FUZZY LOGIC AND SCALE SPACE APPROACH TO MICROCALCIFICATION DETECTION

H. D. Cheng and Jingli Wang

Department of Computer Science, Utah State University
Logan, UT 84322-4205

ABSTRACT

Breast cancer is one of the leading causes of women mortality in the world. Since the causes are unknown, breast cancer cannot be prevented. It is difficult for radiologists to provide both accurate and uniform evaluation over the enormous number of mammograms generated in widespread screening. Microcalcifications and masses are the earliest signs of breast carcinomas and their detection is one of the key issues for breast cancer control. Computer-aided detection of microcalcifications and masses is an important and challenging task in breast cancer control.

This paper presents a novel approach for detecting microcalcification clusters (MCCs). First, mammograms are normalized. Then, fuzzy set theory and fuzzy entropy principle are employed to fuzzify the mammograms. Then, the fuzzified images are enhanced. Finally, scale-space and Laplacian-of-a-Gaussian filter techniques are used to determine the sizes and locations of microcalicifications. A free-response operating characteristic (FROC) curve is used to evaluate the performance. The major advantage of the proposed system is its ability to detect microcalcifications even in very dense breast mammograms. A data set of 40 mammograms (Nijmegen database) containing 105 clusters of microcalcifications is studied. Experimental result show that the microcalcifications can be accurately and efficiently detected using the proposed approach.

Keywords: Fuzzy logic; Maximum entropy principle; S-function; Homogeneity; Microcalcification; Scale Space; Contrast; Laplacian-of-a-Gaussian (LoG).

1. INTRODUCTION

Breast cancer is the second-leading cause of cancer death in women, exceeded only by lung cancer. One of eight women could develop breast cancer at some point during their lifetime [1]. Primary prevention seems impossible since the causes of this disease are still unknown. Early detection is the key to improving breast cancer prognosis. Mammograms have been shown to be one of the most reliable methods for early detection of breast carcinomas. Although computer-aided mammography has been studied for two decades, automated interpretation of microcalcifications is still very difficult. The major reasons are: First, the objects of interest can be extremely small. They lead to potential misidentification. Second, different sizes, various shapes, and variable distributions of microcalcifications appear in mammograms; hence, temple matching seems to be impossible. Third, the regions of interest may be of low contrast. The intensity difference between suspicious areas and their surrounding tissues can be quite slim. Fourth, dense tissues and/or skin thickening, especially in younger women, can cause suspicious areas to be almost invisible. Finally, dense tissues may be easily misinterpreted as calcifications, causing a high false positive (FP) rate. This is a major disadvantage with the existing algorithms. To deal with these problems, a large number of techniques for breast cancer detection have been developed and described in the literature.

In this paper, we propose a novel approach to detect the microcalcification clusters in the mammograms of breasts with various densities. Our approach is based on fuzzy logic techniques. The proposed algorithm consists of the following steps: normalization, fuzzification, enhancement, and microcalcification detection by scale space signatures. We employ fuzzy entropy principle and fuzzy set theory to automatically determine fuzzy membership function. Contrast is defined based on homogeneity measurement and is used to enhance the images. A neural network determines the threshold and the scale space technique are used to decide the size and location of microcalcifications. All 40 mammograms in the Nijmegen database were utilized. The results aptly show that the microcalcification clusters can be accurately and effectively detected even in very dense mammograms.

2. ALGORITHM AND IMPLEMENTATION

The aim of this study is to develop an algorithm to detect microcalcification clusters in mammograms of the

breasts with various densities. The flowchart of the proposed approach is given in Fig. 1.

2.1 Image acquisition

The mammogram images were provided by the University Hospital of Nijmegen. They can be accessed from the website: figment.csee.usf.edu. The database, which has been widely used by researchers, contains 40 digitized mammogram images composed of both oblique and craniocaudal views from 21 patients. Each image has one or more clusters of microcalcifications marked by radiologists. The total number of clusters in the database is 105.

