 ISSN : 2229-4333(Print) | ISSN : 0976-8491(Online)

Segmentation of the Mammogram Images to find Breast Boundaries

[1]Sumit Chopra, [2]Pankaj Bhambri, [3]Baljit Singh
[1]K.C. College of Engineering & IT, Nawanshahr, Punjab, India
[2]Guru Nanak Dev Engg. College, Ludhiana, Punjab, India
[3]BBSBEC, Fatehgarh Sahib, Punjab, India

Abstract
The principal feature in Mammographic Images is to find the region of interest. Extraction of breast region is an important pre processing step in the detection of the area which is required to be processed for further processing. Detection of edges is an image is very important step towards understanding image features. Since edges often occur at image location representing object boundaries , edge detection is used in image segmentation when images are divided into areas corresponding to different objects.The edge detectors present don't show good results in mammographic images. In this paper a new algorithm is developed to find boundary of the breast which decreases the amount of further processing required

Keywords
Boundary detection,Mammograms, Region of Interest, Segmentation.

I. Introduction
Breast cancer is the most frequently diagnoised form of cancer in Canadian women, accounting for approximately 30 % new cases every year. One in nine Canadian women is expected to develop breast cancer in her lifetime and out of every 25 one 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 10-30 % of women diagnosed with breast cancer have false negative mammograms [1]. A significant problem with Mammography is the variability of diagnosis among radiologist and Computer Aided Diagnosis (CAD) methods (that consistently highlight region on mammograms that might warrant futhrer examination) are being developed to assist the radiologists.

One of the first step in CAD is the segmentation of the image into breast and the background region. This has the advantage of simplifying further processing of the image (by eliminating the background) and also provides the reference for the alignment of views when the two views are being compared. Knowledge of breaspt edges also sometimes help with the identification of large masses that have distorted the outline of the breast. Medical Image edge detection is an important work for object recognition of human organs such as lungs and ribs, and it is an essential pre processing step in medical image segmentation [2,3]. For mammograms manifesting masses this corresponds to the detection of suspicious mass regions. A number of image processing methods have been proposed to perform this task. Weight median filtering has been proposed to enhance digital image prior to object identification. Thresholding and fuzzy pyramid linking has been used for detecting mass localization and calcification [4]. Edge detection in image processing is the task of locating pixel value variations in images. The success of edge detection provides a good basis for the performance of higher level image processing task such as object recognition, target tracking and segmentation, since it reduces the amount of information to be processed.

II. Existing Methods of Breast Border Detection
Many methods are available for detecting breast border in mammograms which includes Grey level Co -occurrence Matrix, Watershed Algorithm, Extended Edge Operator, Thresholding, Tracking, Artifical Neural Networks and Modelling. These methods are discussed below

A. Grey Level Co-occurrence Matrix
Grey level Co – occurrence Matrix method is used as a texture descriptor in the process of feature extraction.The selection of certain texture is possible as it is based on the distribution in grey level co- occurrence Matrix (GLCM). Boundaries that separate between textures can be created by separating the gradients in one dimensional GLCP statistical features. The process of GLCP extraction is arbitrary and takes unreliable time. The value of the co-occurrence matrix elements presents relative frequencies with two neighbours pixels separated by distance d appear on the image ,where one has gray level value i and other j .Such matrix is symmetric and also a function of angular relationship between two neighboring pixels. By using gray level co-occurrence matrix we can extract the different features like probability,energy, entropy, variance, inverse moment difference etc. Some of them are defined as follows.

$$\text{Maximum Pr obability} = \max(P_{ij}) \qquad (1)$$

$$\text{Variance} = (\sum (i-\mu_i)^2 \sum Pij(\sum ((j-\mu j)^2 \sum P_{ij}) \qquad (2)$$

$$\text{Correlation} : \sum_i \sum_j (i-\mu_x)(j-\mu_y)P_{ij} / \sigma_x\sigma_y \qquad (3)$$

$$\text{Entropy} = \sum_i \sum_j P_{ij}\log(P_{ij}) \qquad (4)$$

Where μ_x and μ_y are the means and σ_x, σ_y are the standard deviations. Amongst all these features entropy gives good results.

