Joint Conference of 1996 International Computer Symposium
December 19~21, Kaohsiung, Taiwan, R.O.C.

A Computerized System for Detection and Segmentation of Clustered Microcalcifications

Chien-Shun Lo[1] Pau-choo Chung[1,*] Chein-I Chang[2] San Kan Lee[3,4] Chia-Hsien Wen[5]

[1]Department of Electrical Engineering, National Cheng Kung University
Tainan, Taiwan 70101, Republic of China

[2]Department of Computer Science and Electrical Engineering
University of Maryland Baltimore County
Baltimore, MD 21228-5398, U. S. A.

[3]Department of Radiology, Taichung Veterans General Hospital, VACRS
Taichung, Taiwan 40705, Republic of China

[4]Department of Diagnostic Radiology, National Defense Medical Center
Taipei, Taiwan 100, Republic of China

[5]Computer Center, Taichung Veterans General Hospital, VACRS
Taichung, Taiwan 40705, Republic of China

* corresponding author: Email: pcchung@eembox.nuku.edu.tw
Tel/Fax: 886-6-2757575 ext 62373/886-6-2345482

Abstract

Over the past years, there have been extensive studies in developing effective CAD techniques for the detection of microcalcifications. Unfortunately, most of them require manually locate suspicious areas, called regions of interest (ROIs) prior to detection. In this paper, we present a three-stage computer aided diagnostic (CAD) method which not only automatically locates ROIs but also detects and segments clustered microcalcifications. The first stage is to extract the breast from the background followed by a second stage to detect and locate the suspicious areas, and finally the segmentation of clustered microcalcifications are carried out in the third stage. Performance of this method is evaluated by skillful radiologists based on a database developed by Nijmegen University Hospital and mammograms obtained from Taichung Veterans General Hospital. The initial assessment of the proposed system is very encouraging.

Keywords: computer-aided diagnostic (CAD) method; detection; fractal dimension; mammograms; microcalcifications; region of interest (ROI); segmentation

1. Introduction

Breast cancer has become a second leading cause of cancers among women in Taiwan. Microcalcifications generally provide an indication of an early sign of breast cancer. Past screening studies showed that 90% of nonpalpable in situ ductal carcinomas and 70% of nonpalpable minimal carcinomas were visible on microcalcifications alone [1]. As a result, detecting nonpalpable malignant calcifications within the breast before they become metastasized has significant impact on reduction of breast cancer. Thus far, mammography screening still remains to be the only effective technique for early breast cancer detection and cannot be replaced by other diagnostic modalities such as sonography, thermography and MRI [2].

Calcifications are tiny clustered particles and probably the smallest structures which can be identified on mammograms and best visualized using high-resolution imaging techniques or direct radiological magnifications. Early reports from 1960s suggested that clustered microcalcifications associated with benignancy and malignancy usually have distinct characteristics. Even so, it was found that recognizing the presence of these tiny particles from mammograms has been a great challenge for radiologist..

Over the past years, there have been extensive studies in developing effective CAD techniques for the detection of microcalcifications [3-15]. Unfortunately, most of them require manually locate suspicious areas, called regions of interest (ROIs) prior to detection. Strictly speaking, they are not really automatic systems. In this paper, an automatic method to select ROIs is proposed. In order to select ROIs automatically, the

breast region is firstly extracted from a mammogram. In the past, the extraction of the breast from a mammogram has not received attention partly because it is easy to visually identify the breast area from mammograms and the breast extraction is not important from the microcalcification detection point of view. However, from an automatic system point of view, all processes must be computerized and any manual process can cause system slow-down and delay which result in inefficiency. This is particularly true if a mammogram contains a small breast in which case, a large amount of memory will be wasted to store the breast background. So, it is highly desirable to store only the breast area rather than the whole mammogram so as to save computer memory and increase the CAD efficiency.

In this paper, we present a computerized CAD system shown in block diagram 1 which processes mammograms automatically in three stages: breast extraction, detection and location of suspicious areas, and segmentation of clustered microcalcifications. The breast extraction is an initial step of identifying ROIs. It allows the search for ROIs limited to the breast area and avoids unnecessary background processing resulting from direct search for ROIs. The breast is extracted by a process, called block region growing. The detection and location of clusters is first to detect microcalcifications by using fractal-Sobel filtering and then locates suspicious areas by using density window filtering. This fractal-Sobel filter is accomplished by using Sobel operator coupled with the fractal dimension. Finally, the segmentation of clustered microcalcifications is carried out in the last stage by using local thresholding. The proposed system is evaluated by experienced radiologists based on a database developed the Nijmegen University Hospital, Netherlands, and mammograms in Taichung Veterans General Hospital.

