image in Fig. 1 and the reconstructed image is shown in Fig. 3(a).

For the LPHC method, we have attempted three different ways to eliminate the LSBs in the Laplacian images. In the first method, each pixel value was divided by 2^l (l is the number of LSBs to be eliminated) and rounded off to the nearest integer. The pixel value was multiplied by 2^l during reconstruction. The second method was similar to the first, except that the quotient was truncated to an integer. The third method simply set the LSB to 0's by a bitwise AND operation with a bit plane mask and was the most efficient in terms of computational speed among the three. The third method yielded an image different from the second method because the Laplacian image was a difference image that contained negative integers. The truncation method reduced the absolute values of both the positive and negative integers, whereas the bit masking method shifted both the positive and negative integers to lower values (i.e., more negative for negative integers). The error images between the original and the reconstructed images with eight-bit quantization using the three different bit reduction methods are shown in Figs. 3(a)–3(c), respectively. The round-off and the bit masking methods for bit reduction resulted in similar error images, but the error image from the truncation method had obvious noise patterns. As described below, the MSE of the truncation method was much larger than those of the other two methods. The FROC curves for the LPHC techniques shown in Figs. 4 and 6 were obtained with the bit masking method because of its computational speed and its similarities to the round-off method.

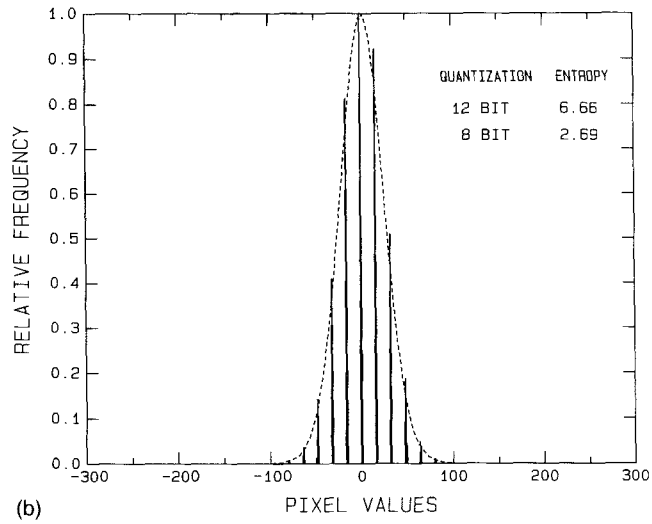
With the LPHC method, both the compression ratio and the reconstruction accuracy depend on the number of levels on the pyramid that the image is decomposed and reconstructed. We first investigated the dependence of the detection accuracy of the computer algorithm on the pyramid level. Figure 4 shows the FROC curves for two to four Gaussian pyramid levels of decomposition and reconstruction. The Laplacian images were quantized to eight bits in all cases. The images decomposed into three levels and reconstructed provided higher accuracy than those of two and four levels. Applying a paired t test to the TP rates at corresponding FP rates and pooled over the range of FPs between about 0.1 and 1 FP per cm^2, as discussed previously[4] and in Sec. IV, we found that the TP rates for the three-level decomposed images are significantly higher, with a two-tailed p value of less than 0.001, than those for the two- or four-level decomposed images.

The dependence of the detection accuracy on reconstruction level was also examined. The images were decomposed to three levels and the Laplacian pyramid images were maintained at 12 bits. The images were then reconstructed to the second level (image size of 256×256 pixels), to the first level (image size of 512×512 pixels), and to the original level (image size of 1024×1024 pixels). Because the first and second level images were already low-pass filtered, the 5×5 kernel was not applied to the difference images in the detection process in order to avoid further reduction in the spatial resolution. The detection results were plotted in Fig. 5. The detection accuracy decreased drastically if the images were not reconstructed to the original zeroth level. The decreases in the TP rates are statistically significant from the original level to the first level ($p<0.01$) and from the original to the second level ($p<0.001$).

Based on these results, images decomposed to the third level and reconstructed to the original level provided the highest detection accuracy. The following studies of the dependence of the detection accuracy on the bit depth of quantization were performed under this condition. The quantization of the Laplacian pyramid images was varied from six bits to nine bits. The FROC curves for these quantization bit depths were plotted in Fig. 6, along with the FROC curve for the original 12-bit images. There are no statistically significant differences among the curves for the 12-bit to 8-bit images ($p>0.05$). As the bit depth decreased further to seven bits, the reduction in the TP rates from those of the 12-bit images in the range of FP between 0.1 and 1 per cm^2 became statistically significant at $p<0.003$. The corresponding average bit rate for each of the conditions is shown in the figure legends. It indicates that, at eight-bit quantization of the La-



(a)



(b)

FIG. 2. (a) Three levels of Laplacian pyramid images of the ROI in Fig. 1. At each level the original 12-bit image is compared to the 8-bit quantized image. Note that the Laplacian pyramid images contain both positive and negative pixel values. For display purposes, a constant was added to every pixel of each image to shift them to the same positive mean level. (b) The gray level histograms of the level-0 Laplacian images at 12-bit (dashed curve) and at 8-bit (solid curve) gray level resolution.

placian images, the images can be compressed to an average of 3.28 bits/pixel without loss of the detectability of the microcalcifications by the computer algorithm.

