A Method For Improve Preprocessing Images Mammography

Sara Dehghani and Mashallah Abbasi Dezfooli

Abstract—**A great volume of mammography images have dark background, these parts are not important in processing of mammography images. We can decrease the picture size and increase the processing speed by deleting these parts. Some parts of images have some notes and labels which consist of information like name, family, hospital's seal, date and…. It is better to omit these labels before processing since they are bright; the same as gray level of some breast tissues and have the same gray level. In the other way it is better to limit processing to the breast region and omit excessive parts. Previous studies over breast tissues show that if the considered region has not been marked clearly, it will influence characteristics and will not give the desired result. For obtaining this aim first find the main breast region, and then omit excessive parts to obtain processing result accurately and rapidly. The work that we do in this paper is introduced in three phases. The first phase is omitting the excessive image parts which are in the two sides of the image; we do this work by the usage of the pixels brightness. The second phase is the distinction of the breast direction and put all images in one direction; we do this work by the usage of threshold limit of gray level of the two halves of the image. The third phase is the breast region segmentation from the background; we do this work by the usage of series of point operations and the growing region method and the result has been reported to 99%.**

Key Words—**Breast Cancer, Mammography, Cropping, Breast Contour, Left Direction, Region Growing, Gray Threshold**

I. Introduction

Breast cancer is the most frequently diagnosed form of cancer in Canadian women, accounting for approximately 30% of all new cancer cases each year. In 2003, 21,200 cases were diagnosed and 5,300 women died of this disease [1]. In 1995, breast cancer accounted for 97,000 potential years of life lost for Canadian women. Since 1984 breast cancer incidence rates for all ages have increased progressively whilst mortality rates have remained relatively stable. One in nine Canadian women is expected to develop breast cancer in her lifetime, and one out of every 25 is expected to die from it. Mammography plays a central role in the process of detecting abnormalities in breast cancer screening. Although mammography is considered the most reliable means of detecting breast cancer, between IO-30% of women subsequently diagnosed with breast cancer have false negative mammograms [28]. It has been shown that an independent second reading of mammograms improves the sensitivity of mammography by as much as 15% [29]. Computerized analysis of mammograms has been envisaged as a means of providing a emulated secon- opinion [30], improving consistency by providing a standardized approach lo mammogram interpretation, and increasing detection sensitivity. Computer- Aided Detection (CADe) is the process of identifying potential abnormalities withm a mammogram, classifying regions of a mammogram as positive or negative. CADe systems use image processing algorithms to analyze mammograms for possible abnormalities including masses, microcalcifications, distortions in breast architecture, and symmetric densities. AAer an initial examination of the mammographic images, a radiologist reviews suspicious regions prompted by the CADe system and determines whether they warrant further follow-up. The CADe system assists the radiologist by confirming the detection of suspicious regions or identifying those that might otherwise have been missed. The precise segmentation of the breast region in mammograms is an essential preprocessing step in the computerized analysis of mammograms. It allows the search for abnormalities to be limited to the region of the breast without undue influence from the background of the mammogram. It also facilitates enhancements to techniques such as comparative analysis, which includes the automated comparison of corresponding mammograms. The breast boundary contains significant information relating to the deformation between two mammograms and is the source of information for relating the position of the nipple relative to the skin surface.

Manuscript received March 23, 2011
Sara Dehghani , Mashallah Abbasi Dezfooli ,Department of Computer engineering, Science and Research Branch, Islamic Azad University, Khouzestan-Iran (email: S_dehghani61@yahoo.com , m.abbasi@khouzestan.srbiau.ac.ir).

II. Works on Breast Region Segmentation

The breast gross-segmentation have been treated widely. Table I shows the tendencies and distribution of methods from the firsts works to recent approaches.

- *Histogram based techniques.* Probably one of the first attempts to separate the breast region was presented by [2] and it was done using simple histogram thresholding. The works of [3], as well as [5], and [7] used a simple thresholding to segment the breast from the background. The work of [6] presents a combination of local thresholding, region growing and morphological filtering. [8] propose their own global histogram thresholding, while [4] proposed a local thresholding method.