This paper is organized as follows. Section 2 describes an algorithm for block region growing to extract the breast area. Section 3 develops a series of preprocessing techniques, fractal-Sobel filter to detect microcalcifications, and location of suspicious areas by using density window filter. Section 4 designs a feasible local thresholding to segment clustered microcalcifications. Section 5 includes experiments and the system performance evaluation. Section 6 draws a brief conclusion.

2. Automatic breast extraction

The breast area in a mammogram generally represents inhomogeneous regions which have high gray-level intensities and variances. On the other hand, the background of the breast tends to have low and smooth gray levels distributions. Based on this mammographic characteristics, we can separate a mammogram into two classes, one consisting of pixels with high gray levels and variances which corresponds to the breast and the other containing pixels with low gray levels and variances which will be assigned to the background. An algorithm designed for such class separation is given as follows.

Automatic Breast Extraction Algorithm

Step 1: Divide a mammogram denoted by I shown in Figure 1 into blocks with the same size of $m \times n$.

Step 2: Block region growing procedure

Let $B = \{b_{11}, b_{12}, \cdots, b_{pq}\}$ be the set of blocks resulting from Step 1 where b_{ij} is the (i, j) th block in B and pq is the total number of blocks in B.

For each b_{ij}, compute the mean

$$\bar{b}_{ij} = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} b_{ij}(x, y) \quad (1)$$

and the variance

$$\sigma^2_{b_{ij}} = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} |b_{ij}(x, y) - \bar{b}_{ij}|^2 \quad (2).$$

Step 3: Compute the energy function defined by

$$E_{b_{ij}} = |\bar{b}_{ij}|^2 + |\sigma^2_{b_{ij}}| \quad (3)$$

The block region growing procedure begins with the block with smallest energy, say b_M, then grow b_M using the 4-neighbor connectivity with prescribed threshold value T. The result is shown in Figure 2.

Step 4: K-means clustering-based thresholding

Although the breast area extracted by Step 3, denoted by *BREAST* is an estimate not necessarily the desired breast area, we can use a K-means clustering-based thresholding to improve this estimate on the basis of *BREAST*. Let *MaxT* and *MinT* be the largest and the smallest gray level means, respectively, of blocks in *BREAST* and compute three seed gray level values as follows.

$$S_1 = \frac{MaxT + MinT}{3} \quad (4)$$

$$S_2 = \frac{2 \times (MaxT + MinT)}{3} \quad (5)$$

$$S_3 = MaxT. \quad (6)$$

Let *MaxB* and *MinB* be the largest and the smallest gray level means of the background. A fourth seed gray level value based on the background can be calculated by

$$S_0 = \frac{MaxB + MinB}{2}. \quad (7)$$

Applying K-means clustering algorithm to the four seed gray level values generated by Eqs. (4)-(7) produces four classes $\{C_0, C_1, C_2, C_3\}$ and each class C_i corresponds to S_j where $\{C_1, C_2, C_3\}$ and C_0 represents the breast area and the background respectively. The gray level separating C_0 from C_1 will be chosen as the optimal threshold and the resulting thresholded image is the desired binary image.

Step 5: Smoothing processing

The image generated by Step 4 still contains scattering pixels which may be misclassified into breast pixels. The main goal of this step is to clean up these.

An example of a reference image generated by Step 5 is shown in Figure 3 where the reference image of the breast identified by the automatic extraction algorithm is delineated by the white line.

3. Detection and location of suspicious areas

Two major mammographic characteristics usually used for the detection of microcalcifications are gray-level intensity and contrast [10]. In general, the pixels representing calcified tissues have high gray-level intensity as well as high contrast compared to their surrounding background. However, one does not necessarily imply the other. For example, it may be the case that calcifications with high intensity have low contrast due to strong background noises. More importantly, these two criteria are always offset by various background interferences such as random noise, structured noise caused by tissue textures and background. An additional degradation is the spatial resolution of a digitizer and digitization effects. All these factors make it difficult to detect and segment microcalcifications.