With the DCT-FFEC approach, we evaluated three conditions, splitting into MSB of 3 and LSB of 9, splitting into MSB of 4 and LSB of 8, and without splitting. The detection accuracy in the reconstructed images for these conditions was compared in Figs. 7(a)–7(c). The parameters used for the compression schemes are listed in Table I in Appendix B. With splitting, the detection accuracy for the microcalcifications was similar to that in the original images at entropy coding ranges of seven to two bits and six to two bits ($p > 0.05$). The coding ranges of six to two bits resulted in an average bit rate of 1.25 for the three- and nine-bit splitting and 1.82 for the four- and eight-bit splitting. The detectability dropped significantly ($p < 0.001$) as the range reduced to five to two bits for both splitting schemes. The error image for the three- and nine-bit splitting with the entropy coding range of six to two bits is shown in Fig. 3(d). Without splitting, the detection accuracy for the entropy coding ranges of eight to two bits and seven to two bits was comparable to that of the original images ($p > 0.05$). The average bit rate for the seven to two bits coding was 1.49. The detectability was significantly lower for the coding range of six to two bits ($p < 0.001$).

The MSE and the bit rate for each of the compression

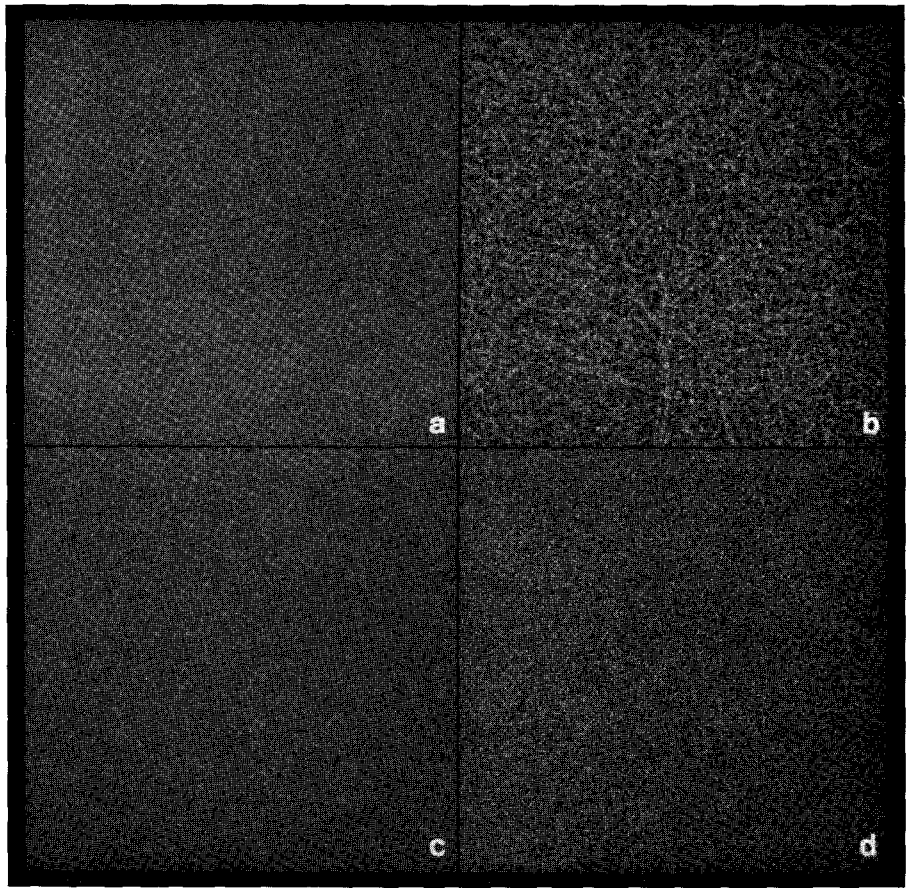


FIG. 3. The error images between the original image in Fig. 1 and the reconstructed images. Note that, for display purposes, a constant was added to every pixel of the error images to shift them to the same positive mean level. (a)–(c). The reconstructed images were obtained with the LPHC compression–decompression scheme. The image was decomposed to the second level of the Laplacian pyramid ($N=3$), quantized to eight bits, and reconstructed to the zeroth level. The eight-bit quantization was obtained by (a) dividing by 2^4 and rounding off to the nearest integers, (b) dividing by 2^4 and truncated to integer, and (c) setting the four LSBs to 0 using a bitwise AND operation with a bit plane mask. (d) The reconstructed image was obtained with the DCT-FFEC technique, three MSB and nine LSB splitting, entropy coding range of six to two bits.

schemes and conditions were averaged over the set of 25 images. The results were plotted on a semilog scale in Fig. 8. Each of the solid curves shows the results for a DCT-FFEC scheme with a different bit splitting parameter. The range of bits for entropy coding was varied along each curve. The results for the LPHC method were plotted as dashed curves. The lowest average bit rate that a scheme could achieve at which the detectability of the microcalcifications by computer was comparable to that of the original images was identified by an arrow for each of the compression schemes. The logarithm of the MSE appeared to be inversely proportional to the bit rate for most compression schemes. The curves for the DCT-FFEC schemes without or with splitting were comparable. The lowest average bit rate achieved by the DCT-FFEC method without a degradation in the detection accuracy is 1.25, corresponding to a compression ratio of about 9.6:1 in comparison to a 12-bit images without compression.