B. Thesholding
For mammograms, thresholding usually involves selecting a single gray level value from an analysis of the grey-level histogram, to segment the histogram into background and breast tissues. All the pixels with grey level value less than the threshold are marked as background and the rest as breast. Thresholding uses only grey level value and no spatial information is considered. Therefore, the major shortcoming of the threshold is that there is often an overlap between grey levels of the objects in the breast and the background.

Micro calcification clusters are the primary indicator of malignant type of breast cancer, the detection is important to prevent and treat the disease. The micro calcification appears in small clusters of a few pixels with relatively high intensity and closed contours compared with there neighboring pixels. However, it is a challenge to detect all the micro calcifications since they appear as spots which are slightly brighter than their background. The micro calcifications in mammograms can be detected by using dual threshold method [5]. Experimental results showed that proposed method can locate the micro calcifications exactly in the mammograms as well as restrain the contours produced by the noises.

C. Tracking

Tracking the breast border means involves implementing a tracking algorithm that marks a pixel as a border pixel if it satisfies certain condition.

Yen. et.al [6] used a four connectivity tracking algorithm to find the breast border. In this computerized scheme is developed for the detection of masses in digital mammograms. Based on the deviation from the normal architecture symmetry of left and right breast , a bilateral subtraction technique is used to enhance the conspicuity of possible masses. The right and left breast images in each pair are aligned manually in each pair. A non linear bilateral subtraction technique that involves linking multiple subtraction images has been investigated and compared to simple linear subtraction model. Various feature extraction techniques are used to reduce false positive detection resulting from bilateral subtraction.

In [7] a semi automatic method of detecting the breast border, utilizing the gradient of gray levels in three user selected regions. The algorithm used thresholding and pair wise pixel differences in specific directions to detect the breast border. Two radiologist and one physic evaluated the result .

D. Watershed Algorithm

Watershed is the ridge that divides areas drained by different river systems. A catchement basin is the geographical area draining into a river or reservation .The watershed transform applies these ideas to gray-scale image processing in a way that can be used to solve a variety of image segmentation problems.

Watershed classifies pixels into regions using gradient descent on image features and analysis of weak points along region boundaries. The image feature space is treated , using a suitable mapping, as a topological surface where higher value indicates the presence of boundaries in the original image data.

E. Artifical Neural Networks

In [8] Multiple linked self organizing neural networks to segment the breast into four components: background, pectoral muscles, fibro glandular tissue and adipose tissue. This method had the advantage of simultaneously identifying the background and the pectoral muscles , but no evaluation of the background segmentation results are given.

A neural network is a graph with patterns represented in terms of numerical values attached to the nodes of the graph and transformation between patterns achieved via simple message passing algorithm. Certain of the nodes in the graph are generally distinguished as being input nodes or output nodes, and the graph as a whole can be viewed as a representation of multivibrate function linking inputs to outputs. Numerical values (weights) are attached to the links of the graph, parameterizing the input and output function allowing it to be adjusted to via a learning algorithm

A broader view of a neural network architecture involves treating the network as a statistical processor, characterized by making particular probabilistic assumptions about the data. Patterns appearing on the input node or output nodes of the network are viewed as samples from probability densities, and a network is viewed as a probabilistic model that assign probabilities to patterns. The problem of learning the weights of the pattern is therby reduced to the problem in statistics – that of finding weight value that look probable in the light of observable data.

F. Extended Edge Operators

Prewitt, Sobel, Kirch and Compass operators are used for finding edges in images. For detecting tumors in mammographic images extended Edge operators for Prewitt, Sobel, Kirsh and Compass operator are proposed.These operators are proposed in the preeceeding section.