3.1 Enhancement

In this section, three enhancement preprocessing techniques proposed in [16] will be used to remedy these disadvantages. They are gradient enhancement, mean contrast enhancement and Gaussian filtering processes which can be briefly described as follows. However, unlike [16] whose preprocessing was applied to selected ROIs, the preprocessing used in this paper only applies to those pixels which have been identified to be breast pixels by the reference image generated by the automatic breast extraction algorithm.

(a) *Gradient Enhancement*

As mentioned previously, one of major mammographic characteristics in calcifications is high intensity in gray level. It is often the case that the surrounding neighborhood of calcified pixels generally have relatively higher gray level intensities. In order to separate the calcified pixels from the background, a gradient enhancement technique is employed to increase the gradient values of these pixels so that they can be enhanced to become more visible. This stage is carried out by the following procedure.

$$\begin{aligned} I_1(x,y) &= \textit{gradient enhancement}\{I(x,y)\} \\ &= g_1(x,y) + I(x,y) \end{aligned} \quad (8)$$

where $g_1(x,y)$ is the gray level of the gradient of the pixel (x,y) calculated based on a 3×3 window mask and given by

$$g_1(x,y) = \frac{1}{9}\sum_{i=-1}^{1}\sum_{j=-1}^{1} |I(x+i,y+j) - I(x,y)|. \quad (9)$$

(b) *Mean Contrast Enhancement*

Basically, the objective of gradient enhancement is to increase the gray levels of calcified pixels. However, while the gradients of these pixels are enhanced, the background noise pixels may be also enhanced in some areas as well, particularly, in their surrounding edges of calcified pixels. In this case, their relative contrast may be low and still remains unchanged even after the gradient values are enhanced. Since contrast is another major mammographic criterion in detection of microcalcifications, increasing relative contrast certainly improves the detection rate. The mean contrast enhancement is to scale down the gray level of each pixel by a normalized local mean based on the following equation.

$$\mu^m(x,y) = \frac{1}{9}\sum_{i=-1}^{1}\sum_{j=-1}^{1}\mu^{m-1}(x+i,y+j) \quad (10)$$

where $\mu^m(x,y)$ is the mth mean image obtained by averaging the original image $I(x,y)$ m times over a 3×3 window mask with central pixel (x,y). For example, $\mu^1(x,y) = \frac{1}{9}\sum_{i=-1}^{1}\sum_{j=-1}^{1}\mu^0(x+i,y+j)$ for

any pixel (x,y) with $\mu^0(x,y) = I(x,y)$. The processed image becomes

$$I_2(x,y) = \frac{\mu^m(x,y)}{L-1} I_1(x,y). \quad (11)$$

(c) *Gaussian Filtering*

As a result of (11), the relative contrast of calcified pixels to noise pixels will be increased, a Gaussian filter G_σ with a proper choice of variance σ^2 is then applied to I_2 in (4) to eliminate the noises at the expense of a small blurring effect on I_2. This is accomplished by subtracting a Gaussian filtered image, $G_\sigma \otimes I_2$ from the image I_2 as given below

$$I_3(x,y) = I_2(x,y) - G_\sigma \otimes I_2(x,y). \quad (12)$$

The images I_1, I_2 and I_3 resulting from the above sequence of enhancement processing given by (8), (11) and (12). The enhanced image is shown in figure 4..

3.2 Fractal-Sobel filter

In Section 3.1 we have used a sequence of enhancement techniques to detect and locate suspicious areas. Since only clustered microcalcifications will be considered to be possibility of malignancy, in this section, we will use a Sobel operator coupled with fractal dimension to identify microcalcifications. Due to the fact that clustered microcalcifications generally possess high gradients, we apply a Sobel operator to image I_3 to generate a gradient image I_g and then use the fractal dimension-generated indices as a thresholding criterion to detect the microcalcifications. This is different from the method proposed in [16] which used local entropy thresholding to achieve detection of microcalcification. The technique to be used for calculation of the fractal dimension is called *Blanket method* which can be briefly described as follows [19][20].