For the LPHC technique, the curve for each bit elimination method was plotted in Fig. 8 for linear quantization from six to nine bits. It can be seen that the round-off and the bit masking methods yielded similar MSE and bit rates for

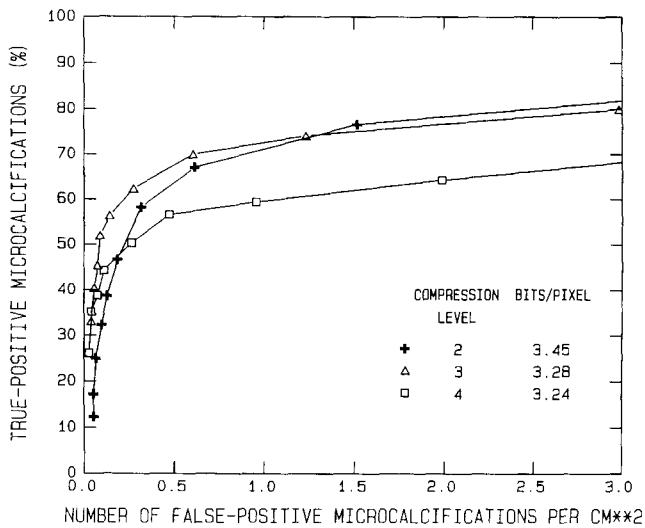


FIG. 4. The FROC curves for images decomposed to the Nth level (refer to the flow diagram in Fig. 9) and reconstructed to the zeroth level; crosses: $N=2$, triangles: $N=3$, and squares: $N=4$. The Laplacian pyramid images were quantized to eight bits.

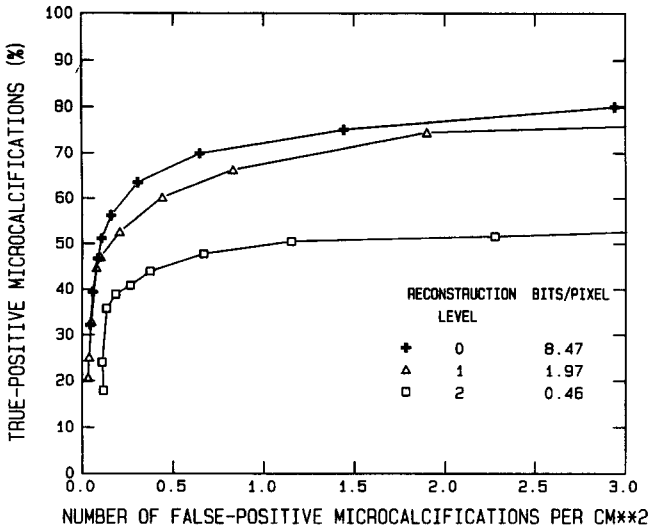


FIG. 5. The FROC curves for images decomposed to the $N=3$ level and reconstructed to the zeroth level (crosses), the first level (triangles), and the second level (squares). The Laplacian pyramid images were used at 12 bits without requantization.

seven to nine bit quantization. For the truncation method, the MSE was much higher than the other two methods for a given number of quantization bits. This is consistent with the visual appearance of the error image shown in Fig. 3(b). The entropy of the Laplacian images from the truncation method was lower than those from the other methods because a larger number of pixel values were truncated to zero from both the positive and the negative pixel values. Although it appeared that the detectability of the eight-bit truncated images did not decrease significantly compared to that of the original images, the detectability dropped more rapidly than the other two methods when the Laplacian images were further truncated to seven bits and six bits. Using the bit masking or round-off method, the lowest bit rate achieved without a degradation in the detectability of the microcalcifications by the computer was about 3.3, which corresponded to a

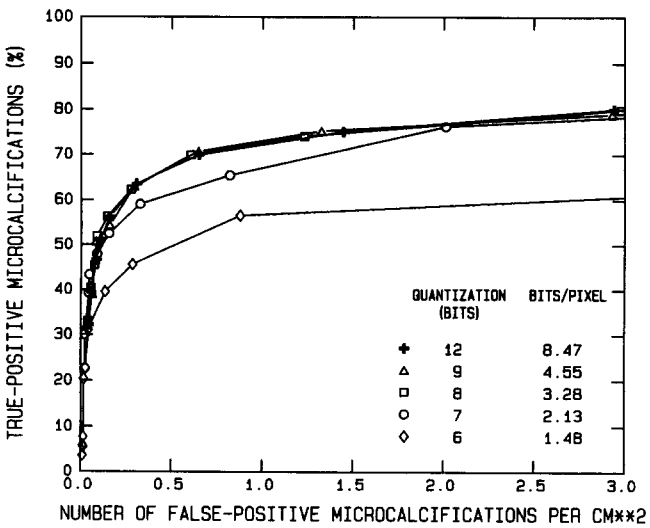


FIG. 6. The FROC curves for images decomposed to the $N=3$ level and reconstructed to the zeroth level. The Laplacian pyramid images were quantized to six to nine bits. The bit rate of 8.47 at 12 bits was calculated for a lossless compression with the LPHC technique.

