On Techniques for Detecting Circumscribed Masses in Mammograms

SHUK-MEI LAI, XIAOBO LI, AND WALTER F. BISCHOF

Abstract—**This paper presents a method for detecting one type of breast tumor, circumscribed masses, in mammograms. It relies on a combination of criteria used by experts including the shape, brightness contrast, and uniform density of tumor areas. The method uses modified median filtering to enhance mammogram images and template matching to detect breast tumors. In the template matching step, suspicious areas are picked by thresholding the cross-correlation values and a percentile method is used to determine a threshold for each film. In addition, two tests are designed to remove false alarms from the resulting candidates. The results obtained by applying these techniques to a set of test images are described.**

I. Introduction

BREAST cancer is not only a leading cause of death among all cancers for women of middle age and older [24], but its incidence is rising. Primary prevention is not possible since the cause of this disease is still not understood. However, current methods of treatment are very effective against breast cancer in its early phase when the balance between the tumor and its host is more favorable [23]. Therefore, removal of the cancer while it is still in its early stages is the most promising way to achieve a significant change in the current breast cancer situation. Of all diagnostic methods currently available for this purpose, mammography is the most reliable method for the detection of early breast cancer [24]. Hence, a mass screening program utilizing mammography is obviously the best weapon against breast cancer.

A major problem expected with such a screening program would involve the interpretation of the large volume of images produced. In addition to a shortage of trained radiologists and the need to improve the cost benefit ratio of such a program, it is difficult for human radiologists to maintain interest in interpreting large numbers of images in which only a small number show abnormalities [6]. Hence, the need to construct computer-aided systems to diagnose breast cancer in mammograms becomes apparent.

The analysis of mammograms by computer can be roughly divided into three steps: 1) enhancement of preselected features and removal of irrelevant details using application-dependent techniques, 2) localization of suspicious areas, and 3) classification of these areas into nontumors, benign, or malignant tumor areas. The analysis is difficult for several reasons. First, it involves the analysis of small features of low contrast superimposed onto nonuniform backgrounds. Second, the diagnostic information content in mammograms is more semantic than statistical in nature as is often the case with radiographic images in general [9]. That is, the features essential for diagnosing a particular disease are determined both by the disease and the class of images. Third, the resolution required in a particular disease application may, for example, be an order of magnitude higher than that necessary for all general diseases using the same class of images.

A large repertoire of image processing techniques has been developed for image enhancement, object localization, and pattern classification. However, due to the problems of radiographic image analysis mentioned above, most image processing techniques which have been applied to biomedical situations have been found to be very application dependent. This phenomenon is analogous to the fact that radiologists adopt different strategies to analyze different types of medical images. Thus, starting with a large repertoire of techniques, one may have to combine and modify some existing techniques to create the best technique for a particular application.

In the past, several groups [1], [23] have demonstrated the potential use of computers in feature-based classification of suspicious areas. Since human assistance is needed to locate these areas before the computer processes the images, these systems are not fully automated. Hand *et al.* [10] have developed a method for identifying and locating abnormal areas in xeromammograms. Their approach, which utilizes texture and shape parameters together with a left–right asymmetry analysis, shows a high false alarm rate of over 50 suspicious areas per xeromammogram. Since xeromammograms are produced using a different recording technique from that used for mammograms, the performance of Hand *et al.*'s method with mammograms cannot be directly determined.

In the present work, we are concerned with the detection and localization of one particular tumor, *circumscribed masses*. In order to design a method for detecting areas corresponding to circumscribed masses or areas that may correspond to circumscribed masses (suspicious areas), the radiologists' descriptions [21], [13] of tumor symptoms must first be translated into computational procedures. Based on these descriptions, we define a *suspi-*

Manuscript received October 4, 1988; revised May 14, 1989. This work was supported in part by the Canadian Natural Sciences and Engineering Research Council under Grants OGP9198 and OGP38521.

The authors are with the Department of Computing Science, University of Alberta, Edmonton, Alta., Canada T6G 2H1.

IEEE Log Number 8929507.

cious area as an area which 1) is brighter than its surrounding tissue, 2) has uniform density inside the area, 3) has an approximately circular shape of varying size, and 4) has fuzzy edges. It must be emphasized that a suspicious area is not necessarily a tumor area. It can also be a benign tumor or an area which radiologists would choose to examine in greater detail. The methods presented here are concerned only with the localization of suspicious areas and no attempt is made to further classify these areas.

Locating suspicious areas in mammograms is difficult for a number of reasons. The small differences in density between normal and tumorous tissues in human breasts create little contrast between a tumor area and its background in the image. This contrast is further reduced in the filming and digitization process of the mammogram images. In addition, the presence of noise and anatomical structures, such as ducts and glands, increases the background variations of tumor areas. The boundaries of tumor areas are fuzzy, and in some instances, only partially visible. Together with the small size of early-stage tumors, this makes any attempt to segment the image by global gray level thresholding technique very difficult.

Two examples of mammogram images are shown in Plates 1 and 2, respectively. The positions of the suspicious areas in these mammograms are indicated by markers on the edge of the plate. The suspicious area in image 1 is easier to detect than that in image 2 due to the fact that there is less background variation near the suspicious areas. The background variations in image 2 are due to the presence of gland and fatty tissue in the breast.

The approach described in this paper takes this problem into account and existing techniques are modified to improve tumor detection. Section II discusses methods for enhancing mammographic features and presents our approach which uses *selective median filtering* to enhance mammograms. Section III discusses the *template matching* technique we used for detecting suspicious areas. In addition, since the candidate areas produced by this matching process are not all suspicious areas, two tests designed for removing false alarms from the resulting candidates are presented in Section IV. Section V discusses the computational efficiency of the template matching process, and the conclusions of this paper are presented in Section VI.