- *Gradient based techniques.* Breast region extraction techniques based on gradient have long been in use, since the early work of [9], who by means of spatial filters and a Sobel edge detector obtains the breast boundary. [10] use a two-level histogram threshold to obtain the breast region and oriented upwards, the region is then divided into three parts to track the boundary using the gradient. An evaluation of the quality of the segmentation is provided using the "accurate" or "near accurate" labels. They compare successfully their results to the work presented by [5]. The work presented by [14] takes advantage of a multiresolution scheme, processing in low-res and extrapolating the result. Using a global thresholding technique they obtain a preliminary region, which is processed using a 3x3 Sobel operator, and the pectoral muscle position is estimated via Hough transform. [12] provide an scheme based on different thresholds to find the breast edge. Using the gradient of two images and its union they obtain a possible breast contour. They found the boundary in 98% of the 500 images tested. The segmezntation presented by [13] was another gradient based method. After subtracting the background via an initial threshold, an edge was found by a line-by-line gradient analysis. [11] presented an improvement of this last approach.

Table I. Classification of breast gross-segmentation proposals.

Methods	1980's	1990's	2000's
Histogram	Hoyer79 [2]	Lau91 [3] Yin91 [5] Bick95 [6] Byng96 [7] Hein98 [8]	Masek00 [4]
Gradient	Semmlow80 [9]	Méndez96 [10] Abdel-Mottaleb96 [12] Morton96 [13] Karssemeijer97 [14]	Zhou01 [11]
Polynomial Modelling		Stomatakis94 [15] Chandrasekhar96 [16] Goodsitt98 [17]	
Active Contours		Ojala99 [18]	Ferrari00 [19] McLoughlin00 [20] Wirth04 [21]
Classifiers		Lou91 [22]	Saha01 [23] Rickard03 [24] Wirth04 [25] Tromans04 [26]

- *Polynomial Modelling based techniques.* An early method proposed by [15] was not a strict polynomial modelling. By means of an image preprocessing technique to enhance the response of non-dark pixels, a noise reduction process and a histogram threshold, they obtain the breast region, but the boundary is smoothed using Cubic B-splines and samples at fixed pixel intervals are extracted. Then a smooth curve is generated through cubic polynomial calculations. One of the firsts, effective and real polynomial modelling was presented by [16]. An initial threshold is used to approximate the breast region. Their method provides around 94% acceptable results from 300 images from MIAS [27] mammogram database. A quadratic/cubic polynomial fitting method was proposed by [17] which is fitted by translation and rotation the axes.

- *Active Contours based techniques.* One of the firsts applications of the active contours on breast segmentation was presented by [18], [20]. They apply a global threshold to obtain an initial result. They statistically model the breast with the pixels inside the region and a snake algorithm is applied to obtain the final boundary. On the other hand, [19] propose a method that firstly enhances the image with a logarithmic transformation, and then an iterative technique (as the Lloyd-Max least-squares) is applied to find and optimal threshold. Finally, they use a B-Spline to approximate the boundary. Recently, [21] propose an active contour to segment the breast. The method obtains two preliminary regions using a convolution matrix to enhance the edges and a dual threshold obtained by different techniques. They obtain the control point for the snake with the comparison of the two regions. They evaluate the method over the MIAS database.

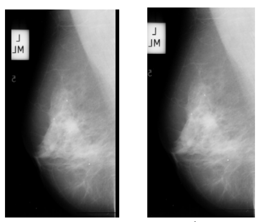
- *Classifiers based techniques.* [22], used a clustering approach to obtain an initial region, estimates the real boundary extrapolating and linking those detected points. [23] use a scale-based fuzzy connectedness algorithm. [24] presents Self-organizing map, a type of unsupervised artificial neural network model. The method applied by [25] was a fuzzy segmentation and evaluates the results in terms of completeness and correctness comparing the images from the MIAS database with a gold standard manually generated. Recently, [26], use a mixture model to obtain a mathematical representation of the image back-ground and the compressed parameters, combined with a Fourier model, using an Expectation Maximization algorithm. Summarizing, the traditional histogram based method has provided good and quick results. This quality sometimes turns on weakness in difficult cases where can be enhanced with local histogram or gradient approaches. The polynomial modelling and active contours provide very good results with accurate profiles.

III. Our Method

To achieve the segmentation we propose a "three-phase" based method.

