A Review on Preprocessing Techniques for Digital Mammography images

Aziz Makandar
Professor,
Department of Computer Science,
Karnataka State Women's University, Vijayapur

Bhagirathi Halalli
Research Scholar,
Department of Computer Science,
Karnataka State Women's University, Vijayapur

ABSTRACT
Mammograms are the soft X-rays kind of imaging technique used for the detection of any lesions or cysts in breasts. Digital mammograms have many kinds of artifacts that affect the accuracy of the detection of tumor tissues in the automated Computer Aided Detection (CAD) system for mammograms. Preprocessing helps to remove such artifacts is an important step. Image preprocessing is used to maintain image efficiency in mammogram images there are many artifacts need to be removed like labels, patient name, muscle part, etc. and enhance the region of interest which helps for efficient segmentation and detection of tumor. The basic objective of this study is to evaluate and discuss different techniques and approaches proposed in order to enhance the breast cancer images and an efficient preprocessing technique for mammography. It aims to find the existing preprocessing techniques for mammography images and discuss the techniques used and their advantages.

General Terms
Digital Image Processing, Preprocessing of Mammography and Image Enhancement.

Keywords
Breast Cancer, Preprocessing, Active Contour, Seeded Region Grow, Contrast Enhancement, Morphology, Watershed, and Region of Interest (ROI).

1. INTRODUCTION
Mammography is a low energy x-ray procedure for the visual image of the internal structure of the breast. The diagnostic technique has been verified to be the foremost reliable technique and it is the essential screening technique for detection of breast cancer. Image pre-processing techniques are necessary for seeking out the orientation of the mammogram, to diagnose and to reinforce the standard of the region of interest [1-9]. Before applying any image processing algorithm the preprocessing is essential to get ready with the enhanced mammogram. Preprocessing is vital to limit the hunt for abnormalities and will not influence from background of the mammogram [19-30].

Basically medical images are difficult to be understood; therefore a preprocessing phase is required to enhance the image quality and to get the segmentation results more accurate. Imaging techniques play an important role by helping to perform a digital mammogram, for such images where abnormal areas cannot be identified physically but can be seen on a enhanced mammogram [10]. The main objective of this method is to enhance the standard of the image quality to make it prepared for more process by removing the unrelated and surplus elements [1-6] within the background of the mammogram. Breast segmentation consists of breast border contour extraction, muscle part extraction by removing unwanted data such as labels, patient name and tape artifacts can be achieved. On preprocessed images obtained from the digital mammography devices segmentation process is much easier. The databases for mammography imaging Mini-MIAS [18] used for experimental results.

2. PREPROCESSING
The basic purpose of medical image processing is to diagnose and examine medical images in a more effective, accurate and efficient manner. This can be obtained through the process of image enhancement shown in figure 1 and preprocessing techniques shown in the table 1.

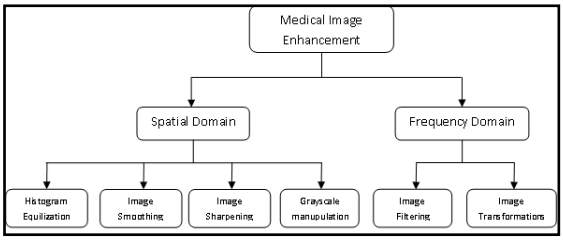


Fig 1. Types of image enhancement techniques

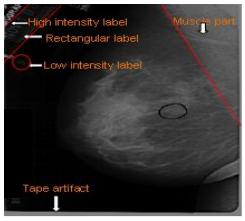


Fig.2 Types of artifacts and noises observed in mammogram image

In the previous survey paper [4] has introduced many of the preprocessing techniques used earlier for mammography images. They have discussed many of the methods. This paper intended to review the latest techniques have been introduced to preprocessing mammography with efficient transformations.

2.1 Active Contour Method
Active contour technique [6] incorporates in two stages, first it removes hurdles like labels, patient name, scanning and taping artifacts using an automatic algorithm based on thresholding. Applying active contour and automatic stopping algorithm it generates the contour that contains the boundary of the muscle part. It extracts the muscle part binary image from the contour. Finally, combine with the muscle part binary image and the original image it obtains the required image and the pectoral muscle.