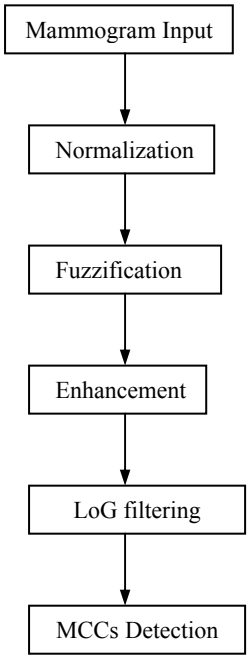

Fig. 1: The flowchart of the proposed algorithm.

2.2 Mammogram Normalization

The mammograms are with different brightness and contrast due to the varying illumination. In order to reduce the variation and achieve computational consistency, the images are normalized. We map all mammograms into a fixed range of the intensities from r_1 to r_2. Assume an image $g_i(x,y)$ whose maximum gray level is $\max g_i$ and minimum is $\min g_i$. We transform $g_i(x,y)$ into $g_k(x,y)$:

$$g_k(x,y) = r_1 + \frac{(g_i(x,y) - \min g_i) \times (r_2 - r_1)}{\max g_i - \min g_i} \qquad (1)$$

In our experiments we choose r_1 = 60 and r_2 = 210 because all the microcalcification intensities fell within this range based on the experimental results of a huge amount of mammograms.

2.3 Fuzzy Entropy and Membership Function

Fuzzy set theory is useful to deal with uncertainty. It is well known that mammograms have some degrees of fuzziness such as indistinct borders, ill-defined shapes, different densities, etc. Due to the nature of mammography and breast structure, fuzzy logic would be a better choice to handle the fuzziness of mammograms than traditional methods. We use the standard S-function [2-3] to fuzzify the images. The standard S-function is defined as:

$$\mu_{bright}(g) = S(g,a,b,c) = \begin{cases} 0 & g \le a \\ \dfrac{(g-a)^2}{(b-a)(ea)} & a \le g \le b \\ 1 - \dfrac{(g-c)^2}{(c-b)(ea)} & b \le g \le c \\ 1 & g \ge c \end{cases} \qquad (2)$$

where g is a variable representing the gray level, and a, b, and c are the parameters that determine the shape of the S-function.

2.4 Mammogram Enhancement

Since many mammograms are low contrast, blur and fuzzy. It is difficult to detect the microcalcifications. Mammogram enhancement is essential and important. We use fuzzy homogeneity to define the contrast and enhance the contrast. The method will use both the global and local information, therefore, it has much better performance [4].

2.5 Microcalcification Clusters Identification

Microcalcifications are small, subtle abnormalities that appear as isolated bright spots in mammograms. Because the diameters of the microcalcifications are between 0.1-0.3 mm, we can adapt the method in [5] to find the local maxim in Laplacian convoluted images when the size of filter kernel is chosen appropriately.
Two-dimensional Gaussian function:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}} \qquad (3)$$

Laplacian-of-a-Gaussian is:

$$LoG(x,y) = \frac{1}{\pi\sigma^4}\left(1 - \frac{x^2+y^2}{2\sigma^2}\right)e^{\frac{-(x^2+y^2)}{2\sigma^2}} \quad (4)$$

where σ is the standard deviation and σ >0.

To locate microcalcifications, we first compute LoG values. We choose the window size *h* to obtain an M× M kernel where M=3*h and* $\sigma = \frac{h}{2\sqrt{2}}$ [6].

Second, we calculate the convolution of the original image I with the LoG filter as follows:

$$[\text{LoG}(h) * \text{I}](x, y). \quad (5)$$

where *h*=1, 2,…, $h_{\max}$.

In our experiments, we choose $h_{\max}$ =12 because no appropriate spots as the candidates can be found when *h* value is equal to or greater than 12.