II. Image Enhancement

There are two possible approaches in enhancing mammographic features. One is to increase the contrast of suspicious areas and the other is to remove background noise. Some techniques for contrast enhancement of film mammograms have been suggested earlier [6], [8]. Those methods are based on adaptive neighborhood processing with a set of contrast enhancement functions to enhance the contrast of mammographic features. In our research, we take the other approach and enhance the images by removing background noise while preserving the edge information of suspicious areas in the images. We investigated four known selective averaging schemes and a new method, a modification of median filtering, with respect to their performance in enhancing mammogram images.

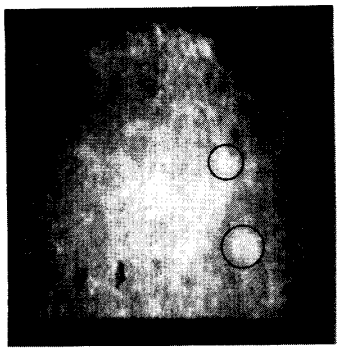

Plate 1. Mammogram image 1 with easily detectable suspicious areas. Suspicious areas are marked by circles.

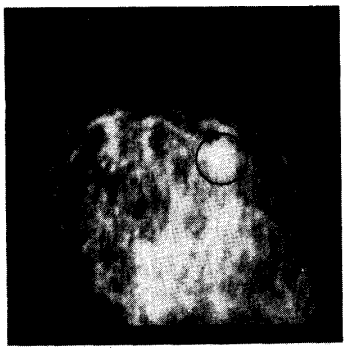

Plate 2. Mammogram image 2 with a suspicious area which is difficult to detect. The suspicious area is marked by a circle.

A. Selective Averaging

Edge-preserving noise removal can be achieved using an averaging scheme that avoids averaging across edges. One way to achieve this is to restrict averaging to a subset of neighborhood pixels such that the most likely edge in a given neighborhood does not cross this subset. Several such methods proposed in the literature were tested in our image enhancement stage.

Edge-Preserving Smoothing: This method [14] tries to search for a homogeneous neighborhood in different directions of a given pixel and averages in this neighborhood only. The search is done by rotating a window around a given pixel, using the variance within the window as a measure of homogeneity of an area, and selecting the smallest variance window for averaging.

Half-Neighborhood Method: This method [18] operates like unweighted averaging at pixels in the interior of a region. For pixels that lie on an edge, the neighborhood is subdivided in various ways, and the half neighborhood whose average level differs the most from that of the other half is chosen for averaging. It assumes that edges are straight and are composed of consecutive neighborhood pixels.

The k-Nearest Neighbor Method: This method [5] is based on the idea that pixels in the same region should have similar gray values. Therefore, under this scheme, the average is taken using the k neighbors within a given neighborhood region whose gray levels are closest to that of the given pixel. It makes no attempt to detect the presence of edges and does not require that the neighbors involved in the averaging are adjacent.

Directional Smoothing: This method [16] is similar to the half-neighborhood method in that for pixels in the interior of a region, averaging is done using all neighborhood pixels. If an edge is present, a directional average is taken on those neighbors that lie in the direction along the edge. The edge position and orientation are obtained by convolving the image with a set of four oriented edge detection masks.

B. Median Filtering

Median filtering has been found to be very powerful in removing noise from two-dimensional signals without blurring edges [3]. This makes it particularly suitable for enhancing images.

To apply median filtering to a digital picture, we replace the value at a pixel by the median of the values in a neighborhood of the pixel. Given a set of n numbers $\{x_1 \cdots x_n\}$ and we define the ordered set $\{x_1^* \cdots x_n^*\}$. The median of the set is given then by

$$\text{median } \{x_1 \cdots x_n\} = \begin{cases} x^*_{n/2} & n \text{ even} \\ x^*_{(n+1)/2} & n \text{ odd.} \end{cases}$$

Two-dimensional median filters can be defined for arbitrary sizes and shapes of filter windows $W(i, j)$, such as line segments, squares, circles, and crosses. The two-dimensional median filtering operation is defined as follows. For a two-dimensional filter window $W(i, j)$ centered at image coordinates (i, j) of a picture $\{x_{ij}: (i, j) \in Z^2\}$, the median filtering output is

$$\hat{x}_{ij} = \text{median } \{x_{r,s}: (r, s) \in N(i, j)\}, \ (i, j) \in Z^2$$

where $N(i, j)$ is the area in the image covered by window $W(i, j)$.

Median filters have several properties that make them superior to low-pass filters [11]. If an image has impulse-like noise, median filtering can remove it without significantly distorting the signal, and if an image contains edges, median filtering can preserve them due to the fact that only a small fraction of the neighborhood overlaps the edge. Further, Bovik *et al.* [4] have analyzed generalized filters based on linear combinations of order statistics and demonstrated that among those, the median filter is nearly optimal for suppressing impulse noise or noise which is characterized by a large percentage of outliers.

C. Selective Median Filtering

Experiments showed that the edge preservation power of the standard median filter is not sufficient for enhancing mammogram images due to the fuzziness of the boundaries of suspicious areas. We introduce a modification of the median filter, selective median filter (SMF), which is defined as follows.

For a window $W(i, j)$ centered at image coordinates (i, j), the output of the selective median filter is

$$\hat{x}_{ij} = \text{median } \{x_{r,s}: (r, s) \in N(i, j) \text{ and } |x_{r,s} - x_{ij}| < T\}, \ (i, j) \in Z^2$$

where $N(i, j)$ is the area in the image covered by window $W(i, j)$ and T is a threshold.

In computing the median, the set of pixels is restricted to those with a difference in gray level no greater than some threshold T. By adjusting the parameter T, the amount of edge smearing can be controlled. If T is small, the edge-preserving power of SMF is strong, but its smoothing effect will be small. If T is large, the SMF behaves the other way around. This modification of the median filter is related to selective averaging schemes developed for linear filters [16] that show good results in improving the edge preserving power of linear low-pass filters.