Blanket Method

Peleg et al.[19] extended the blanket method, described by Mandelbrot[20], to three dimensions. The surface area [i,j,f(i,j)] of an object can be estimated by measuring the volume between an upper blanket $u_r(i,j)$ and a lower blanket $b_r(i,j)$ which not further than a distance r above and below the surface to be measured. The upper and lower blanket are defined, respectively, as follows:

$$u_r(i,j) = \max\{u_{r-1}(i,j)+1, \max_{|(k,l)-(i,j)\le 1|} u_{r-1}(k,l)\} \quad (13)$$

$$b_r(i,j) = \max\{b_{r-1}(i,j)-1, \max_{|(k,l)-(i,j)\le 1|} b_{r-1}(k,l)\} \quad (14)$$

where $u_0(i,j) = b_0(i,j) = f(i,j)$. Peleg et al. Define the area as half of the volume increase

$$A(r) = \tfrac{1}{2}\sum_{i,j}\{[u_r(i,j) - u_{r-1}(i,j)] + [b_{r-1}(i,j) - b_r(i,j)]\} \quad (15)$$

Using the relation

$$A(r) \propto r^{2-D}$$

This relation can be represented as Eq. (16).

$$A(r) = k * r^{2-D} \quad (16)$$

Using Eq. (16), we can calculate the fractal dimension D based on log(A(r)) versus log(r). Generally speaking, The area A(r) always represents the variance of any fixed size of regions and D represents the similarity of different images. Different images with the similar variance would produce the same fractal Dimension D. The higher is A(r), the higher is variance. The higher is variance, but the lower is D. The variance of regions, such as regions which contain clustered microcalcifications, usually have high gradient and high variance. Therefore, D is an important index to identify the regions of clustered microcalcificaitons. Then, microcalcifciations are filtered by a fractal threshold in this gradient image I_g.

3.3 Density window filter

Although the microcalcifications are produced by using Fractal-Sobel Filter in Section 3.2, only clustered microcalcifications show high correlation with benignancy and malignancy. Therefore, it is necessary to locate the clustered microcalcifications and eliminate unnecessary microcalcifications. Density window filtering is a method used to detect the location and the scope of clustered microcalcifications in this paper. A fixed-size window is selected to move around the breast region pixel by pixel. Let N(i,j) be the number of microcalcifications within a window with the center located at (i,j). W is the window size. The density of a window DW is defined as

$$DW = N(i,j)/W \quad (17)$$

The densities of windows are then used for identifing the location of clustered microcalcifications. The high density window usually occur on the clusters of microcalcifications. Fig. 7 shows a cluster which consists of several windows. Under the processing of density window filtering, scattered microcalcification

would be eliminated and are not added to compute density.

4. Segmentation of clustered microcalcifications

Though from the previous procedures, microcalcifications and their clustering structures are obtained. However, by the previous steps the obtained results are still rough, since gradient plays the major role in the segmentation. Hence in this section the thresholding technique is further conducted on each clustered regions to obtain more accurate microcalcifications. As we know, microcalcifications under different background tissues would have different intensities. Therefore, it is impossible to use a global threshold to segment all microcalcifications on a mammogram. In our approach, a local threshold is selected from each clustered microcalcification region. The local threshold is computed from the average grey level value of each clustered microcalcification region. Based on the obtained threshold criterion, fine segmentation within this clustered regions is conducted. This procedure will be repeated from one clustered region to another until all the clustered microcalcification regions are perfomed.

5. Experiments and system evaluation

In this section, we use the following databases to evaluate our proposed CAD system.

5.1 Nijmegen database

40 mammograms in the database developed by the Department of Radiology, Nijmegen University Hospital, Netherlands are from 21 patients and of size 2048×2048 with 12 bit gray level resolution digitized by an Eikonix 1412 12-bit CCD camera with a fixed calibration. Optical density 0.18 corresponds to the maximum output level (4095). A sampling aperture of 0.5 mm in diameter and 0.1 mm sampling distance were used. All mammograms were corrected for inhomogeneity of the light source (Gordon plannar 1417) and recorded by a Kodak MIN-R/SO177 screen/film combination using various types of equipments. All mammograms show one or more clustered microcalcifications.