2.2 Seeded Region Growing Technique
Initially, a contrast enhancement by using the contrast limited CLAHE technique [21]. Then defining the rectangle to

separate the muscle part of the ROI and finally suppress the muscle part using modified Seeded Region Growing (SRG) algorithm. The muscle part have a rather higher intensity compared to the remainder of the breast tissue and seem in upper left corner of MLO view of the mammogram (the orientation of right breast mammograms are flipped horizontally to left). After contrast enhancement applied SRG algorithm on the extracted rectangle, to suppress the pectoral muscle. Region growing is a procedure that groups pixels or sub regions into larger regions based on predefined criteria [3, 8, 10]. The basic approach is to start with a seed point and from this it grows into regions by appending to each seed those neighboring pixels that have properties similar to that seed. Selection of the seed depends on the nature of the problem. A problem in the seeded region growing algorithm is the formulation of stopping rules. Basically, growing a region should stop when no more pixels satisfy the criteria for inclusion in that region. The steps followed in the seeded region growing algorithm are

Step 1: Read the inverted right angle triangle area from left top corner.

Step 2: The selection method is obtained by subtracting average intensity from pixel intensity and then dividing it by the difference between maximum intensity and average intensity.

Step 3: If the pixel intensity is greater than 0 and less than equals to 1, then pixel will be merged to the region growing and the intensity value will be 0 else it will be remain unchanged.

$$0 < \frac{I(x,y) - Iavg}{IMax - Iavg} \leq 1 \qquad (1)$$

The I(x, y) is the intensity of a pixel, IAvg, IMax is average intensity and maximum intensity calculated earlier from the traversed seeds. The region growing will be restricted within the boundary of the inverted right angle triangle area of the cropped image. Using this algorithm pectoral muscle, that is an unwanted part of the mammogram will be eliminated.

2.3 Morphological Operations

Morphological operations are basic and efficient preproprocessing techniques for mammography. The basic morphological operations are dilatation, erosion, opening and closing [16]. Erosion is used to fill the gaps or holes in the mammography then dilated gradient mask shows the cells nicely but still some holes remains in the interior of the tumor. To eliminate these holes light structure is connected to image border [1]. In image the structuring element disc is applied with radius 5 starting with dilation followed by erosion, opening, closing and ends with Top hat and Bottom hat transforms [2]. It was originally defined for binary images, later on extended to grayscale images, and then for a complete lattices [5]. Artifacts can also be removed by morphological open operation followed by reconstruction operation [8].

2.4 Curve Fitting

The curve fitting techniques involve computing matrix and derive it in order to determine the highest inflection point at breast tissues pectoral muscle interface. By referring to equation (2) slop of breast equation was obtained at a specific calculated point [3].

$$Mb = 3aI^2 + 2bI + c \qquad (2)$$

Where, I is highest variation point.

Through this slope the inverted right angle was found that was the pectoral muscle.

2.5 Image Cropping and Pruning

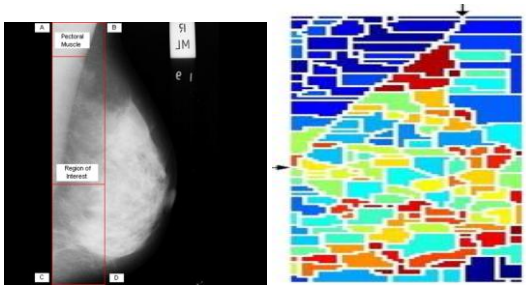
The spatial domain to reduce the effect of higher brightness and lower darkness in the image [7]. Min-Max techniques helps assign lower darkness value is 0 which is min and higher brightness value is 255 which max. The intermediate pixel values are changed according to equation (3).

$$Dp = \frac{D - Dmin}{Dmax - Dmin} * 255 \qquad (3)$$

Where D is intensity of the image, Dmin is minimum intensity and Dmax is maximum intensity.

2.6 Watershed Transformation

When the gradient of the ROI of a mammogram [9] was treated with the watershed transform [16, 28, 30], the results, given in Figure 3, showed a strong indication of the presence of the pectoral muscle boundary with a set of properties: (i) There is a unique continuous watershed line (hereafter it is called the watershed line of interest) which starts from the top and ends at the left of the image [19]. The width, extending from the left-most pixel of each row (starting from the top row) to its current position, is gradually decreasing and becomes zero when it reaches the left-most position at the end. (ii) The watershed line of interest encloses a triangular region covering the left top region of the image. It has a curved shape. (iii) The pectoral muscle is oversegmented; this is caused by several irrelevant regional minima within the pectoral muscle. The irrelevant regional minima may appear due to noise, local variations, etc.