Stage A: omit the excessive parts of the image

Excessive background parts which do not cover breast region, are the right and left part of the image. Where as the backgrounds are dark and their gray level are near zero and these gray level do not have any difference with each other, we can use the amount of bright points for recognizing these parts. fig1(a) show this reality. For omitting extra image parts in this project we sweep the whole image from right to left and left to right and we go forward until the time the vertical brightness amount of the image is more than the specific threshold limit or the difference of vertical brightness amount of pixels is smaller and fewer than the specific threshold limit, or more than the specific threshold limit. If the difference of vertical brightness amount of pixels is less than the smaller threshold limit, and more than the bigger threshold limit or the vertical brightness amount of the image is more than the specific threshold limit, it means we are still in the background restriction. So we calculate the difference of vertical brightness amount of the next pixels. We repeat this work until the time the amount of grey level changes is more than the smaller threshold limit or less than the bigger threshold limit. The amount of threshold limit is obtained experimentally. These parts are excessive and should be omitted of the images. You can see an example in figure 1.

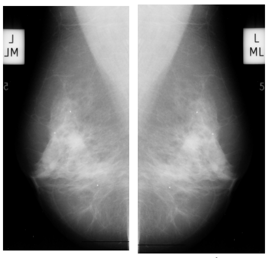


a b

Fig 1. a) original image size 1024x1024 b) excessive background has been omitted and the image size has been decreased to 510x1024

Stage B: diagnose the image direction and put it to one direction

MIAS's database images which are used in this project are belonged to both right and left breast. In the other way breast image in some image is put in the right side and in the other images in the left side. To make processing easier we put whole the breast image to the left side of the image. To do this work, first we should find the breast direction. First we divide the image to two halves in comparison with vertical axis then we calculate the grey level of threshold limit of each two image halves. Surely the threshold limit under the image which the breast is in it, is more than the under of the other image which is the background. After finding the breast direction we put it to left direction. This work is done by 180 degree direction with regard to vertical axis. In figure 2 there is an example of changing direction.

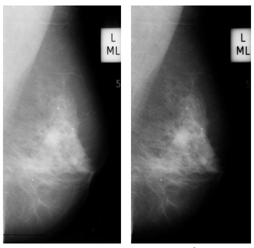


a b

Fig 2. a) the original breast image is in the right side b) the image is put in the left side for easier processing

Stage C: omit labels and breast region from background.

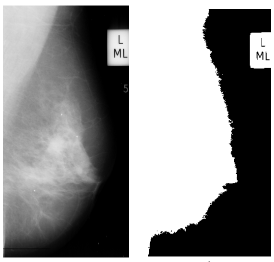
For finding the tissue boundary of the breast, we act as follow; first we change the image from unsigned region of unit 8 to double decimal, then we obtain the image energy which is equal to the second power of decimal image. A sample of image energy is in figure 3.



a b

Fig 3. a) a sample of left sided image b) a sample of image energy

Then according to the grey level of threshold limit we make the original image binary. There is a sample of binary image in figure 4.



a b

Fig 4. a) a sample of left sided image b) a sample of binary image

With attention to put image in one direction in the previous stage and in the spite of this point that we know all images are in the left side, then we go forward by the region's growing method and measuring images from left to right and up to down until the time the image energy is more than the specific threshold limit or the amount of pixels's brightness of binary picture is equal to 1. Then with high attention the breast region is separated from the background and the other parts (the grey black surface) become zero. With this action in addition to breast tissue segmentation, notes have been omitted simultaneously. Figure 5 shows the breast region. The same as you can see the labels in the upper part of the image which is considered as a noise is omitted in this stage.

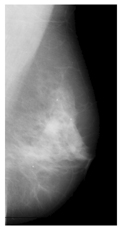


Fig 5. breast region segmentation and omitting probable labels

IV. Experimental Results

We have used the public database MiniMIAS [27] to test our method. It is a reduced version of the original MIAS Database (digitized at 50 micron pixel edge) that has been reduced to 200 micron pixel edge and clipped or padded so that every image is 1024x1024 pixels.

Figure 6 shows three representative results. We have tested over 60 images, and we have obtained a 99% of "near accurate" results, which include the "accurate" results.

