

Pre-processing of Mammography Image for Early Detection of Breast Cancer

Aziz Makandar

Department of Computer Science
Karnataka State Women's University
Vijayapura

Bhagirathi Halalli

Department of Computer Science
Karnataka State Women's University
Vijayapura

ABSTRACT

Breast cancer is one of the most prevalent causes of death among women worldwide. Hence, the early detection helps to save the life of the women. Mammography is the basic screening test for breast cancer. It consist many artefacts, which negatively influences in detection of the breast cancer. Therefore, removing artefacts and enhancing the image quality is a required process in Computer Aided Diagnosis (CAD) system. The accuracy and efficiency of the CAD is increased by providing exact Region of Interest (ROI). Extracting ROI is a challenging task in preprocessing because the presence of pectoral muscle influences the detection of abnormality. Here, the proposed show that the wiener filter and Contrast Limited Adaptive Histogram Equalization (CLAHE) techniques efficiently aids for enhancing the quality of the image, thereby it also removes the unwanted background and the pectoral muscle by using thresholding and modified region growing technique respectively. Furthermore, the proposed algorithm was tested on mini-MIAS database; the result obtained was compared with completeness and correctness for pectoral muscle removal and was reported as 98% and 97% respectively. Collectively, these results suggest that the proposed method is well suited for improving the quality of mammography image for Auto-CAD system.

General Terms

Computer Science, Digital Image Processing, Cancer Imaging.

Keywords

Breast Cancer, Mammography, Preprocessing, Region Growing, Wiener Filter.

1. INTRODUCTION

Breast cancer ranks as second leading cause of death in women worldwide [1]. According to American Cancer Society, about 1 in 8 women will have breast cancer in her lifetime and only 5% to 10% of breast cancers occur in women with clearly defined genetic link [2]. Hence, the early detection will help to have better quality of life, economical treatment and mental peace of patient and family. With a low dose of X-ray imaging, mammography is most basic screening test for breast cancer and also records better visualized internal details of the breast [3]. Usually mammography images consist of many artefacts and noises and makes medical images too difficult to detect and understand the cancer at the primary stages [4]. Therefore, standardization of image quality and extraction of ROI is essential to limit the hunt for abnormalities.

Pre-processing of mammography helps to identifies abnormal areas that cannot be experienced physically or visualized, but can be detected through CAD. Imaging techniques play an important role in CAD. The CAD helps doctors and

radiologists to diagnose and identify the abnormality quickly and easily [5]. Here, the proposed method helps standardize the image quality and extracting the target ROI. By breast part extraction, muscle part removal and enhancement of mammogram. The proposed algorithm helps radiologist to diagnosis the disease more accurately and takes decision swiftly. The types of noise observed in the mammogram image are marked in the Figure 1. In the proposed algorithm, it abolish all these unwanted and surplus noises from the mammogram image.

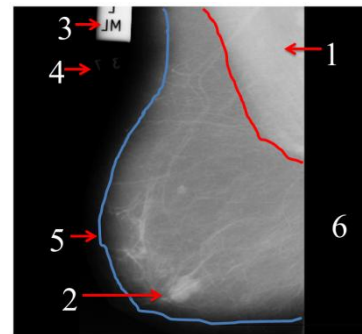


Fig 1 .Types of noises observed in original image and marked with numbers as 1. Pictorial Muscle, 2.Tumor, 3.High Intensity, 4.Low Intensity, 5. Breast Part and 6. Background.

Several methods have been reported for pre-processing mammography images since 1980 because of its influences in detection of cancer [7-9]. The techniques like adaptive median filter, mean filter, adaptive mean filter, histogram equalization, histogram modified local contrast enhancement, breast region and pectoral muscle extraction, CLAHE technique and morphological have been discussed earlier [10-12].In this paper explored that how the selection of significant parameters for wiener filter and CLAHE influences in the preprocessing of CAD system. Section 2 and 3 indicates the dataset used and proposed methodologies. Section 4 explores the test and results of proposed method.