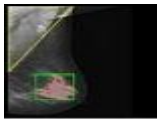
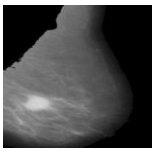
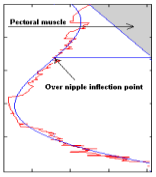
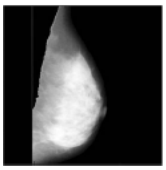
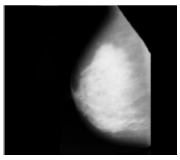
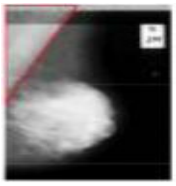
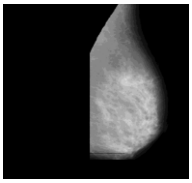
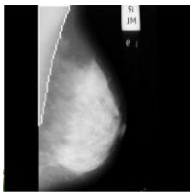
(a) **(b)**

Fig 3: (a) Original image labeled with ROI (b) Watershed transform the muscle part edge is marked with arrow

3. RESULT AND ANALYSIS

In this paper preprocessing techniques for mammogram images are compared and shown in the table 1.

Table 1. Comparison of existing preprocessing techniques

Sl.No	Authors name	Methods	Results	Advantages
1	In 2014 Naishil N. Shah, et al [1]	Edge Detector, Morphological filtering and Region Growing		morphological filtering helps to includes grayscale dilation, hole-filling, grayscale erosion and clear border.
2	In 2014 K.Vennila,et al [2]	Otsu Thresholding and Dilation followed by erosion, opening, closing and ends with Top hat and Bottom hat transforms.		The artifacts are removed by using Matlab Command bwareaopen which keeps the breast portion and removes all the non-breast profile.
3	In 2014 Farag H.et al.[3]	Curve fitting algorithm		It helps to detect the pectoral muscle upper location and ower limiting point against the nipple location
4	In 2013, Prakash Bethapudi, et al [5]	Morphological Techniques and Thresholding		Efficient in getting the malignant breast cancer region
5	In 2013, Farhan AKRAM et.al [6]	A Preprocessing for the CAD System of Mammograms Using the Active Contour Method		The image with a fair intensity difference between the pectoral muscle region and the remaining breast region
6	In 2012, Rajashekar K.R [7]	Image cropping operations , pruning and Min max equalization mathods.		It helps to remove artifacts, noise and background.
7	In 2012, R.Poongothai, et al [8]	Morphological open operation followed by reconstruction operation, Raster Scanning and Thresholding methods		Rectangle shaped image helps retrive the entire pectoral muscles.
8	In 2012, Indra Kanta Maitra, Sanjay Nag, Samir Kumar Bandyopadhyay [10,23]	CLAHE and SRG		Robust techniques which include contrast enhancement and denoising also need not perform separately.
9	In 2011, K. Santle Camilus, et al [9]	Watershed transformation and merging algorithm		This new merging algorithm helps choose a pectoral muscle region as the initial seed by using one of the attributes of the pectoral muscle.

Although the above conferred techniques are good in enhancing and removing artifacts. However, in some cases the preprocessing techniques may remove the breast details along with pectoral muscle; they may destroy the shape of the tumor. This may leads to miss classification and wrong decision in diagnosis. The transformation techniques can be taken into account because the Contourlet is best to represent the edges than the wavelet and region growing is good enough to add most similar part for muscle extraction.

4. CONCLUSION

The cervical, breast and endomental cancers are most often affect to women. Breast cancer is the second leading cause for women surveillance. For early detection of breast cancer the basic screening test is mammography. Mammography is poor contrast, noisy, blurred and labeled by patient name, age etc. To remove these artifacts the preprocessing technique is needed. In this review paper the existing preprocessing techniques have been discussed with their advantages. But for pectoral muscle there should be taking care of breast part because along with muscle some time breast details also are removed. Hence, for our next research work we are more concentrating on pectoral muscle removal and contrast enhancement. The region growing techniques and CLAHE respectively will be the best suitable techniques for this process.