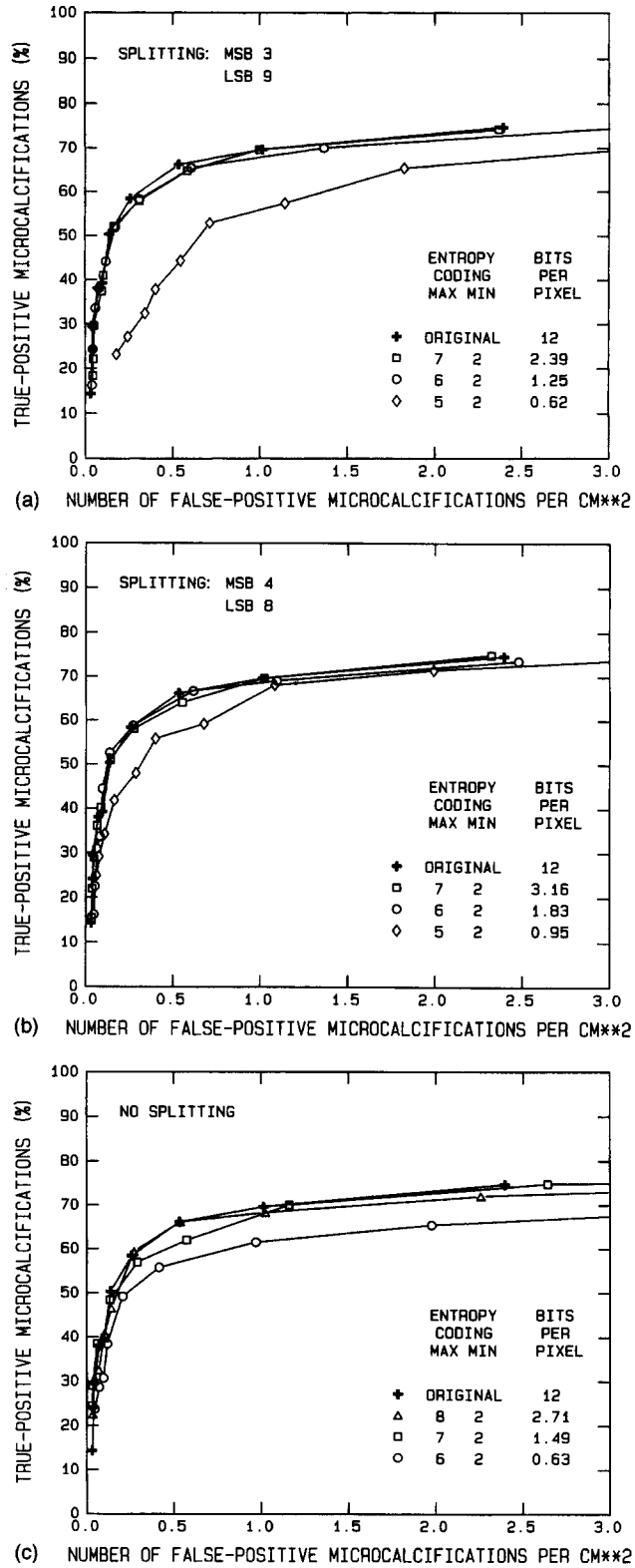


FIG. 7. The FROC curves for images compressed and decompressed with the DCT-FFEC technique, (a) with splitting at MSB=3 and LSB=9, (b) with splitting at MSB=4 and LSB=8, and (c) without splitting.

compression ratio of 3.6:1 in comparison to a 12-bit image without compression.

IV. DISCUSSION

The results of this study indicate that the DCT-FFEC method can provide a higher compression ratio than the LPHC method. The DCT-FFEC method with bit splitting of three MSB and nine LSB and entropy coding of six to two bits can achieve a compression ratio of 9.6:1 without significant degradation in the detectability of microcalcifications by a computer algorithm. On the other hand, with the LPHC method and the range of parameters studied, the compression ratio is only about 3.6:1 if the detectability of subtle microcalcifications has to be preserved. The DCT-FFEC method is thus about three times more efficient than the LPHC method if the detectability of microcalcifications by computer is used as the criterion of image fidelity. Although computer vision can be very different from human vision and the results cannot be simply generalized to image compression to be used for human readers, our results indicate that the DCT-FFEC method can retain high-frequency information, such as that of the microcalcifications better than the LPHC method.

Examples of the error images obtained from subtracting the decompressed image from the original image for the different compression schemes with the selected parameters are shown in Figs. 3(a)–3(d). It can be seen that the error image from the DCT-FFEC technique [Fig. 3(d)] appears to be more random and contain higher frequencies than those from the LPHC technique. The error image from the LPHC technique with truncation [Fig. 3(b)] contained visible structures of the mammogram. This problem can be avoided by a constant shift of all pixel values of the Laplacian pyramid images to positive integers before division and truncation. The resulting image and detectability of microcalcification will then be similar to the round-off and bit-masking methods.

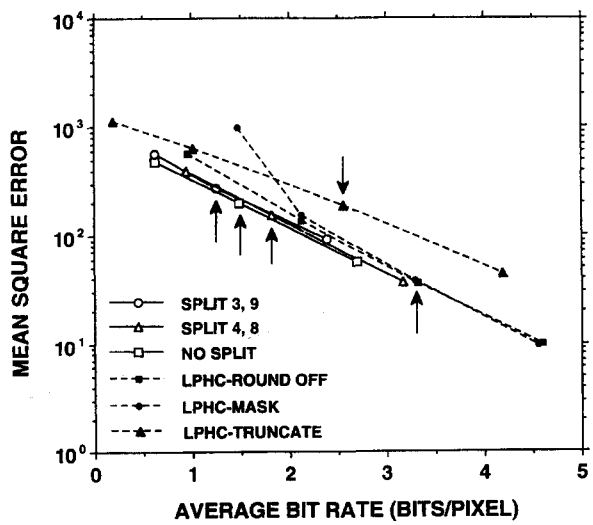


FIG. 8. The relationship between the mean square error (MSE) and the average bit rate for the various compression techniques. The solid curves are for DCT-FFEC techniques and the dashed curves are for the LPHC techniques.