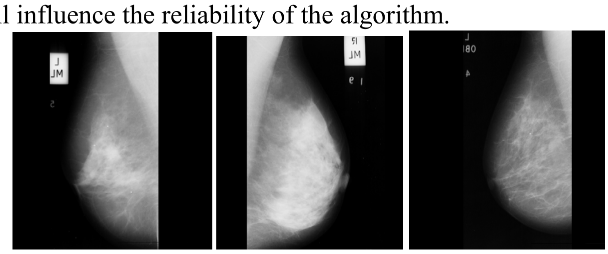
To summarize, the results obtained by the method show that it is a robust approach but it can be improved in terms of accuracy. Even so, we accept this method because it provides useful regions (there is no meaningful loss of information).

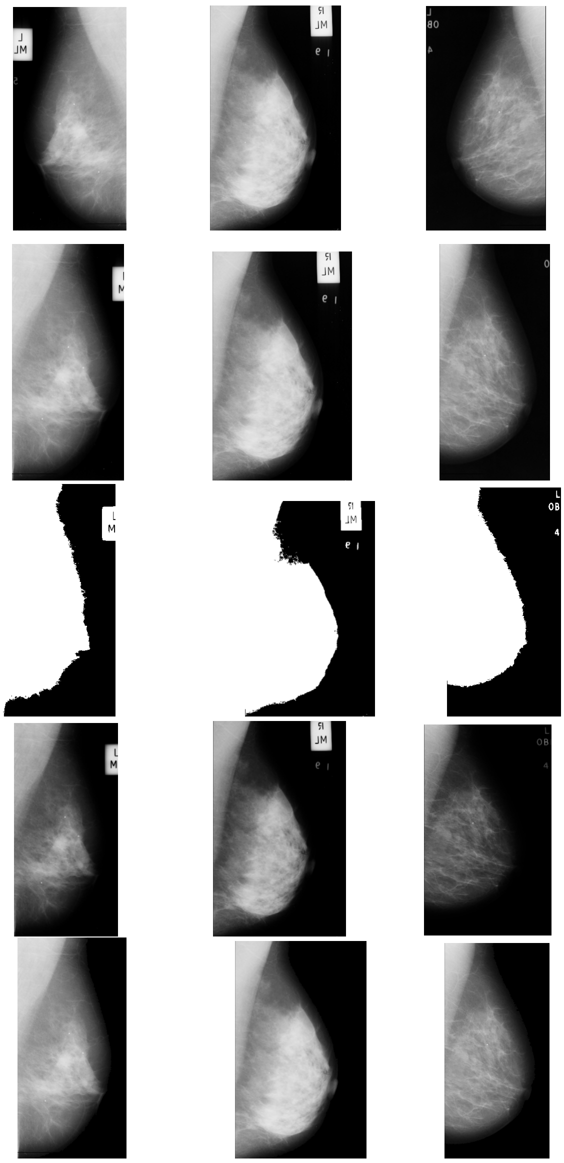
V. Conclusions

The literature survey will be a useful resource for others researching in this area.

A new method to segment the breast has been presented. The results obtained over MiniMIAS database have shown a general good behavior.

The results have shown that problems with the image acquisition, background noise, artifacts and scratches could all influence the reliability of the algorithm.





(a) (b) (c)

Fig 6. An example of the performance of the presented approach on the segmentation of the profile of three different breasts.