To achieve strong noise suppression, one can either use a filter with a large window size or the filter can be applied repeatedly [7], [15], [19]. The first approach has several drawbacks [7], [15]. The median filter, designed to act as a low-pass filter in homogeneous areas, will respond more and more like a bandpass filter as window size is increased. Further, increasing the window size leads to increased noise suppression, but also to increased signal distortion. On the other hand, they showed iterated median filtering reduces an original signal to an invariant signal, called *root* signal, and that only piecewise constant images are roots to the median filters, implying that edge information is not lost by iterating the filtering process.

Stein [19] has presented a modification of the median filter, the adaptive recursive filter, which is related to our selective median filter. Stein's method can be loosely described as a "half-neighborhood" median filter, whereas the SMF could be described as a "nearest neighbor" median filter. Although the two methods differ, we would expect their performance to be comparable.

D. Evaluations and Comparison

In enhancing mammogram images, we require an enhancement technique which clears noise while preserving edges of suspicious areas. In this section, the performance of the neighborhood averaging methods discussed and the proposed selective median filtering method are compared and evaluated.

Performance comparison of the different averaging schemes was done visually (attempts to evaluate the techniques using quantitative measurements [9] proved to be unsatisfactory since they could not adequately measure the preservation of the image structure, such as the preservation of important edges). To facilitate the evaluation of the edge-preserving power of the various enhancement techniques, the boundary information of the enhanced im-

ages was extracted for visual evaluation. A high-pass (Laplacian) filter was used for this purpose, the high-pass filter output being defined as

$$g(x, y) = \left| \sum_{i=-1}^{1} \sum_{j=-1}^{1} f(x + i, y + j) - 9f(x, y) \right|$$

where $f(x, y)$ is an enhanced image. The high-pass filtered image was then converted to a binary image using a fixed threshold.

Among the five techniques implemented, the edge preservation power of the directional smoothing method was the worst and it tended to preserve straight edges better than curved ones. However, the shape of suspicious areas is approximately circular, and therefore, edges do not always lie in a straight line. In this case, no matter which direction the average is computed, the edge pixel is always being averaged with pixels not belonging to the same region. As a result, the edges of suspicious areas are blurred, and as the window size increases, this effect is more prominent.

The k-nearest neighbor method cannot preserve edges as well as the half-neighbor or SMF method. It assumes that the five nearest neighbors (in gray level) must belong to the same region as the center pixel in a window. This is not always true, and the five nearest neighbors may include pixels that lie outside the region. The half-neighborhood method preserves edges better because it treats edge pixels differently and restricts that the five neighbors selected for averaging must be consecutive pixels. This latter restriction reduces the chance of averaging an edge pixel with pixels belonging to another region. Using a threshold to select pixels from which the median is chosen, the SMF achieves the same effect.

In comparing the high-pass results of the various enhanced images, it was noted that only the edge preserving smoothing method (Plate 3) and the SMF method (Plate 4) preserve the edges of the suspicious areas adequately. In both plates, the edges of the suspicious areas appear as approximately closed rings. However, the noise cleaning power of the SMF method is better than that of the edge-preserving smoothing method. This can be illustrated by the fact that there are fewer noise edges in Plate 4 than in Plate 3. Based on these observations, it was concluded that the SMF method with a 5×5 window size enhances the mammogram images better than the other methods.

E. Performance of the Proposed Method

To further evaluate the performance of the SMF, a set of 24 mammogram films was used. They were randomly selected from mammogram files and were diagnosed by an expert radiologist who circled the suspicious areas in each test case. His markings were then translated into x, y coordinates representing the approximate center and radius measures.

The SMF has three parameters that can be adjusted to adapt the filter to the noise characteristics of the mammograms, the three parameters being the threshold T, the number of iterations, and the window size W. To set the values of these three parameters, a training set of seven mammograms is used. The other 17 mammograms were grouped into a test set for testing the performance of the proposed method after the values of the parameters had been determined. Mammograms can differ quite considerably in their gray level histogram due to exposure differences in mammographic screening, and thus it can be quite difficult to determine a single set of parameter values applicable to all mammograms. To avoid this problem, all mammograms were normalized to a gray level range of 0–255 before filtering.

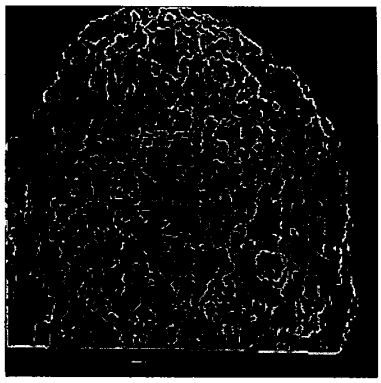

Plate 3. Laplacian of mammogram image 1 which is enhanced by the edge-preserving smoothing method.

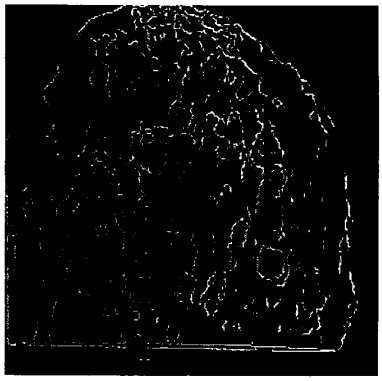

Plate 4. Laplacian of mammogram image 1 which is enhanced by the selective median filtering method.

To estimate the maximum value for the parameter T, the contrast between pixels in the suspicious areas (of the training set images) was checked. It was noted that the contrast ranged from 5 to 15 gray levels. Using different threshold values within this range, the SMF was applied to the training set. By observing the amount of noise edges and "relevant" edges in the high-pass results, it was concluded that although a large threshold value (e.g., 10) removed noise more effectively, a smaller threshold value (e.g., 5) preserved edges of all suspicious areas better.