5. REFERENCES

[1] Naishil N. Shah, Tushar V. Ratanpara, C.K.Bhensdadia, "Early Breast Cancer Tumor Detection on Mammogram Images," International Journal of Computer Vol. 87, No.14,pp.14-18, February 2014.

[2] K.Vennila, K.Sivakami, R.Padmapriya," Detection of Mass in Digital Mammograms," International Journal of Computer Applications, Vol.104, No.5,pp.22-23, October 2014.

[3] Farag H. Alhsnony, Maher. G. M. Abdolrasol, Samei. G. M. Abadelrsool, "Auto-Identification of Pectoral Muscle Region in Digital Mammogram Images," International Journal of e-Education, e-Business, e-Management and e-Learning, Vol. 4, No. 1, pp.15-18, February 2014.

[4] K. Meenakshi Sundaram, D. Sasikala, P. Aarthi Rani, "A Study On Preprocessing A Mammogram Image Using Adaptive Median Filter," IJIRSET, Vol. 3, Issue 3, pp.10333-10337,March 2014.

[5] Prakash Bethapudi Member IEEE, Dr. E. Srinivasa Reddy, Dr.Madhuri.P, "Detection of Malignancy in Digital Mammograms from Segmented Breast Region Using Morphological Techniques," IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE) , Vol. 5, Issue 4, pp.09-12, May- Jun 2013.

[6] Farhan Akram, Jeong Heon Kim, Inteck Whoang, and Kwang Nam Choi, " A Preprocessing Algorithm for the CAD System of Mammograms Using the Active Contour Method," Applied Medical Informatics Vol. 32, No.2, pp: 1-13,2013.

[7] Rajashekar K.R, " Detection and Classification of Tumors in a Digital Mammogram," Special Issue of International Journal of Computer Application,s International Conference on Electronic Design and Signal Processing (ICEDSP),pp.25-28, 2012.

[8] R.Poongothai, I.Laurence Aroquiaraj, D.Kalaivani, " A Novel Pectoral Removal Approach in Digital Mammogram Segmentation and USRR Feature Selection," International Journal of Engineering Research and Applications (IJERA) Vol. 2, Issue4, pp.1846-1851, July-August 2012.

[9] K. Santle Camilus, V. K. Govindan, P.S. Sathidevi, " Pectoral muscle identification in mammograms", Journal of Applied Clinical Medical Physics, Vol. 12, No. 3, pp. 215-230, 2011.

[10] Indra Kanta Maitra, Sanjay Nag, Samir Kumar Bandyopadhyay, "Technique for preprocessing of digital mammogram, " computational methods and programs in biomedicine 107 , pp. 175–188,2012.

[11] Samir Kumar Bandyopadhyay, "Pre-processing of Mammogram Images," International Journal of Engineering Science and Technology, Vol. 2,No.11, pp. 6753-6758,2010.

[12] D.Narain Ponraj, M.Evangelin Jenifer, P. Poongodi, J. Samuel Manoharan, "A Survey of the Preprocessing Techniques of Mammogram for the Detection of Breast Cancer," Journal of Emerging Trends in Computing and Information Sciences, Vol. 2, No. 12, pp.656-664, December 2011

[13] Maciej A. Mazurowski, Joseph Y. Lo, Brian P. Harrawood, Georgia D. Tourassi, "Mutual information-based template matching scheme for detection of breast masses: From mammography to digital breast tomosynthesis," Journal of Biomedical Informatics 2011.

[14] Jwad Nagi, "Automated Breast Profile Segmentation for ROI Detection Using Digital Mammograms," IEEE EMBS Conference on Biomedical Engineering & Sciences (IECBES), pp. 87 - 92, 2010

[15] Aziz Makandar, Bhagirathi Halalli, " Comparision of denoising techniques for mammography," pre science congress of Indian Science Congress 2015 aurangabad chapter, 2014.

[16] R.C., Gonzalez, R.E. Woods, Digital image processing, 3rd edition, pp. 104-289, 462-776, 2013.

[17] National Cancer Institute (NCI) Web site, http://www.cancernet.gov.