2.6 Detection of Microcalcification Clusters

Local maximums in the image filtered by LoG are considered as microcalcification candidates. We use a 3×3 window to identify local maximums. The pixel is regarded as a candidate when its value is maximum in a 3×3 neighborhood.

Coarse-to-fine tracking refines the localization of candidate microcalcifications. The images are mapped into scale spaces with different values of *h*. If the candidate response is larger than a predefined threshold, C>T(*h*), it is marked as candidate microcalcification. T(*h*) is a given threshold that is dependent on the size *h* of the spot. Here, we choose T(*h*) = $\varphi_I \times$ MaxT(*h*). φ_I is the threshold factor of image *I*. MaxT (*h*) denotes the maximum response.

In order to detect microcalcifications with high true positive (TP) rate, we must remove the isolated bright pixels. If three or more microcalcifications are within the region of 1 cm^2 , we consider that a cluster exists. In the given database, true clusters were detected and marked by radiologists. Regions with no microcalcifications that were detected by the machine will be counted as a false positive (FP) cluster. If the detection result is consistent with the one from expert radiologists, it will be counted as true positive (TP).

2.7 Thresholding

Threshold T is used to control the sensitivity of the detection. If T is small, it may cause too many FPs. On the other hand, if T is high, it may cause too many FNs (false negatives). The threshold value must be determined according to the characteristics of the mammograms.

We will use statistical values (standard deviation and mean) as the features to train a neural network whose output will be φ_I . We need to find the solutions φ_I of the function:

$$\varphi_I = f(\sigma_I, \mu_I) \quad (6)$$

In general, it is a nonlinear function.

For solving it, a multi-layer, feed-forward, error backpropagation neural network (BPN) is used. The parameters for the neural network used in our experiments are: Mean μ_I , and Standard deviation σ_I . Threshold factor φ_I is the output. Totally, 50 ROIs (region of interest) are employed. Training and testing sets were chosen randomly. The selection of optimal threshold value is trained by a set of 38 ROIs, consisting of 85 microcalcification clusters. The testing set includes 12 ROIs consisting of 20 clusters. The training data are not used during the testing stage.

The final structure of the neural network is:

- Architecture: 3 layer backpropagation neural network;
- Input of the neural network: mean and standard deviation;
- Output of the neural network: threshold value;
- Momentum: 0.5;
- Learning Rate: 0.5;
- Learn Rule: Delta-Rule;
- Activation Function: Sigmoid.

The training process is terminated when the number of epochs reaches 5000. A matrix of connection weights is obtained for the trained neural network. Finally, we test the network performance using the data in the testing set. We lost only one cluster for all 20 test clusters. Therefore, using the neural networks to determine the threshold factor is quite effective. The performance could be improved if we have a larger mammogram base.

3. RESULTS AND COMPARSIONS

In this study, 50 ROIs contain 105 clusters. The performance of the proposed approach is evaluated by a free-response receiver operating characteristic (FROC). Our FROC curve (Fig. 2) shows that the proposed method can archive accuracy greater than 97 true positive rates with the FP rate of three clusters per image.

We have compared our experimental results with those in [7] and [8] due to the facts:

1. The same set of images was used by these methods;
2. [7] used the scale-space approach as well;
3. [7] had compared its results with those in [8].

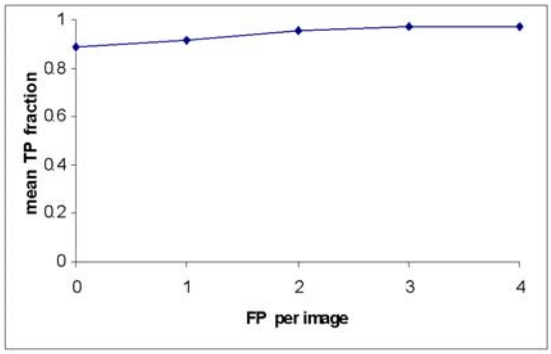

Fig. 2: FROC curve.
FROC is based on 105 TP and 50 ROI.