It can be seen from Fig. 8 that, at the lowest average bit rate without degradation of detectability by the computer, the MSE from the different methods varied from about 36 to 274. However, the MSE did not correlate with the detectability of the microcalcifications by the computer. For example, at a higher MSE of 274, the DCT-FFEC method with three- and nine-bit splitting and entropy coding of six to two bits provided a higher detectability than the LPHC technique with seven-bit quantization at an MSE of 154. The relative image quality and the information content in the decompressed images therefore cannot be judged by comparison of the MSE. Experimental measurement of the detectability of signals in the decompressed images has to be performed for both human and machine observers in order to determine the loss of image information. The acceptable degree of information loss will also be dependent on the detection task.

For the purpose of this study, the detectability of microcalcifications by computer on the decompressed images was compared to that on the original images when the same preprocessing method for SNR enhancement in the CAD algorithm was used for each image compression method. For example, for the DCT-FFEC approach, a bandpass filter used in our previous studies[4,25] was used to extract the difference image. For the LPHC method, the Laplacian pyramid images were used to reconstruct the difference image. It can be seen by comparing the highest detection curves in Figs. 6 that the multiresolution Laplacian pyramid images can provide a higher detectability than the bandpass filtered images. Therefore, although the LPHC method is less efficient than the DCT-FFEC method for image compression, the Laplacian pyramid decomposition can be useful for SNR enhancement in the CAD program. This was also the motivation that we chose to evaluate the LPHC method for image compression in CAD applications.

It may be noted that, in the LPHC method, we chose to use linear quantization for compression of the Laplacian pyramid images. It is possible that some other methods can compress more efficiently these high-frequency bandpass images and improve the performance of the LPHC method. The ranges of parameters tested in the DCT-FFEC technique were also somewhat limited. We could not explore exhaustively the different image compression methods or all possible combinations of parameters for a specific method in this study. However, our investigation did indicate the utility of the DCT-FFEC technique and the importance of proper evaluation of information loss for mammographic image compression. More extensive comparison of various compression techniques is warranted in future studies.

One principal reason that digital mammograms require an extremely high resolution is due to the potential appearance of subtle microcalcifications, which may or may not be associated with breast cancer. In general, a mammogram is characterized as an image with predominantly low-frequency contents, except for microcalcifications and subtle speculations and margin characteristics of masses, which may indicate an early breast cancer. In other words, only a very small portion of the mammogram contains clinically significant image patterns. While these patterns can be greatly distorted

by an image compression technique, to our knowledge no global error measurement can quantify the effects. Many conventional compression techniques, therefore, can achieve a large compression ratio and obtain a low MSE without producing obvious visual degradation. However, these image compression techniques may degrade clinically significant information such as microcalcifications. In this study, our machine observer indicated the potential loss of detectability due to improper compression methods. The results of this preliminary study emphasize the need that special attention should be paid to the evaluation of image fidelity when compression techniques are applied to radiological images with potential subtle disease patterns. Optimization of image compression techniques based on analysis of detailed image features has recently been pursued by Lo *et al.*[26]

As found in our previous study,[4] the FROCFIT program[27] could not provide good fits to our FROC curves. Our attempt of applying the alternative FROC analysis[28] to the detection data and subsequently the CLABROC program[29] to the pair of correlated AFROC curves also failed to obtain reasonably fitted curves. This problem was probably caused by the correlation of the individual FP signals detected in an image due to the clustering criterion used in the detection process. We therefore could not use a fitted FROC curve or a single index such as the area under the AFROC curve[28] for comparison of the detection performance among different conditions. Because a rigorous statistical test of the significance of the differences between pairs of FROC curves was not yet available, we applied a paired t test to the TP values at a given FP 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 over a range of FP values of interest. The inclusion of TP pairs over a range of FP values took advantage of the consistency of the differences between the two FROC curves over the range of interest, similar to a curve fitting approach. However, the statistical significance might be somewhat overestimated because of the potential correlation between the TP pairs at the different SNR thresholds. An alternative test may be a paired t test of the differences in the partial areas under the FROC curves over the range of FP values of interest for the corresponding image pairs. The validity of these tests may be evaluated when a rigorous statistical significance test for FROC curves is developed.

V. CONCLUSION

We evaluated two image compression methods in this study. It was found that the DCT-FFEC method with bit splitting is more efficient than the LPHC method with linear quantization for compression of mammographic images, without degradation of the detectability of subtle microcalcifications by our automated detection algorithm. The highest compression ratio achieved without significant loss in detection accuracy was 9.6:1. It was demonstrated that the MSE was a poor indicator for comparison of information loss due to image compression. The evaluation of the acceptability of an image compression technique should therefore be based on human and machine observer studies.

ACKNOWLEDGMENTS

This work is supported by USPHS Grant No. CA 48129 and U.S. Army Grant No. DAMD 17-93-J-3007 (through Subgrant No. GU RX 4300-803UM from Georgetown University). The content of this publication does not necessarily reflect the position of Georgetown University or the government, and no official endorsement of any equipment and product of any companies mentioned in the publication should be inferred. The authors are grateful to Diane Williams for secretarial assistance.