In addition, it was noted that the performance of SMF does not improve much after the fifth iteration if the other two parameters (W and T) are kept constant. This is in agreement with Gallagher and Wise's [7] observation that

every signal can only be reduced down to a certain point no matter how many times the median filter is applied.

Best performance of the SMF was found for window size $W = 9 \times 9$, number of iterations = 5, and edge threshold $T = 5$. Using these parameter values, the SMF was applied to the mammogram image 1 (Plate 1) and 2 (Plate 2) and the filtered images are shown in Plates 5 and 6. Compared to the original images, the background variation is reduced in the filtered images, while the boundaries of all suspicious areas are preserved. The filtered images have, however, a mottled appearance and this is caused by the iteration of the filter. This effect is not desirable and it may produce false alarms in the tumor detection process. This problem can be easily overcome by applying a false alarm test using the original unfiltered image. This test is discussed in detail in Section IV.

In applying this SMF on the test of 17 mammogram images, the performance of the SMF was found to be satisfactory. The background variations in the filtered images is reduced and the boundaries of all suspicious areas are preserved.

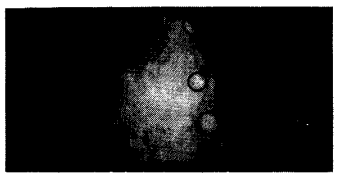

Plate 5. Mammogram images 1 enhanced by the selective median filtering method using five iterations, window size of 9, and $T = 5$. The suspicious areas reported by the proposed tumor detection method are marked by circles.

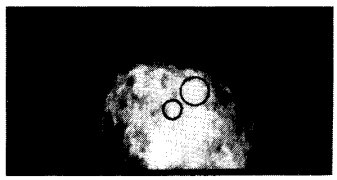

Plate 6. Mammogram images 2 enhanced by the selective median filtering method using five iterations, window size of 9, and $T = 5$. The suspicious areas reported by the proposed tumor detection method are marked by circles.

III. Tumor Detection

The approximately circular shape and the brightness homogeneity of tumor areas are important characteristics used in detecting suspicious areas. General techniques developed for object detection, such as pixel-based and region-based segmentation techniques, can make use of the brightness homogeneity property of regions to segment the image, but they are blind to the shape of tumors. Since brightness homogeneity is not a unique characteristic of tumor areas, it is very difficult to use this criterion alone to segment the image successfully into tumor and background regions.

One way to detect the approximately circular tumors is to extract image edges and then look for ring-like structures. However, in noisy or lightly textured images, a large number of noisy edges is extracted and edge tracking becomes very difficult, if not impossible.

The other approach to using shape information in tumor detection and the one we used in our work is based on template matching. With this approach, the shape and homogeneity characteristics of tumor areas can be used as the match criterion by defining them in the templates. The location and size of detected suspicious areas can then be obtained from the output of the matching operation. However, as will be discussed below, this approach has to be complemented by several additional tests for good performance due to the fact that the tumors to be detected can differ quite considerably from an "ideal" shape.

A. Definition of a Tumor-Like Template

A tumor-like template is defined based on three characteristics of tumor areas, namely, 1) brightness contrast (i.e., a bright object in a dark background), 2) uniform density, and 3) approximately circular shape.

The template used to match tumors with a diameter of five pixels is shown in Fig. 1. A circular patch of 1's in the center of the template represents a tumor area having uniform density. To allow tumor shape to deviate slightly from a perfect circle, the patch of 1's is bounded by a ring of 0's. This is a "don't care" area in the match. The background of the patch is filled with -1's instead of 0's because we are looking for a light object on a dark background. The size of the ring of 0's and the background in each template increase in proportion to the template size. In addition, the shape of the template is circular instead of square so as to increase the sensitivity of the match. By using a circular template, all the neighboring pixels which locate evenly around the tumor are checked in the matching process.

```
          -1 -1 -1
       -1 -1  0 -1 -1
    -1 -1  0  1  0 -1 -1
 -1 -1  0  1  1  1  0 -1 -1
 -1  0  1  1  1  1  1  0 -1
 -1 -1  0  1  1  1  0 -1 -1
    -1 -1  0  1  0 -1 -1
       -1 -1  0 -1 -1
          -1 -1 -1
```

Fig. 1. A tumor-like template for matching with tumors of five pixels in diameter.

B. Similarity Measure

To measure the similarity between a true suspicious area and the template, we need a similarity measure which is not sensitive to average brightness. In addition, the measures produced by templates of different size must be comparable. Hence, the normalized cross-correlation measure [12] is used.

Let S be the image, an $L \times L$ array of pixels, each taking one of K gray levels, and let W be the $M \times M$ template with $M \ll L$. Each $M \times M$ subimage of S can be uniquely referenced by its upper left corner coordinates (i, j) and there are $(L - M)(L - M + 1)$ such subimages. The normalized cross-correlation measure is de-

fined as

$$R(i,j) = \frac{\sum_{k=1}^{M}\sum_{m=1}^{M}\left\{\left(W(k,m)-\mu_w\right)\left(S(i+k-1,j+m-1)-\mu_s(i,j)\right)\right\}}{\sqrt{\sum_{k=1}^{M}\sum_{m=1}^{M}\left(W(k,m)-\mu_w\right)^2\sum_{k=1}^{M}\sum_{m=1}^{M}\left(S(i+k-1,j+m-1)-\mu_s(i,j)\right)^2}}$$

where μ_w is the mean of the template and μ_s is the mean of the subimage centered at image point (i, j).

As mentioned earlier, one disadvantage of using template matching is that the orientation and size of the target object may disturb the match. However, to the extent that the tumor area is circular, the match is orientation invariant. To cope with size variations of the object, multiple templates were used, with the circle radius ranging from 3 to 14 pixels.