V. REFERENCE

[1] National Cancer Institute of Canada: Canadian Cancer Statistics. Toronto, Canada, 2003.
[2] Hoyer, A., Spiesberg, W.: Computerized mammogram processing. In: Phillips Technical Review. Volume 38. (1979) 347-355
[3] Lau, T., Bischoff, W.: Automated detection of breast tumors using the asymmetry approach. In: Computers and Biomedical Research. Volume 24. (1991) 273-295.
[4] Masek, M., Attikiouzel, Y.: Skin-air interface extraction from mammograms using an automatic local thresholding algorithm. In: ICB, Brno, CR (2000) 204-206.
[5] Yin, F., Giger, M.: Computerized detection of masses in digital mammogram: analysis of bilateral subtraction images. In: Medical Physics. Volume 28. (1991) 955-963.
[6] Bick, U., Giger, M.: Automated segmentation of digitized mammograms. In: Academic Radiology. Volume 2. (1995) 1-9.
[7] Byng, J., Boyd, N.: Automated analysis of mammographic densities. In: Medical Physics. Volume 41. (1996) 909-923.
[8] Hein, J., Kallargi, M.: Multiresolution wavelet approach for separating the breast region from the background in high resolution digital mammography. In: Digital Mammography, Nijmegen, Kluwer Academic Publishers (1998) 295-298.
[9] Semmlow, J., Shadagopappan, A.: A fully automated system for screening xero-mammograms. In: Computers and Biomedical Reseach. Volume 13. (1980) 350-362.
[10] Mendez, A., Tahoces, P.: Automatic detection of breast border and nipple in digital mammograms. In: Computer Methods and Programs in Biomedicine. Volume 49. (1996) 253-262.
[11] Zhou, C., Chan, H.: Computerized image analysis: Estimation of breast density on mammograms. In: Med. Phys. Volume 28. (2001) 1056-1069.
[12] Abdel-Mottaleb, M., Carman, C.: Locating the boundary between the breast skin edge and the background in digitized mammograms. In: Digital Mammography. (1996) 467-470.
[13] Morton, A., Chan, H., Goodsitt, M.: Automated model-guided breast segmentation algorithm. In: Med Phys. (1996) 1107-1108.
[14] Karssemeijer, N., te Brake, G.: Combining single view features and asymmetry for detection of mass lesions. In: IWDM. (1998) 95-102.
[15] Stomatakis, E., Cairns, A.: A novel approach to aligning mammograms. In: Digital Mammography. (1994) 255-364.
[16] Chandrasekhar, R., Attikiouzel, Y.: Gross segmentation of mammograms using a polynomial model. In: International Conference of the IEEE in Medicine and Biology Society. Volume 3. (1996) 1056-1058.
[17] Goodsitt, M., Chan, H.: Classification of compressed breast shapes for the design of equalisation filters in x-ray mammography. In: Medical Physics. Volume 25. (1998) 937-947.
[18] Ojala, T., Liang, J.: Interactive segmentation of the breast region from digitized mammograms with united snakes. Technical Report 315, Turku Centre for Computer Science (1999).
[19] Ferrari, R., Rangayyan, R.: Segmentation of mammograms: Identification of the skin boundary and the pectoral muscle. In: IWDM. Volume 23. (2000).
[20] McLoughlin, K., Bones, P.: Locating the breast-air boundary for a digital mammogram image. In: Image and Vision Computing. (2000).
[21] Wirth, M., Stapinski, A.: Segmentation of the breast region in mammograms using snakes. In: Canadian Conference on Computer and Robot Vision. (2004) 385-392.
[22] Gauch, J.: Image segmentation and analysis via multiscale watershed hierarchies. In: IEEE Transactions on Image Processing. Volume 8. (1999) 69-79.
[23] Saha, P., Udupa, J.: Breast tissue density quantification via digitized mammograms. In: IEEE Transactions on Medical Imaging. Volume 20. (2001) 792-803.
[24] Rickard, H., Tourassi, G., Elmaghraby, A.: Self-organizing maps for masking mammography images. In: IEEE EMBS. (2003) 302-305.
[25] Wirth, M., Lyon, J., Nikitenko, D.: A fuzzy approach to segmenting the breast region in mammograms. In: IEEE FI. Volume 1. (2004) 474-479.
[26] Tromans, C., Brady, J., Warren, R.: A high accuracy technique for breast air boundary segmentation and the resulting improvement from its use in breast density estimation. In: IWDM. (2004) 17-18.
[27] Suckling, J., Parker, J.: The mammographic images analysis society digital mammogram database. In: Exerpta Medica. International Congress Series. Volume 1069. (1994) 375-378.
[28] RE, Bird, T.W. Wallace, and B.C. Yankaskas, "Analysis of Cancers Missed at Screening Mammography", Rodoiogy, vol.184 , pp.613-617, 1992.
[29] E.L. Thur@Il. RA. Lemevall, and A.A.S. Taubc, '.Benefit of Independent Double Reading in a Population-based Mammography Screening Program", Radiology, vol.191,p p.241-244, 1994.
[30] R.L. Birdwell, D.M. Ikeda, K.F. O'ShaugImessy, and E.A. Sickles, "Mammographic Characteristics of I I5 Missed Cancers later Detected with Screening Mammography and the Potential Utility of Computer-Aided Detection", Rodiology, vol .219, pp.192-202,2001.

Sara Dehghani took his Ms.c in computer Engineering, Software from Science and Research Branch, Islamic Azad university, Khouzestan-Iran. His area of research interest is Image processing.