1. Extended Prewitt Edge Detector

Sobel operator is applied by convolving the image with the Sobel operator to find the derivative in both horizontal and vertical direction. By convolving with the horizontal direction Sobel kernel we find the gradient in horizontal direction and by convolving in the vertical direction we find the gradient in vertical direction. Typically it is used to find the absolute approximate magnitude at each point in an input gray scale image. The operator consists of 3 X 3 convolution kernel as shown below.

$$\begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & -2 & 1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

Fig. 1 (a) Fig.1 (b)

This implies result of Sobel operator at an image point which is a region of constant image intensity is a zero vector and at a point on an edge is a vector which points across the edge ,from darker to brighter value.

To achieve better results for the segmentation of Mammograms the 3 X 3 kernels are modified to 5 X 5 kernels. One kernel is simply the other modified by 900. These extensions to Sobel kernel are shown below

$$\begin{pmatrix} 2 & 2 & 4 & 2 & 2 \\ 1 & 1 & 2 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & -2 & -1 & -1 \\ -2 & -2 & -4 & -2 & -2 \end{pmatrix} \qquad \begin{pmatrix} 2 & 1 & 0 & -1 & -2 \\ 2 & 1 & 0 & -1 & -2 \\ 4 & 2 & 0 & -2 & -4 \\ 2 & 1 & 0 & -1 & -2 \\ 2 & 1 & 0 & -1 & -2 \end{pmatrix}$$

Gx Gy

Fig. 2(a) Fig. 2(b)

2. Extended Prewitt Operator

Prewitt is a method of edge detection in image processing which calculates the maximum response of a set of convolution kernels to find local edge orientation for each pixel. The Prewitt edge detector is an appropriate way to estimate the magnitude and orientation of an edge.In all the edge detection methods following algorithm is used.

(a)Smooth the input image

$$(\hat{f}(x, y) = f(x, y) * G(x, y) \tag{5}$$

 ISSN : 2229-4333(Print) | ISSN : 0976-8491(Online)

$$(\hat{f}_x(x,y) = \hat{f}(x,y) * M_x(x,y) \quad (6)$$

$$(\hat{f}_y(x,y) = \hat{f}(x,y) * M_y(x,y) \quad (7)$$

$$magn(x,y) = |\hat{f}_x| + |\hat{f}_y| \quad (8)$$

$$dir(x,y) = \tan^{-1}(\hat{f}_y \,|\, \hat{f}_x) \quad (9)$$

If magn(x,y) > T, then possible edge point.

The 3 X 3 kernel over segmented the mammographic images. So to reduce the over segmentation they are modified to 5 X 5 kernels. The extension of Prewitt edge detector for 00 and 900 are shown below.

$$\begin{pmatrix} 2 & 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ -1 & -1 & -1 & -1 & -1 \\ -2 & -2 & -2 & -2 & -2 \end{pmatrix} \qquad \begin{pmatrix} 2 & 1 & 0 & -1 & -2 \\ 2 & 1 & 0 & -1 & -2 \\ 2 & 1 & 0 & -1 & -2 \\ 2 & 1 & 0 & -1 & -2 \\ 2 & 1 & 0 & -1 & -2 \end{pmatrix}$$

Gx Gy

Fig. 3(a) Fig. 3(b)

3. Kirsch Edge Detector

The kirsch operator is a non linear edge detector that finds the maximum edge strength in a few predetermined directions. The kirsch edge detector detects edges using eight filters are applied to the image with maximum being retained for the final image. The eight filters are a rotation of basic compass convolution filters. The 3 X 3 kernels over segmented the image, so to reduce their artifacts they are extended to 5 X 5 kernels and their extensions are shown below.

$$\begin{pmatrix} 9 & 9 & 9 & 9 & 9 \\ 9 & 5 & 5 & 5 & 9 \\ -7 & -3 & 0 & -3 & -7 \\ -7 & -3 & -3 & -3 & -7 \\ -7 & -7 & -7 & -7 & -7 \end{pmatrix} \qquad \begin{pmatrix} 9 & 9 & -7 & -7 & -7 \\ 9 & 5 & -3 & -3 & -7 \\ 9 & 5 & 0 & -3 & -7 \\ 9 & 5 & -3 & -3 & -7 \\ 9 & 9 & -7 & -7 & -7 \end{pmatrix}$$

Gx Gy

Fig. 4(a) Fig. 4(b)

All the above discussed techniques suffer from some drawback and the new algorithm developed has tried to overcome the results and results are compared with the existing techniques.