C. Criteria in Selecting Suspicious Areas

The template matching operation produces 12 output images (template radii 3–14 pixels) in which each pixel value is the result of cross correlating the template and the subimage centered at that point. An appropriate method is required to interpret these values so that all suspicious areas are detected and nonsuspicious areas are excluded.

Due to the design of the templates, a circular dark object on a bright background will produce a large negative cross-correlation value. Hence, locations in the images which have negative cross-correlation values are not considered suspicious areas.

An effective criterion for selecting suspicious areas must be able to solve three problems. First, most suspicious areas have the maximum cross-correlation value when being matched with one of the 12 templates, but we have no guarantee that this is always true. Second, we do not have prior knowledge of the size and the number of tumors in a mammogram film. Third, some mammograms have a very rich image texture due to the presence of glands and fatty tissue, leading to high cross-correlation values in the template matching stage. Some of these values may even be larger than those for some suspicious areas in other images. Consequently, a single global threshold for the template matching stage is bound to miss tumors in many images and to produce large numbers of false alarms in other images.

Alternatively, one can use a percentile method [1] and classify a fixed percentage of locations as suspicious. This (possibly large) set of suspicious areas can then be analyzed with further tests to reduce the number of false alarms. The fixed percentage should be chosen in such a way as to have no misses (i.e., no tumor or suspicious area should be rejected) and to have a reasonably small number of false alarms so as to keep the amount of computation spent on false alarm tests within reasonable limits. Hence, by analyzing the normalized cross-correlation distribution of an image, we choose a threshold value R such that q percent of the locations in the image having a correlation larger than R are considered as suspicious areas. This corresponds to mapping q percent of the locations into suspicious areas. In our implementation, we chose $q = 2.5$ percent as this value leads to no misses and a reasonably small number of false alarms. The choice of q primarily affects the number of locations that are subjected to false alarm tests and only to a minimal extent the number of locations that are classified as suspicious after the false alarm tests.

The template matching process only considers local information; therefore, it cannot adjust to the global texture of each image. The percentile method improves the template matching step by taking into account global image information. In general, if many locations in an image (with rough texture) produce large cross-correlation values, a large threshold will be selected to minimize the number of nonsuspicious areas being considered as suspicious areas. In the case of an image with smooth texture, a smaller threshold is used to ensure the detection of suspicious areas. The cross-correlation distributions for the images in Plates 5 and 6 are shown in Figs. 2 and 3, respectively. Note that in using a q value of 2.5 percent, a large threshold (0.67) was selected for the image with rough texture (Plate 6), while a smaller threshold (0.58) was selected for the smoother image (Plate 5).

The tumor detection method is applied to mammogram images that are prefiltered using the selective median filter, although it is possible to apply it directly to the original images. Experiments showed, however, that direct application leads to a significant degradation in performance, with a higher number of misses and false alarms, and thus that the noise reduction achieved by the prefiltering process is essential for the success of the subsequent tumor detection stage.

IV. False Alarm Tests

The percentile method produces a set of candidate points which are not exclusively suspicious areas. We therefore need some selection methods to remove as many false alarms as possible. One important requirement of a false alarm test is that they must not respond to true suspicious areas to avoid degrading the overall performance. After analyzing the characteristics of the many false alarm areas, two tests were designed. They examine a mammogram image at the locations reported as suspicious by the template matching procedure and try to discriminate false alarms from true suspicious areas.

A. Neighborhood Test

This test is based on the observation that the template matching typically produces a sharp cross-correlation peak

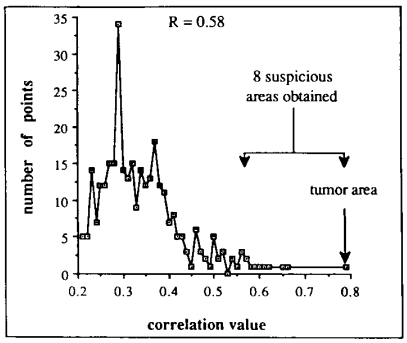

Fig. 2. A distribution curve of normalized cross-correlation values computed for mammogram image 1. Using a q value of 2.5, the threshold R of this image is 0.58.

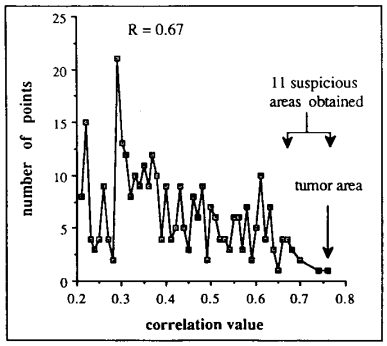

Fig. 3. A distribution curve of normalized cross-correlation values computed for mammogram image 2. Using a q value of 2.5, the threshold R of this image is 0.67.

in false alarm areas whereas cross correlation drops more gradually in the immediate neighborhood of tumors. Recall that the boundary of a suspicious area is fuzzy and the shape of a suspicious area is circular. Therefore, a high match between a template and a suspicious area will not be confined to a single pixel centered at the suspicious area. Instead, the pixels in the immediate neighborhood of the center pixel will also produce high cross-correlation value when being matched with a template having the same size as the suspicious area. On the other hand, in false alarm areas where a bright, homogeneous, and circular area does not exist, the cross-correlation value may fall off quickly away from the center. This is illustrated in Figs. 4 and 5 which show the cross-correlation values computed around the center of a suspicious area and a false alarm area.

As mentioned earlier, some false alarms are due to an artificial phenomenon created by the filtering process. The neighborhood test also aims at removing this type of false alarm by applying the test to the original image instead of the filtered image. This is based on the belief that when the matching is applied to the original image, a cluster of high cross-correlation values should not exist in these false alarm areas because they should appear as normal tissue in the unfiltered image.