2. DATASET USED

For testing and analysis of the proposed algorithm, randomly selected 100 images of Mammographic Image Analysis Society (mini-MIAS) database, organized by J Suckling et al.in 1994 [13]. The team developed a database of digital mammograms. Films taken in the UK National Breast Screening Programme (NBSP).Images were digitized to 50 micron pixel and represented with an 8-bit word of each pixel. It reduced to a 200 micron pixel and padded, all the images are in 1024×1024 size. The database consists of 322 digitized mammograms (Among which it consist 202 normal and 120 abnormal images). It also includes radiologist's markings on the locations of abnormalities if present. Mammographic

images are available online in Pilot European Image Processing Archive (PEIPA) at the University of Essex.

3. PROPOSED METHOD

The main purpose of a breast cancer CAD system is to help the radiologist and doctors to take decision swiftly. By providing exact ROI will help to identify abnormality. The proposed method works in three stages as explained in figure 2. The first step is to remove the back ground artefacts (label 3,4 and 6) identified in the figure 1. The second step is to reduce the pectoral muscle (label 1) identified in figure 1 and the digital mammography enhanced by using wiener filter and CLAHE.

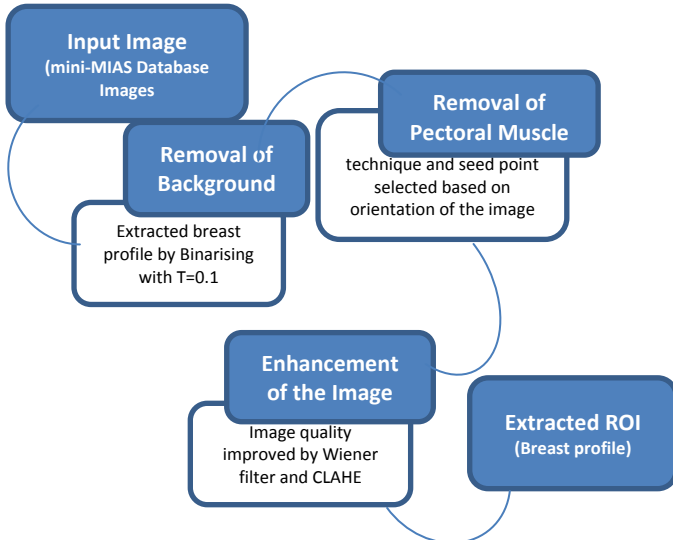


Fig 2 Block diagram of proposed method

3.1 Background Removal

Initially image was binarized with threshold value 0.1 then the connected component organized in descending order to extract the largest blob which is the breast profile but consists of pectoral muscle.

3.2 Suppression of pectoral muscle

The second stage was used to reduce the pectoral muscle part by using modified region growing technique. The seeded region growing is one of the image segmentation methods [14], it works in two ways based on selected pixel locational value and other is selection of seed point. The seed point may be selected adaptively or manually. In the proposed method, seed point is selected automatically by considering the orientation of the mammography. This approach determines the neighboring pixels of seed point and examines whether the next pixels should be added to the region or not. The process is iterated till to extract the complete ROI [15].

3.3 Image enhancement

Third stage was used to enhance the quality of the image using wiener filter and CLAHE [16].

4. RESULTS AND DISCUSSION

The proposed algorithm was experimented on 100 mammogram images of mini-MIAS database that covers all types of images such as Fatty, Glandular and Dense and the proposed techniques automatically extracted the ROI as shown in the figure 3.