APPENDIX A: LAPLACIAN PYRAMID HIERACHICAL CODING (LPHC)

A schematic of the LPHC technique for image compression and decompression is shown in Fig. 9. An input image G_0 is low-pass filtered with a local and symmetric weight kernel $w(m,n)$ and then subsampled by every other pixel to different levels sequentially according to the following relationship:

$$\text{REDUCE}(G_{k-1}) = G_k(i,j) = \sum_{m=-2}^{2} \sum_{n=-2}^{2} w(m,n) G_{k-1}(2i+m, 2j+n), \tag{A1}$$

where $1 \leq k \leq N$ is an index of compression level, N is the number of levels in the pyramid, and (i,j) is the pixel location in the image. The matrix size of an image at the kth level of the pyramid, G_k, is reduced by a factor of 4 compared with the $(k-1)$th level image, and is referred to as a "reduced" version of G_{k-1}. The reduced image is then expanded with a similar operation:

$$\text{EXPAND}(G_k) = E_{k-1}(i,j) = 4 \sum_{m=-2}^{2} \sum_{n=-2}^{2} w(m,n) G_k\left(\frac{i-m}{2}, \frac{j-n}{2}\right), \tag{A2}$$

where the summation is performed over the terms for which $(i-m)/2$ and $(j-n)/2$ are integers. An error image in level $k-1$ is given by the difference between G_{k-1} and E_{k-1}:

$$L_{k-1} = G_{k-1} - \text{EXPAND}(G_k). \tag{A3}$$

By performing the reduction N times, as shown in Fig. 9, a sequence of N low-pass filtered images with successively reduced spatial resolution and reduced sampling rate is obtained. Since one of the important convolution weight kernels resembles the Gaussian probability distribution, this se-

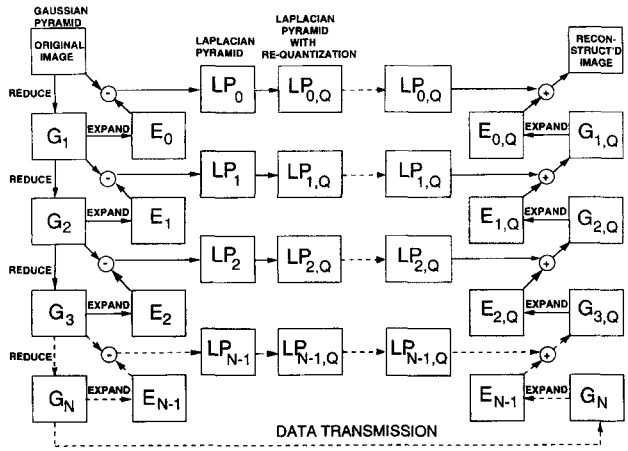


FIG. 9. Schematic diagram of the Laplacian pyramid hierarchical coding (LPHC) and decoding technique.

quence of low-pass filtered images is referred to as the Gaussian pyramid.

By expanding the reduced image and calculating the error image at each level, a sequence of N subband images of reduced sampling rate is also obtained. This sequence of error images is composed of bandpass filtered images and is referred to as the Laplacian pyramid. The scale of the Laplacian operator doubles from level to level of the pyramid, while the center frequency of the passband is reduced by an octave.

The original image can be reconstructed using the highest level Gaussian pyramid image, G_N, and the sequence of Laplacian pyramid images, L_k, $k=0,\ldots,N-1$, as shown on the right-hand side of Fig. 9:

$$G_{k-1}=L_{k-1}+\text{EXPAND}(G_k). \tag{A4}$$

If no lossy compression has been applied to G_N and the Laplacian pyramid images, the original image can be recovered without loss.

In this study, we decomposed the top level Gaussian pyramid image by the differential pulse code modulation (DPCM) technique.[30,31] The number of bits required for encoding the DPCM decomposed image was then determined by the entropy of its pixel value distribution, i.e., the histogram of its gray levels. For the Laplacian pyramid images, because their pixels values were decorrelated and could be assumed to be statistically independent, then the minimum number of bits per pixel required to exactly encode the image was given by its entropy. This optimum might be approached in practice through techniques such as variable-length encoding. The entropy of the pixel value distribution of an image was given by

$$\text{Entropy}=-\sum_{i=0}^{4095} f(i)\log_2 f(i), \tag{A5}$$

where $f(i)$ was the observed probability of occurrence of gray level i. Assuming that the variable-length code words were used in data transmission to take advantage of the nonuniform distribution of pixel values, the effective number of bits for a given Laplacian pyramid level was its entropy times its matrix size. The effective number of bits per pixel for the encoded image was thus the sum of the number of bits for all levels of the component images divided by the matrix size of the original image.

Following the approach described by Burt and Adelson,[21] a 5×5 kernel of weights $w(m,n)$ that was separable to $w(m,n)=h(m)h(n)$ was used in this study. Here h was a symmetric function, such that $h(i)=h(-i)$, for $i=0,1,2$. The weights were subject to the constraint that all nodes at a given level contributed the same total weight to nodes at the next higher level. Therefore, $h(0)=a$, $h(-1)=h(1)=\frac{1}{4}$, $h(-2)=h(2)=\frac{1}{4}-a/2$. The constant a was chosen to be 0.4 in this study to obtain a Gaussian-like function. A uniform quantization by eliminating the least significant bits was applied to the Laplacian pyramid images. Although these parameters might not be optimal choices for mammographic images, they were selected as typical values for study of the effects of the LPHC method on mammograms.