16	26	35	38	36	31	24	17	11	08	03
12	24	34	41	42	39	33	26	20	14	08
07	21	34	42	47	47	41	34	27	19	12
05	15	32	45	53	55	48	42	32	23	16
01	13	30	46	54	60	54	47	38	26	20
01	13	29	43	55	62	58	50	43	32	27
02	11	24	38	49	54	52	47	44	38	34
04	14	22	30	39	44	45	43	43	42	40
04	14	22	28	29	33	32	33	39	40	40
02	10	17	22	23	24	21	24	31	35	38
00	06	13	17	19	18	17	20	25	28	32

(a)

23	26	28	37	39	40	39	36	32	28	24
29	32	37	40	42	43	40	36	30	26	23
29	33	37	40	42	42	40	35	29	24	20
27	31	34	38	40	40	38	34	28	22	18
25	28	30	34	36	38	36	32	43	34	27
24	26	29	30	33	61	59	53	44	38	32
25	27	28	30	30	56	56	49	43	39	36
26	27	29	29	29	51	50	45	39	39	38
27	28	28	28	27	46	44	40	37	36	37
25	28	28	27	25	23	38	38	36	35	35
22	25	27	27	25	22	34	34	33	34	34

(b)

Fig. 4. (a) The normalized cross-correlation values computed around the center of a suspicious area located in the image shown in Plate 5. (b) The normalized cross-correlation values computed around the center of a false alarm area located in the image shown in Plate 5.

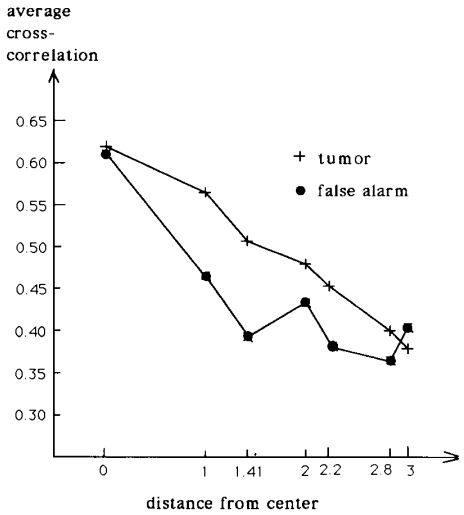

Fig. 5. Average cross correlation as a function of the distance from the cross-correlation peak for a tumor and a false alarm.

To implement this test, the immediate neighborhood of the center of a suspicious area is defined by a 3 × 3 cross-shaped window. Within this window, the original image is matched with the tumor template that produced the largest cross-correlation value in the template matching stage. If the cross-correlation value fall off sharply in the immediate neighborhood of the center of the suspicious area, the average cross-correlation value within the neighborhood window will be small. Otherwise, the average value will be large and, in fact, should be larger than the threshold R selected by the percentile method. Therefore, if the average cross-correlation value is less than the threshold R, the suspicious area is discarded as a false alarm.

B. Histogram Test

If we compute a gray-level histogram within a window containing both a tumor and background, we should ide-

ally obtain two peaks, with one peak corresponding to the tumor and one corresponding to the background. There is no reason to expect a double-peaked gray-level histogram in other areas, and therefore all image locations that do not show a double-peaked gray-level histogram are discarded as false alarms. This is the essential idea of the second false alarm test, based on the gray-level thresholding method described in [16].

The window has to be large enough to contain both tumor and background, and its size can be determined from the tumor radius reported by the template matching process, more precisely, a circular window with a radius 30 percent larger than the tumor radius is used for the gray-level histogram. The gray-level histogram for this window is then smoothed and the number of histogram peaks is determined. All locations with single-peak histograms are discarded as false alarm locations.

Two typical examples of smoothed histograms are shown in Figs. 6 and 7, one for a tumor location (Fig. 6), and one for a false alarm location (Fig. 7). The latter fails the histogram test since it is not double-peaked, and thus provides no evidence for the presence of a circumscribed mass.

V. Experimental Results

In comparing the diagnostic results produced by the computer and the radiologist, we cannot expect that the location and size of a suspicious area determined by the radiologist and the computer are exactly the same, so some relaxed criterion for determining a positive detection was established. A suspicious area was considered detected if the circle placed around a suspicious area by the radiologist and the nearest area mapped by the computer overlapped by at least 50 percent. If the overlap was less than 50 percent, it was counted as a miss. The result of this comparison based on the training set is shown in Table I which lists the size and location of the suspicious area(s) diagnosed by the radiologist along with the areas located by computer. Also included are the total number of suspicious areas reported in each test case.

The result of comparing the performance of the proposed method using different percentile thresholds in the template matching stage is shown in Table II. This table shows that as the percentile threshold q increases, both the total number of suspicious areas detected and the false alarm rate increase. On the other hand, although the false alarm rate is zero when a $q = 0.5$ is used, the hit rate drops to 54 percent. As a result, $q = 2.5$ was chosen for the percentile method since it produces a hit rate of 100 percent and a moderate false alarm rate of 1.1 per mammogram. In Plates 5 and 6, the suspicious areas reported by the proposed method are circled in the mammogram images 1 and 2.

The result of applying the proposed method to a test set of 17 images (SMF-enhanced images) is shown in Table III with a summary in Table IV. The hit rate in detecting suspicious areas remains 100 percent, that is, all suspicious areas detected by the radiologist were detected by

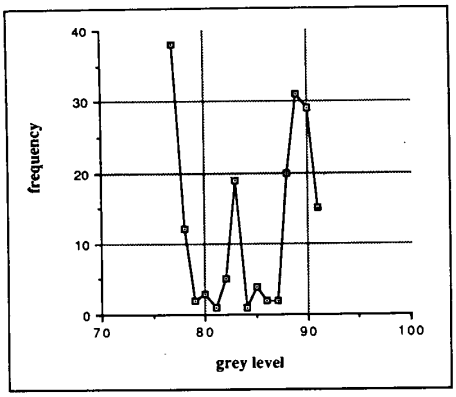

Fig. 6. Gray-level histogram constructed in a suspicious area located in the image shown in Plate 5.