III. Algorithm for Finding boundary of breast in Mammographic Images

The process of finding the boundary of the breast in mammogram images is divided into two steps. Each mammogram image contains a text label which needs to be removed prior to finding boundary of the image.The algorithm for finding the boundary of the image is given below:-

1. Divide a mammogram into pq blocks of size M X N, {B11, B12,…., Bpq
2. Use the block region growing method to eliminate the background.

(i) For each block Bij, compute its mean , variance and energy function given by the equation

$$E_{ij} = (\mu_{ij})^2 + (\sigma_{ij})^2 \quad (10)$$

(ii) Begin with the block with the smallest energy, then use the four – neighbor connectivity rule to grow a region with a prescribed tolerance TE.

(iii) The region obtained by step 2.2 will be considered to be the breast background and will be eliminated.

3. Apply k- mean clustering based thresholding method to extract breast region.

After removing the text label, next step is to find the edges in mammogram images. The algorithm for finding edges in mammogram images is given below:-

(i) Accept the input image f(x,y) of size m x n.

(ii) Convert the given image into type double

(iii) Convert the image into binary image by using suitable threshold value

(a) Select an initial estimate for T.

(b) Segement the image using T.

(c) This will produce two groups of pixel: 1,consisting of intensity value > =T, and G_2 consisting of pixels with values <T.

(d) Compute the median values μ_1 and μ_2 for pixels in the region G_1 and G_2.

(e) Repeat the steps 2 to 4 untill the difference in T in successive iterations is smaller than a predefined parameter.

(iv) Let [x,y] = size of the binary image.

(v) Initialize b, c to 1 and x_1,y_1 to 0..

(vi) Repeat the steps 7 to 8 for i = 1 to x and j = 1 to y

(vii) If any pixel in an given image is 1 and the pixels to the right and left of that pixel is 0 then x_1(c) = i and y_1(c) = j and break.

(viii) Plot the points (y_1, x_1,'w')

(ix) After the edge detection convert the image into gray level image using Otsu's threshold method [9].

IV. Results and Discussion

This is a new method of finding boundaries of mammogram images. The traditional techniques of finding edges in an image don't work well in case of mammogram images. In this the given mammogram image the result are compared with the existing edge detection techniques such as Sobel, Prewitt, Laplacian of Gaussian and Canny edge detector. They are evaluated that how the existing technique is better than the already present. In the given technique selecting an appropriate threshold value is the key. If the proper threshold value is not selected then it will lead to missing of the pixels which are true if the threshold value selected is high and if the threshold value is low then it will leads to the resultant image containing the false pixels.

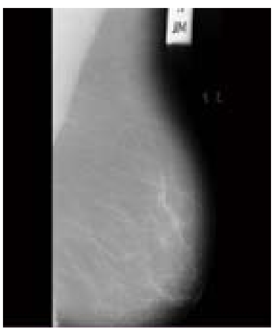

Fig. 5(a)

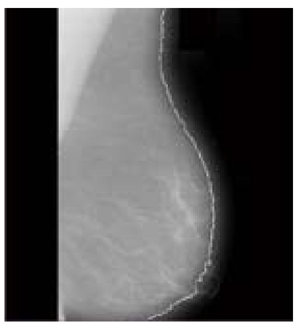

Fig. 5(b)

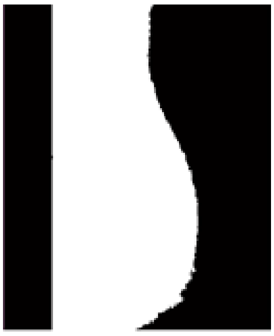

Fig. 5(c)