APPENDIX B: DISCRETE COSINE TRANSFORM-FULL FRAME ENTROPY CODING (DCTF-FEC)

A schematic diagram of the FFEC technique with splitting and remapping is shown in Fig. 10. An input image of $(n+k)$ bits is split into n most significant bits (MSB) and k least significant bits (LSB). Because the MSB in medical images are highly correlated, they can be encoded by correlation encoding such as Lempel–Ziv (LZ) coding or run-length/Huffman coding with high compression ratio. For the LSB image, the bits are remapped in order to convert the residual data into an image with a more continuous tone. The remapping of the LSB, denoted as RLSB, for an image with gray levels $g(x,y)$ can be expressed as

$$\text{RLSB}_k(g(x,y))=\text{LSB}_k(g(x,y)), \quad \text{for } [g(x,y)\&2^k]=0, \tag{B1}$$

$$\text{RLSB}_k(g(x,y))=2^k-1-\text{LSB}_k(g(x,y)),$$

$$\text{for } [g(x,y)\&2^k)\neq 0, \tag{B2}$$

where "&" is the "AND" operation for the bit map of the integers. The splitted and remapped image, RLSB $(g(x,y))$

TABLE I. The parameters used in the discrete cosine transform-full frame entropy coding technique.

Zone	Frequency region	No. of bits for coding	
1	0–63	Floating point	
2	64–127	12	
		Maximum	Minimum
3	128–1023	8	2
		7	2
		6	2
		5	2

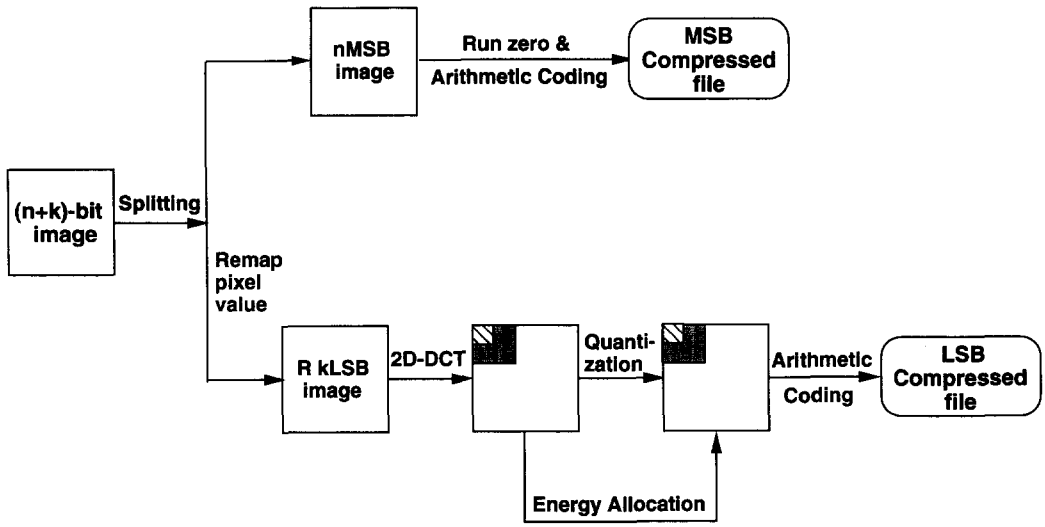


FIG. 10. Schematic diagram of the discrete cosine transform-full frame entropy coding (DCT-FFEC) technique with image splitting and remapping.

is then subject to a two-dimensional DCT. The spatial frequency domain image is divided into three zones for linear quantization. The zone boundaries and the number of bits used in each zone are tabulated in Table I. In the low-frequency zone, the DCT coefficients are stored and transmitted as the original floating point values so that no information is lost. In the mid-frequency zone, the coefficients are quantized to 12 bit integers. In the high-frequency zone, the coefficients are quantized to a range of bits specified by an input maximum and minimum number. The compression ratio is large when the maximum number of bits allowed is small. A specific number of bits to be used for a given coefficient in this zone is determined by an energy allocation scheme.[6,7] The ranges of seven to two bits, six to two bits, and five to two bits were compared for the compression schemes with splitting in this study. The quantized coefficients were then submitted to a statistical coding routine for data packing. The standard statistical coding schemes included arithmetic and Huffman coding, of which the former was used in this study.

For FFEC without bit splitting and remapping, the procedure is similar to those described above. The only difference is that there is no MSB image to be encoded. The entire image undergoes DCT, zonal quantization, and arithmetic coding, as illustrated in the lower path of Fig. 10. For energy allocation, the ranges of eight to two bits to six to two bits were evaluated.

To decompress the FFEC image, the reverse operation of compression is employed. The quantized coefficients are decoded, followed by reverse quantization, and then by the inverse DCT. Because of the quantization process that reduces the DCT coefficients of real numbers to integers with a finite number of bits, the reverse quantization cannot recover the original image information in its entirety. The degree of information loss with the FFEC technique depends on the information content of the input images and the compression ratio.

[a]Correspondence and reprint address: Heang-Ping Chan, Ph.D., University of Michigan, Department of Radiology, Taubman Center 2910, 1500 E. Medical Center Drive, Ann Arbor, Michigan 48109-0326. Telephone: (313) 936-4357; Fax: (313) 936-9723; Electronic mail: chanhp@umich.edu
[b]Department of Radiology, Georgetown University.
[c]Current address: Radiological Sciences and Technology, Massachusetts General Hospital.
[d]Current address: Department of Diagnostic Radiology, Stanford University.
[e]Department of Radiation Oncology, University of Michigan.