5.2 Mammograms in Taichung Veterans General Hospital

This database obtained from Taichung Veterans general Hospital consists of 50 patient cases with 75 mammograms. These mammograms were biopsy proven and taken from two views cephalocaudal (CC) and medicalateral oblique (MLO). All the mammograms were digitized by a film digitizer Truvel made by Division of Vidar System Corporation with 260 DPI (approximately equal to spatial resoluation 0.1 mm/pixel) and the software TruScan PC. Version. 6.3 and processed on a PC486/66 Hz with programs written in C language.

The above two databases are used to evaluate the proposed system. The detection rates are measured based on the average true positive (TP) and the average false positive (PF), which are obtained as 94% and 16%, respectively. For an illustrative purpose, an experimental result is shown in Figures 1-7, demonstrating a sequence of implementing the proposed computerized system for detection and segmentation of clustered microcalcifications. Figure 1 is the original mammogram. Figure 2-3 is the reference image generated by the automatic breast extraction algorithm. In order to demonstrate the clustered microcalcifications clearly, an ROI region detected from our method is cut and shown in Figure 4. Figures 5 is an enhanced image obtained by three enhancement preprocessing techniques. Figure 6 is the fractal-Sobel filtering image. Finally, Figures 7 produce the results of clustered microcalcifications by using density window filtering.

6. Conclusions

In this paper, a computerized CAD system for the detection and segmentation of clustered microcalcifications is presented. Unlike most of CAD methods which ROIs are selected manually, the proposed system can automatically select the areas of interest and identify regions which will be considered to be clustered microcalcifications. It then applies a series of preprocessing procedures to enhance and smooth the selected ROIs for segmentation. The clustered microcalcifications are detected from the background using a fractal-Sobel operator, identify clusters by using density window filtering, and finally are segmented by using local cluster thresholding which can be considered as local thresholding processing. The system is evaluated by Nijmegen database and mammograms in Taichung Veterans General Hospital and proofread by skillful radiologists. The average result which are evaluated by Nijmegen database are clusters TP/per image and clusters FP/per image.

REFERENCES

[1] Feig, S.A. and Galkin, B.M. Breast microcalcifications: early warning for cancer?. Diagnostic Imaging. 54: 132-138; 1990.

[2] Bassett, L.W. Mammographic ananlysis of microcalcifications. Radiological Clinics of North America. 30(1): 132-138; 1992.
[3] Sickles, E.A. Breast calcifications: mammographic evaulation. Radiology. 160: 289-293; 1986.
[4] Lanyi, M. Microcalcifications in the breast-a blessing or curse?. Diagnostic Imaging Clin. Med. 54: 126-145; 1985.
[5] Kegelmeyer, M.P., et al. Computer-aided mammographic screening for spiculated lesions. Radiology. 191: 331-337; 1994.
[6] Fam, B.W., et al. Algorithm for the detection of fine clustered calcifications on film mammograms. Radiology. 169: 333-337; 1988.
[7] Davies, D.H.; Dance, D.E. Automatic computer detection of clustered calcifications in digital mammograms. Phys. Med. Biol. 35(8): 1111-1118; 1990.
[8] Chan, H.P., et al. Image feature analysis and computer-aided diagnosis in digital radiology. I. Automated detection of microcalcifications in mammography. Med. Phy. 14(4): 538-548; 1987.
[9] Chan, H.P., et al. Improvement in radiologists' detection of clustered of microcalcifications on mammograms. Invest. Radiol. 20: 1102-1110; 1990 14(4): 538-548; 1987.
[10] Dengler, J. et al. Enhancement of chest and breast radiographs by automatic spatial filtering. IEEE Trans. on Medical Imaging. 10(3): 330-335; 1991.
[11] Tahoces, P.G., et al. Segmentation of microcalcifications in mammograms. IEEE Trans. on Medical Imaging. 12: 634-642; 1993.
[12] Dhawan, A.P., et al. Analysis of mammographic calcifications using gray-level image structure features. IEEE Trans. on Medical Imaging. 13(2): 263-274; 1994.
[13] Morrow, W.M., et al. Region-based contrast enhancement of mammograms. IEEE Medical Imaging, 11(3): 392-406; 1992.
[14] Shen, L., et al. Applications of shape analysis to mammographic calcifications. IEEE Medical Imaging, 13(2): 263-274; 1994.
[15] Laine, A.F. Mammographic feature enhancement by multiscale analysis. IEEE Med. Imag. 13(4): 725-740; 1994.
[16] Hsu, B.-C. et al. An -automated system for detection and classification of microcalcifications in digital mammograms. Proc. CVGIP'96, Taiwan, ROC.
[17] Chaudhuri, B.B. et al. Texture segmentation using fractal dimension. IEEE Trans. Pattern Analysis and Machine Intelligence. 17(1): 72-79; 1995.
[18] Chung, P.C. et al. An algorithm for detection and segmentationof clustered microcalcifications on mammograms. 2nd Medical Engineering Week of the World. May 26-30, Taipei, Taiwan, ROC. p. 102; 1996
[19] S. Peleg, J. Naor, R. Hartley, and D. Avnir, "Multiple resolution texture analysis and classification," IEEE Trans. Pattern Anal. Machine Intell. 6, 518-523 (1984).
[20] B. B. Mandelbrot, The Fractal Geometry of Nature (Freeman, New York, 1977).