[7] employed an adaptive noise equatlizaion (ANE) in which noise characteristics is estimated from the mammograms. It obtained a maximum likelihood estimation of microcalcification patterns for all 40 mammograms in the Nijmegen database by using adaptive and fixed iso-precision scaling, and a logarithmic conversion. [8] employed Laplacian to detect the microcalcifications. It used statistical variation of the estimated contrast (SVEC). We compare the proposed approach (fuzzy logic and scale space algorithm, FLSS) with ANE and SVEC. The microcalcifications are detected at 1.0 false positive per image. The proposed approach misses only 10 clusters out of 105, compared to 12 clusters in the case of SVEC and 17 clusters in the case of ANE. Considering FPs, the proposed method obtains a much better result compared to ANE and SVEC. The number of FPs of the proposed method is 14 and those for ANE and SVEC are 39 and 42, respectively.

4. CONCLUSIONS

Breast cancer is one of the leading causes of death for women. Primary prevention seems impossible since the causes of this disease still remain unknown. Mammograms have been shown to be one of the most reliable methods for early detection of breast carcinomas.

In this paper, we use fuzzy set theory, fuzzy contrast enhancement and scale space to automatically detect microcalcification clusters in digitized mammograms. The proposed approach is very efficient for locating microcalcifications in the mammograms of breasts with various densities. Since microcalcifications are quite fuzzy and blur in mammograms, fuzzy set theory is preferable to ordinary methods for detecting microcalcifications clusters. The advantages of the proposed approach are as follows.

- The microcalcifications are accurately detected even in mammograms of very dense breasts;
- Mammogram enhancement is more adaptive and robust;
- Definition of the contrast based on fuzzy homogeneity uses both local and global information and the contrast enhancement algorithm can enhance the main features while suppressing noise;
- Some parameters can be altered to control different levels of true positive and false positive rates, and to generate FROC curve;
- The neural network uses the mean and standard deviation of the image intensities to determine the threshold factor.

ACKNOWLEDGMENT

Mammograms were provided by courtesy of the National Expert and Training Center for Breast Cancer Screening and the Department of Radiology at the University of Nijmegen, the Netherlands.

5. REFERENCES

[1] P. A. Wingo, T. Tong, and S. Bolden, "Cancer statistics," *A Cancer Journal for Clinicians*, Vol. 45, pp. 8-30, 1995.

[2] H. D. Cheng, J. R. Chen, and J. Li, "Threshold selection based on fuzzy c-partition entropy approach," *Pattern Recognition*, Vol. 31, No. 7, 857-862, July 1998.

[3] H. D. Cheng and J. R. Chen, "Automatically determine the membership function based on the maximum entropy principle," *Information Sciences*, Vol. 96, No. 3/4, pp. 163-182, Feb. 1997.

[4] R. C. Gonzalez and R. E. Woods, *Digital Image Processing,* 2nd Edi., Prentice-Hall, 2002.

[5] D. Blostein and N. Ahuja, "A multiscale region detector," *Computer Vision, Graphics, and Image Processing*, Vol. 45, pp. 22-41, 1989.

[6] F. Neycenssac, "Contrast enhancement using the Laplacian-of-a-Gaussian filter," *CVGIP: Graphical Models Image Processing,* Vol. 55, No. 6, pp. 447-463, 1993.

[7] N. Karssemeijer, "Adaptive noise equalization and recognition of microcalcification clusters in mammograms," *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 7, No. 6, pp. 1357-1376, 1993.

[8] T. Netsch and H. Peitgen, "Scale space signatures for the detection of clustered microcalcifications in digital mammograms," *IEEE Transactions on Medical Imaging,* Vol. 18, No 9, pp. 774-786, Sep. 1999.