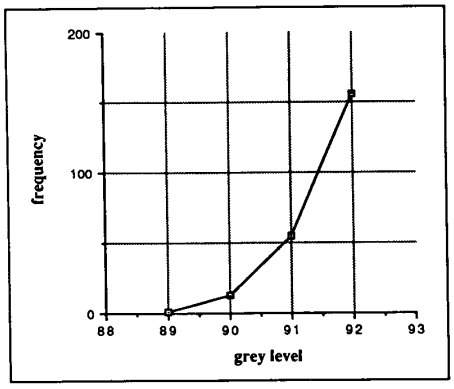

Fig. 7. Gray-level histogram constructed in a false alarm area located in the image shown in Plate 5.

TABLE I
Comparison of the Computer's and the Radiologist's Interpretation on Seven Mammograms in the Training Set ($q = 2.5$)

case number	Coordinates of suspicious areas						Total number of areas detected
	Radiologist			Computer*			
	X	Y	radius (mm)	X	Y	radius (mm)	
136801	170	116	6	166	116	5	2
	180	176	11	181	178	10	
151138	100	186	8	102	185	8	4
	150	93	10	150	96	9	
006356	186	208	8	184	206	6	2
	200	149	4	199	149	3	
63387a	34	165	4	35	165	4	1
119343	160	118	12	159	119	11	2
142322	188	192	9	187	194	8	5
89993a	134	128	8	135	128	8	3
	192	101	5	194	103	4	

*Coordinates of suspicious area most closely correlating with the radiologist.

TABLE II
Comparison of the Performance of the Proposed Method Using Different q Values

q	Hit-rate*	Total number of false alarm detected	Average number of false alarm per film
0.5	54%	0	0
2.5	100%	8	1.1
5.0	100%	47	6.7

*In the training set, 11 suspicious areas are reported by the radiologist.

TABLE III
COMPARISON OF THE COMPUTER'S AND THE RADIOLOGIST'S INTERPRETATION ON 17 MAMMOGRAMS IN THE TEST SET ($q = 2.5$)

	Coordinates of suspicious areas						
	Radiologist			Computer*			Total number of areas detected
case number	X	Y	radius (mm)	X	Y	radius (mm)	
89993b	171	140	6	174	140	6	5
	101	126	4	104	126	4	
124133a	122	134	6	122	136	5	4
	175	159	8	172	158	7	
124133b	130	42	10	131	44	10	3
	88	112	10	88	110	9	
129194a	90	150	6	92	149	7	3
129194b	198	70	4	200	71	5	4
145079	198	208	4	196	208	5	3
40274a	200	64	3	199	64	3	5
63387b	125	194	4	124	194	4	1
40274b	38	162	3	38	164	3	1
125758	118	68	7	118	66	9	1
069591	56	22	5	55	22	5	5
152506	192	164	10	192	164	8	1
004948	-	-	None	-	-	-	0
004703	103	71	4	104	70	5	7
129194	100	44	4	100	46	5	2
012914	138	76	13	140	78	12	2
933302	125	73	7	124	71	6	2

*Coordinates of suspicious area most closely correlating with the radiologist.

TABLE IV
SUMMARY OF EXPERIMENTAL RESULTS

Number of films tested	= 17
Total number of true suspicious areas	= 19
Average number of true suspicious areas per film	= 1.1
Hit-rate of the technique	= 100%
Average number of false alarms found per film	= 1.7

the proposed method and the false alarm rate increased slightly to 1.7 per mammogram compared to the 1.1 obtained with the training set.

VI. COMPUTATION CONSIDERATION

The computational cost of our method is high: about 15 min are required to complete an analysis of one mammogram on a VAX 11/780 computer. Most of the time is spent in the template matching stage which requires matching a mammogram with 12 templates of size 10^2–45^2.

One way to speed up the computation is to reduce the number of locations that are fully analyzed by using a coarse-to-fine template matching method [17], [22]. In the first stage of this method, a reduced-resolution template is matched with a reduced-resolution version of the image. Only those locations which produce a match value larger than a predetermined threshold are used in the second stage where a full resolution template and image are matched. If the coarse template matching is done at half the resolution, its computational cost is reduced by a factor of about 16 compared to the original "one-stage" template matching. The cost of the second stage depends on the threshold. If the threshold is high, the cost of the second stage is low, but some suspicious areas may have a correlation value smaller than the threshold, and therefore will not be used in the second stage matching. This is particularly true for small-size tumors having a low contrast with their surroundings. They are often smeared in low-resolution images, and thus produce correlation values smaller than the threshold in the first stage, leading to a higher miss rate. On the other hand, if the threshold is small, the saving in computational time will become insignificant.

Experiments showed a clear speed-accuracy tradeoff of the coarse-fine template matching method. Selecting a threshold value that saved about 30 percent in computation time produced a miss rate of 9 percent, i.e., the suspicious areas were not detected in two out of the 24 test images.

VII. CONCLUSION

This paper presents a method for breast tumor detection in mammograms. The first step towards tumor detection is image enhancement by noise removal. Several algorithms for noise cleaning in digital images were analyzed. Some techniques based on selective averaging were implemented due to their simplicity and their promising properties for edge preservation and noise removal. In addition, a new enhancement method, selective median filtering, was developed. A training set of seven mammogram images was used to adjust the parameter values used in this method. This technique was then applied to a test set of 17 mammograms in which suspicious areas were identified by a radiologist. The selective median filtering method was found to be more effective than the other techniques implemented in enhancing mammogram images.

The second step is concerned with tumor detection. Our method is based on template matching and is capable of detecting suspicious areas in mammograms independent of their size, orientation, and position. To reduce the false alarm rate, two false alarm tests were designed to examine the output of the template matching process. These tests can discriminate most noise areas from true suspicious areas. The routines were applied to 17 test mammogram images, and the results show that the routines can correctly identify all suspicious areas while producing only a small false alarm rate of 1.7 per mammogram image.