[18] J. Suckling, J. Parker, D.R. Dance, S. Astley, I. Hutt, C.R.M. Boggis, I. Ricketts, E. Stamatakis, N. Cernaez, S.L. Kok, P.Taylor, D. Betal, J. avage, "The mammographic image analysis society digital mammogram database," in: Proceedings of the 2nd International Workshop on Digital Mammography, York, England, 10–12 July 1994, Elsevier Science, Amsterdam, The Netherlands, , pp. 375–378, 1994.

[19] H.S.Sheshadri and A. Kandaswamy, "Detection of Breast Cancer Tumor based on Morphological Watershed Algorithm," ICGST-GVIP Journal, Volume (5), Issue (5), May 2005.

[20] B. Senthil kumar, G. Umamaheswari, "A Novel Edge Detection Algorithm for the Detection of Breast Cancer," European Journal of Scientific Research, Vol.53 No.1, pp.51-55, 2011.

[21] Antonis Daskalakis, et al., "An efficient CLAHE-based, spot-adaptive, image segmentation technique for improving microarray genes' quantification," in: 2nd International Conference on Experiments/Process/

SystemModelling/Simulation and Optimization, Athens, 4–7 July, pp. 1-6, 2007.

[22] Aziz Makandar, Daneshwari Mulimani, Mahantesh Jevoor, "Comparative Study of Different Noise Models and Effective Filtering Techniques," International Journal of Science and Research (IJSR),Vol. 3 Issue 8,pp 458- 464, August 2014.

[23] Indra Kanta Maitra, Sanjay Nag, Samir K Bandyopadhyay, "Accurate Breast Contour Detection Algorithms in Digital Mammogram," International Journal of Computer Applications, Vol. 25, No.5, pp. 207-215, July 2011.

[24] Roshan Dharshana Yapa, Koichi Harada, "Breast Skin-Line Estimation and Breast Segmentation in Mammograms using Fast-Marching Method," International Journal of Biological and Life Sciences, pp. 217-225, 2007.

[25] Aziz Makandar, Bhagirathi Halalli, "Image Enhancement Techniques using Highpass and Lowpass Filters," International Journal of Computer Applications, Vol. 109, No. 14, pp. 12-15,January 2015

[26] Samir Kumar Bandyopadhyay, "Formation of Homogeneous Blocks for Segmentation of Mammograms," International Journal of Engineering Science and Technology, Vol. 2 (12), pp.7444-7448, 2010.

[27] H. Mirzaalian, M.R Ahmadzadeh.S, Sadri, M., Jafari, "Pre-processing Algorithms on Digital Mammograms," IAPR Conference on Machine Vision Applications, Tokyo, Japan, pp. 118-121, May 16-18, 2007.

[28] R.B. Dubey, M. Hanmandlu, S.K. Gupta, "A comparison of two methods for the segmentation of masses in the digital Mammograms," computerized Medical Imaging and Graphics, No. 34, pp.185–191, 2010.

[29] Sumit Chopra, Pankaj Bhambri, Baljit Singh, "Segmentation of the Mammogram Images to find Breast Boundaries," IJCST Vol. 2, Issue 2, pp.164-167, June 2011.

[30] Luc Vincent and Pierre Soille, "Watersheds in Digital Spaces: An Efficient Algorithm based on Immersion Simulations," Luc Vincent and Pierre Soille, IEEE transactions on pattern analysis and machine intelligence, Vol.13, No.6, pp. 583–98, 1991.

[31] Dr.Aziz Makandar, Anita Patrot , Bhagirathi Halalli, "Color Image Analysis and Contrast Stretching using Histogram Equalization," International Journal of Advanced Information Science and Technology (IJAIST), Vol.27, No.27,pp. 119-125, July 2014.

[32] Shavi Gupta, Mohd. Sadiq, Mona Gupta and Naseem Rao, "Semi Automatic Segmentation of Breast Cancer for Mammograms Based on Watershed Segmentation," Proceedings of the 5th National Conference; INDIACom-2011, Computing For Nation Development, March 10 – 11, 2011.

[33] Huo Z, Giger ML, Vyborny CJ. "Computerized analysis of multiple-mammographic views: Potential usefulness of special view mammograms in computer-aided diagnosis," IEEE Transactions on Medical Imaging, Vol.20, pp.1285-1292, 2001.