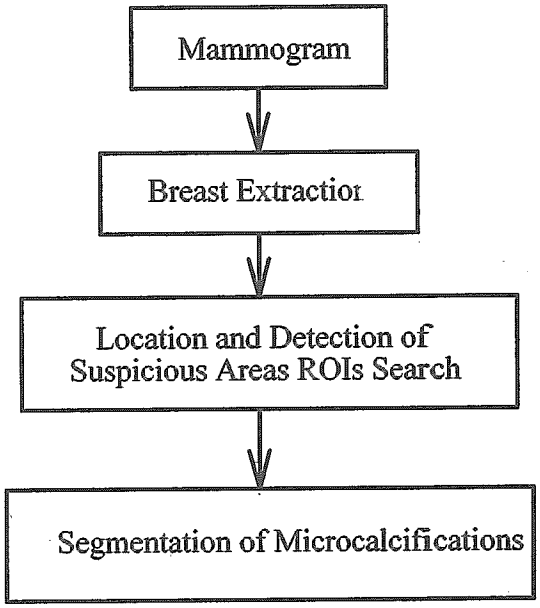

Diagram 1: Block diagram of segmentation of microcalcifications.

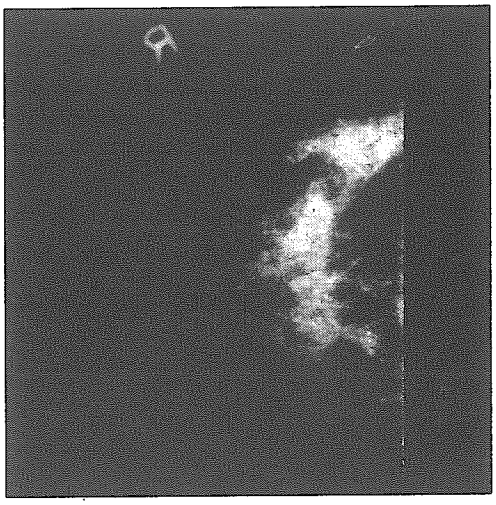

Fig. 1 The original mammogram

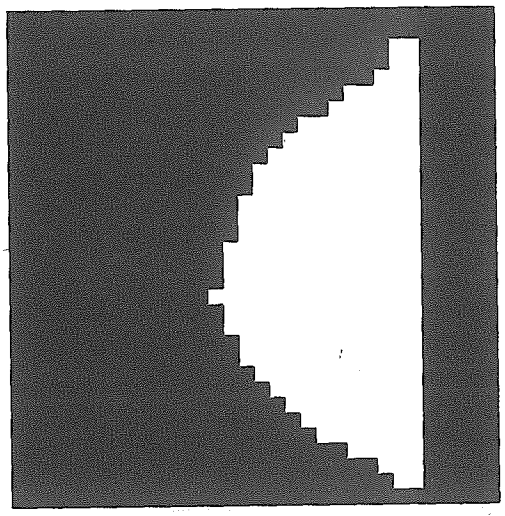

Fig. 2 The block region growing result

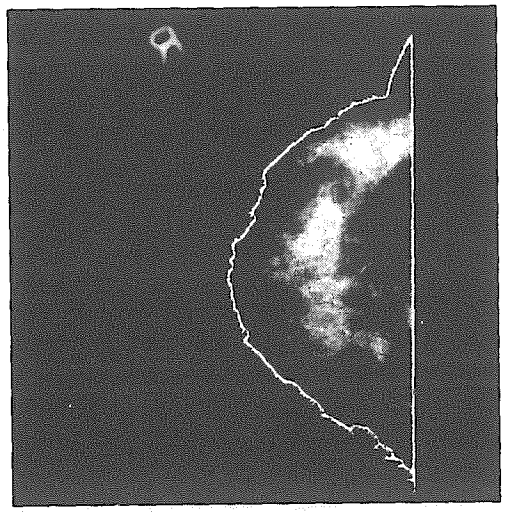