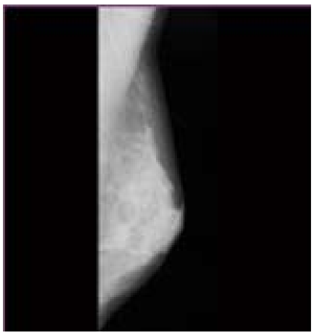

Fig. 6(a)

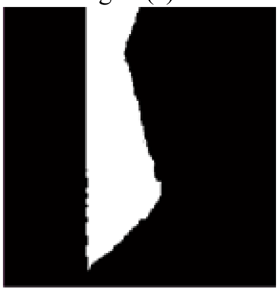

Fig. 6(b)

Fig. 6(c)

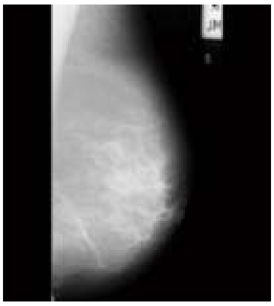

Fig. 7(a)

Fig. 7(b)

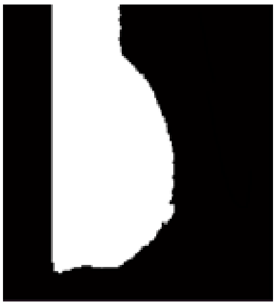

Fig. 7(c)

Fig. 8(a)

Fig 8(b)

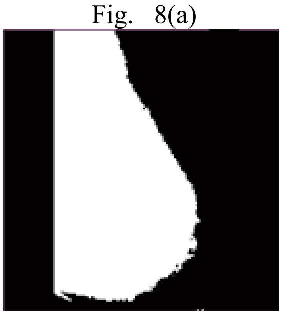

Fig 8(c)

Fig.s in the first column shows the original input mammographic images , Second column shows the output of applying the algorithm for finding the edges in mammographic images and the third column shows the output of applying the thresholding on the mammographic images.

References

[1] E. Bird, T.W. Wallace , B.C. Yankaskas. "Analysis of Cancers missed at screening Mammography Radiology", 184: pp. 613 – 617,1992.

[2] M.L. Rajab, M.S. Woolfson, S. P. Morgan,"Application of region based segmentation and neural network edge detection to skin lesions", Computerized Medical Imaging and Graphics ,vol. 28, pp. 61-68, 2004.

[3] W. Qian, L.P.Clarke, M. Kallegri, R. A. Clark.,"Tree structured non linear filters in digital Mammography", IEEE Trans. Med. Imag.,vol. 8,no.4, pp. 377- 386, Dec 1989.

[4] D. Brazokovic, X.M. Luo, P. IBzrakovic, "An approach to automated detection of tumors in Mammography", IEEE Trans. Med. Imag., Vol 9,No. 3, pp. 233-241, Sept 1990.

[5] Y. Wu,Q. Huang, Y. Peng, W. Situ, "Detection of micro calcification in digital mammograms using dual threshold", pp. 347 – 354 , 2006.

[6] Yin F, Giger M. Doi K, Metz C, Vyborny C, Schmidit R. "Computerized detection of masses in digital mammograms : analysis of bilateral subtraction images , Med Physics 1991; 18, pp. 955-63

[7] Bick U, Gigger M. , Schmidit R.Nishikawa R. ,Wolverton D., Doi K. "Automated segmentation of digitized Mammograms. Acid Radiol 1995 ;2, pp. 1-9.

[8] Suckling J., Dance D.,Moskovic E. , Lewis D., Blacker S.,"Segmentation of mammograms using multiple linked self self organizing neural networks". Med Phys. 1995 : 22; pp. 145 -52.

[9] Otsu N. , "A threshold Selection Method from gray level histogram", IEEE Transaction on System, Man and Cybermetics, Vol. 9 No.1, 1979, pp. 62-66.

Sumit Chopra received his B.Tech Degree from Punjab Techniacal University, Jalandhar and M.tech Degree again from PTU, Jalandhar. He is now working as a lecturer in K.C.College of Engineering & IT, Nawanshahr. His research areas include Digital Image Processing, Software Engineering and Database systems.