[1]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).
[2]D. Winfield, M. Silbiger, G. S. Brown, L. Clarke, S. Dwyer, M. Yaffe, and F. Shtern, "Technology transfer in digital mammography: Report of the Joint National Cancer Institute–National Aeronautics and Space Administration Workshop of May 19–20, 1993," Invest. Radiol. **29**, 507–515 (1994).
[3]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).
[4]H. P. Chan, L. T. Niklason, D. M. Ikeda, K. L. Lam, and D. D. Adler, "Digitization requirements in mammography: Effects on computer-aided detection of microcalcifications," Med. Phys. **21**, 1203–1211 (1994).
[5]S.-C. B. Lo and H. K. Huang, "Compression of radiological images with 512, 1024, and 2048 matrices," Radiology **161**, 519–525 (1986).
[6]S.-C. B. Lo, B. Krasner, S. K. Mun, and S. C. Horii, "Full-frame entropy coding for radiological image compression," Proc. SPIE **1444**, 265–277 (1991).
[7]S.-C. B. Lo, E. L. Shen, S. K. Mun, and J. Chen, "A method for splitting digital value in radiological image compression," Med. Phys. **18**, 939–946 (1991).
[8]T. Ishigaki, S. Sakuma, M. Ikeda, Y. Itoh, M. Suzuki, and S. Iwai, "Clinical evaluation of irreversible image compression: Analysis of chest imaging with computed radiography," Radiology **175**, 739–743 (1990).
[9]H. MacMahon, K. Doi, S. Sanada, S. M. Montner, M. L. Giger, C. E. Metz, N. Nakamori, F.-F. Yin, X.-W. Xu, H. Yonekawa, and H. Takeuchi, "Data compression: Effect on diagnostic accuracy in digital chest radiography," Radiology **178**, 175–179 (1991).
[10]D. R. Aberle, F. Gleeson, J. W. Sayre, K. Brown, P. Batra, D. A. Young, B. K. Stewart, B. K. T. Ho, and H. K. Huang, "The effect of irreversible image compression on diagnostic accuracy in thoracic imaging," Invest. Radiol. **28**, 398–403 (1993).

[11]J. W. Sayre, B. K. T. Ho, M. I. Boechat, T. R. Hall, and H. K. Huang, "Subperiosteal resorption: Effect of full-frame image compression of hand radiographs on diagnostic accuracy," Radiology **185**, 599–603 (1992).

[12]S.-C. B. Lo, H. Li, B. H. Krasner, M. T. Freedman, and S. K. Mun, "Full-frame compression of discrete wavelet and cosine transforms," Proc. SPIE **2431**, 195–202 (1995).

[13]P. W. Jones, S. Daly, R. S. Gaborski, and M. Rabbani, "Comparative study of wavelet and DCT decomposition with equivalent quantization and encoding strategies for medical images," Proc. SPIE **2431**, 571–582 (1995).

[14]M. Smith and S. Eddins, "Analysis/synthesis techniques for subband image coding," IEEE Trans. Acoust. Speech, Signal Proc. **38**, 1446–1456 (1991).

[15]J. Wang and H. K. Huang, "Three-dimensional medical image compression using a wavelet transform with parallel computing," Proc. SPIE **2431**, 162–172 (1995).

[16]M. A. Goldberg, M. Pivovarov, W. W. Mayo-Smith, M. P. Bhalla, J. G. Blickman, R. T. Bramson, G. W. L. Boland, H. J. Llewellyn, and E. Halpern, "Application of wavelet compression to digitized radiographs," Am. J. Roentgenology. **163**, 463–468 (1994).

[17]H. Benoit-Cattin, O. Baudin, A. Baskurt, and R. Goutte, "Coding mammograms using wavelet transform," Proc. SPIE **2164**, 282–290 (1994).

[18]H. S. Huang, L. Guan, and H. Kung, "Compression of digital mammogram databases using a near-lossless scheme," Proc. ICIP **2**, 21–24 (1995).

[19]A. J. Maeder, "Mammogram compression using adaptive prediction," Proc. SPIE **2431**, 216–231 (1995).

[20]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).

[21]P. J. Burt and E. H. Adelson, "The Laplacian pyramid as a compact image code," IEEE Trans. Commun. **COM-31**, 337–345 (1983).

[22]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).

[23]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).

[24]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).

[25]H. P. Chan, S.-C. B. Lo, B. Sahiner, K. L. Lam, and M. A. Helvie, "Computer-aided detection of mammographic microcalcifications: Pattern recognition with an artificial neural network," Med. Phys. **22**, 1555–1567 (1995).

[26]S.-C. B. Lo, H. Li, J. Wang, M. T. Freedman, and S. K. Mun, "On optimization of orthonormal wavelet decomposition: Implication of data accuracy, feature preservation, and compression effects," Proc. SPIE **2707**, 201–214 (1996).

[27]D. P. Chakraborty, "Maximum likelihood analysis of free-response receiver operating characteristic (FROC) data," Med. Phys. **16**, 561–568 (1989).

[28]D. P. Chakraborty and L. H. L. Winter, "Free-response methodology, alternate analysis and a new observer-performance experiment," Radiology **174**, 873–881 (1990).

[29]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).

[30]C. C. Cutler, "Differential quantization of communication signals," Patent 2-605-361, Application June 1950, Issuance July 1952.

[31]R. P. Abbott, "A differential pulse-code-modulation coder for videotelephone using four bits per sample," IEEE Trans. Commun. Tech. **COM-19**, 907–913 (1971).