Fig. 3 The extracted breast region

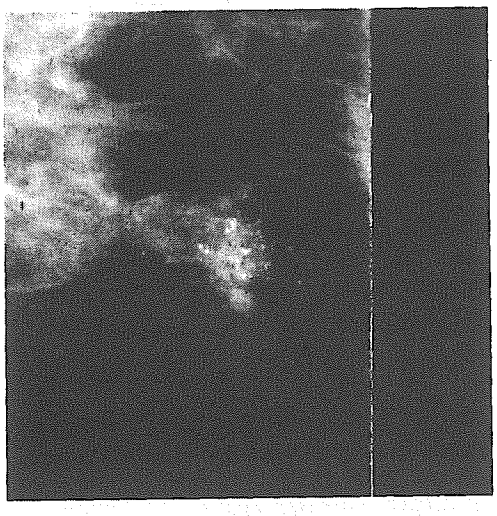

Fig. 4 the ROI

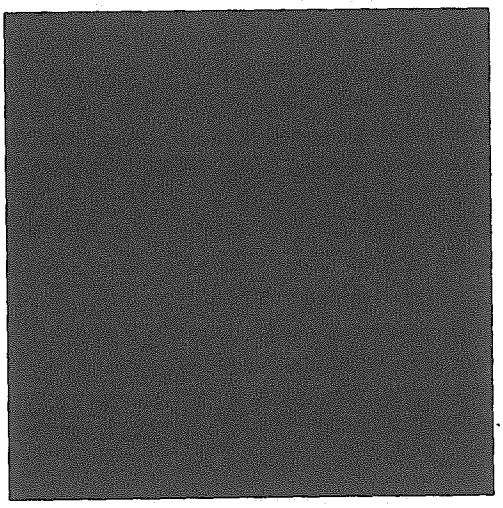

Fig.5 Enhanced image

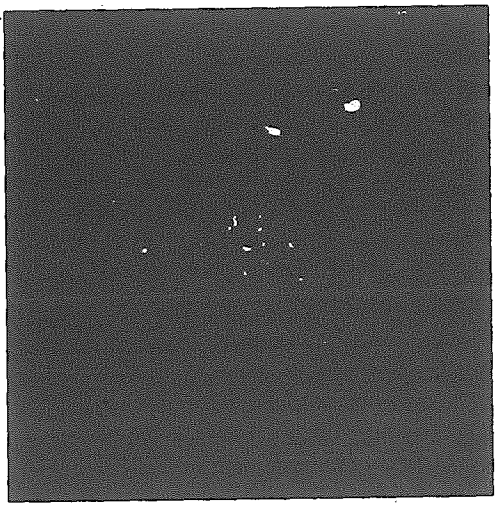

Fig. 6 Fractal-Sobel Filtered image

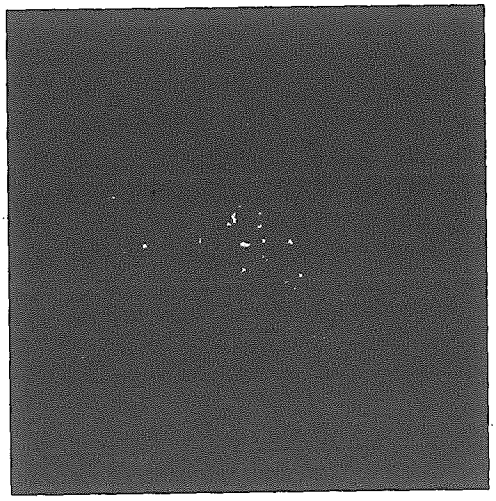

Fig. 7 The segmented result