One disadvantage of our method is the high computational cost. A coarse-fine template matching method was implemented, and the results show that the computational cost of the proposed method can be reduced by such an approach.

The results obtained with our method are quite encouraging. By combining three criteria, namely, the contrast, the uniform density, and the circular shape of tumor areas, the detection algorithm is capable of locating all tumor areas in the test of 24 images. Nevertheless, to test the stability of this method, more testing is required using a larger number of cases.

As was discussed in the Introduction, expert radiologists use the following four criteria for the detection of circumscribed masses: 1) brightness, 2) brightness homogeneity, 3) circular shape of various size, and 4) fuzziness of boundaries. The first three criteria were used in

the design of the image enhancement and tumor detection stage; the fourth criterion was used in the design of the neighborhood test for eliminating false alarms. Performance of the system may be further improved by directly analyzing the boundaries of suspicious regions with respect to their fuzziness similar to the analyses proposed in [2].

The method presented in this paper is successful in detecting one type of breast cancer, circumscribed masses. To assess its clinical utility, it has to be tested on much larger sets of data. Furthermore, a clinically useful system for breast cancer detection must be able to detect all types of breast cancers, not just circumscribed masses. Work is currently under progress on a system that will be able to detect all major types of breast cancers.

REFERENCES

[1] L. V. Ackerman and E. Gose, "Breast lesion classification by computer and xeroradiograph," *Cancer*, pp. 25-1035, 1972.

[2] W. F. Bischof and T. M. Caelli, "Parsing scale-space and spatial stability analysis," *Comput. Vision, Graphics, Image Processing*, vol. 42, pp. 192-205, 1988.

[3] A. C. Bovik, T. S. Huang, and D. C. Munson, "The effect of median filtering on edge estimation and detection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-9, pp. 181-194, Mar. 1987.

[4] ——, "A generalization of median filtering using linear combinations of order statistics," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 1342-1350, 1983.

[5] L. S. Davis and A. Rosenfeld, "Noise cleaning by iterated local averaging," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-8, pp. 705-710, Sept. 1978.

[6] A. P. Dhawan, G. Buelloni, and R. Gordon, "Enhancement of mammographic features by optimal adaptive neighborhood image processing," *IEEE Trans. Med. Imaging*, vol. MI-5, pp. 8-12, Mar. 1986.

[7] N. C. Gallagher and G. L. Wise, "Theoretical analysis of the properties of median filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 1136-1141, Dec. 1981.

[8] R. Gordon and R. M. Ranagayyan, "Feature enhancement of film mammograms using fixed and adaptive neighborhoods," *Appl. Opt.*, vol. 23, no. 4, pp. 560-564, 1984.

[9] E. L. Hall, R. P. Kruger, S. J. Dwyer *et al.*, "A survey of preprocessing and feature extraction techniques for radiographic images," *IEEE Trans. Comput.*, vol. C-20, pp. 32-1044, Sept. 1971.

[10] W. Hand, J. L. Semmlow, L. V. Ackerman *et al.*, "Computer screening of xeromammograms: A technique for defining suspicious areas of the breast," *Comput. Biomed. Res.*, vol. 12, pp. 445-460, 1979.

[11] F. Kuhlmann and G. L. Wise, "On second moment properties of median filtered sequences of independent data," *IEEE Trans. Commun.*, vol. COM-29, pp. 1374-1379, 1981.

[12] X. Li, C. Shanmugamani, T. Wu *et al.*, "Correlation measures for corner detection," in *Proc. Comput. Vision Pattern Recognition*, 1986, pp. 643-647.

[13] J. E. Martin, *Atlas of Mammography, Histologic and Mammographic Correlation*. Baltimore/London: Williams and Wilkins, 1982, pp. 65-101.

[14] M. Nagao and T. Matsuyama, "Edge preserving smoothing," *Comput. Graphics Image Processing*, vol. 9, pp. 394-407, 1979.

[15] T. A. Nodes and N. C. Gallagher "Median filters: Some modifications and their properties," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, Oct. 1982.

[16] A. Rosenfeld and A. C. Kak, *Digital Picture Processing*, vol. 1 and 2, 2nd ed. New York: Academic, 1982.

[17] A. Rosenfeld and G. J. Vanderfrug, "Coarse-fine template matching," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-7, pp. 4-107, Feb. 1977.

[18] A. Scher, F. R. D. Velasco, and A. Rosenfeld, "Some new image smoothing techniques," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-10, no. 3, pp. 153-158, 1980.

[19] R. A. Stein, "A novel median filter for edge enhancement in CT head scans," in *Proc. 8th IEEE Conf. Eng. Med. Biol. Soc.*, vol. 2, Dallas, TX, Nov. 1986, pp. 1126-1129.

[20] P. Strax, "Mass screening of asymptomatic women," in *Breast Cancer, Diagnosis and Treatment*, I. M. Ariel and J. B. Cleary, Eds. New York: McGraw-Hill, 1987, pp. 145-151.

[21] L. Tabar and P. B. Dean, "Basic principles of mammographic diagnosis," *Diag. Imaging Clin. Med.*, vol. 54, pp. 146-157, 1985.

[22] G. J. Vanderbrug and A. Rosenfeld, "Two-stage template matching," *IEEE Trans. Comput.*, vol. C-26, pp. 384-393, Apr. 1977.

[23] F. Winsberg, "Detection of radiographic abnormalities in mammograms by means of optical scanning and computer analysis," *Radiology*, vol. 89, pp. 211-215, 1967.

[24] H. C. Zuckerman, "The role of mammograph in the diagnosis of breast cancer," in *Breast Cancer, Diagnosis and Treatment*, I. M. Ariel and J. B. Cleary, Eds. New York: McGraw-Hill, 1987, pp. 152-172.