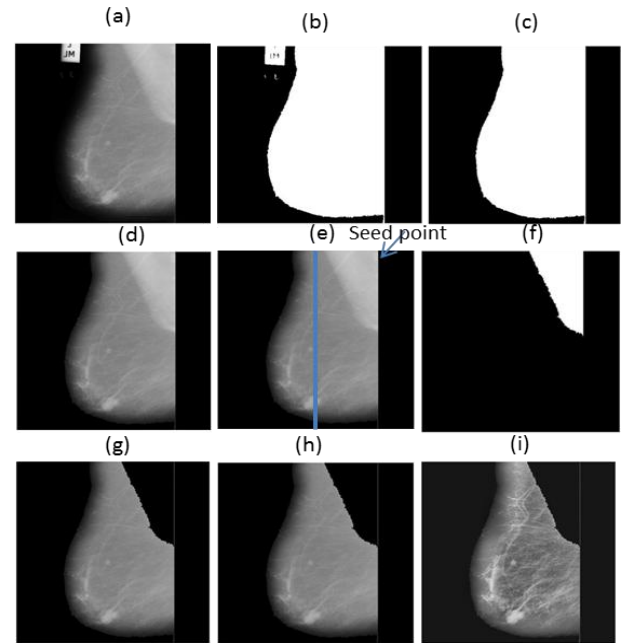


Fig 3 Experimental results proposed method (a) Original Image (b) Binary Image with threshold value 0.1 (c) Breast Part extracted (d) Multiplication of (a) and (c) which consist only breast part without background (e) seed point marked for region growing, (f) pectoral muscle segmented, (g) suppressed from original image, (h) wiener filter, (i) result of CLAHE

The background artefacts are removed by binarization with threshold value 0.1 (fig 3b) and all the connected components are organized in largest to smallest in size to extract the largest blob (fig 3c). Then that blob is multiplied with original image to get original breast profile (fig 3d).

The region growing method used to reduce the pectoral muscle part. The proposed method helps to select the seed point automatically (fig 3e) to remove the pectoral muscle (fig 3f) using modified region growing technique. The conventional selection of seed point is modified based on the orientation of the image. The mini-MIAS dataset consist either left oriented or righted oriented images. Hence, the seed point is either left topmost or right topmost first nonzero pixel. Orientation of the image found by dividing image into half and counting the non-zero pixels if left oriented, left part consist more pixels else right part consist more pixels.

The image quality was enhanced by using wiener filter (fig 3h) and CLAHE (fig 3i). The evaluation of image quality is crucial for medical imaging systems such as compression, transmission and enhancement [17-20]. Quality of the image measured by traditional methods [21] Root Mean Square Error (RMSE), Peak Signal to Noise Ratio (PSNR) and Image Quality Index (IQI) on different images with different levels of Contrast Index (CI). The RMSE values and PSNR values are reciprocal to each other and calculated by using following equation 1 and 2 respectively.

$$RMSE = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N (x(i,j) - y(i,j))^2} \quad (1)$$

$$PSNR = 10 \log \frac{(2^n - 1)^2}{RMSE} \quad (2)$$

IQI is measured if $x=\{x_i \mid i=1,2, \dots, M\}$ and $y=\{y_i \mid i=1,2, \dots, N\}$ are original and test image signals respectively. The IQI is measured as equation 3.

$$Q = \frac{4\sigma_{xy} x' y'}{(\sigma_x^2 + \sigma_y^2)[x'^2 + y'^2]} \quad (3)$$

Where

$$x' = \frac{1}{M} \sum_{i=1}^M x_i \quad \text{and} \quad y' = \frac{1}{N} \sum_{j=1}^N y_j \quad \sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - x')(y_i - y'), \quad \sigma_x^2 = \frac{1}{M} \sum_{i=1}^M (x_i - x')^2 \quad \text{and} \quad \sigma_y^2 = \frac{1}{N} \sum_{i=1}^N (y_i - y')^2$$

The strong reason is for using wiener filter and CLAHE for image enhancement. Comparing the median filter, adaptive min max and wiener filter we got high PSNR for all the images tested as shown in the figure 4.

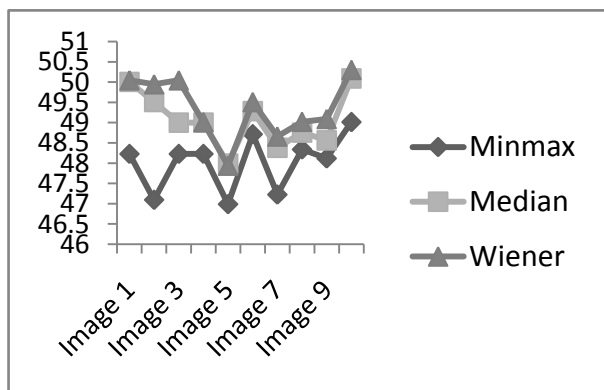


Fig 4 Comparison of different filters

The above graph strongly proves that Wiener filter is suitable for noise removal of mammography image because it gives and high PSNR compared to min-max and median filter. Once wiener filter selected, tested with different filtering mask from [1 1] to [8 8] to select significant filter mask (fig 4) for wiener filter.

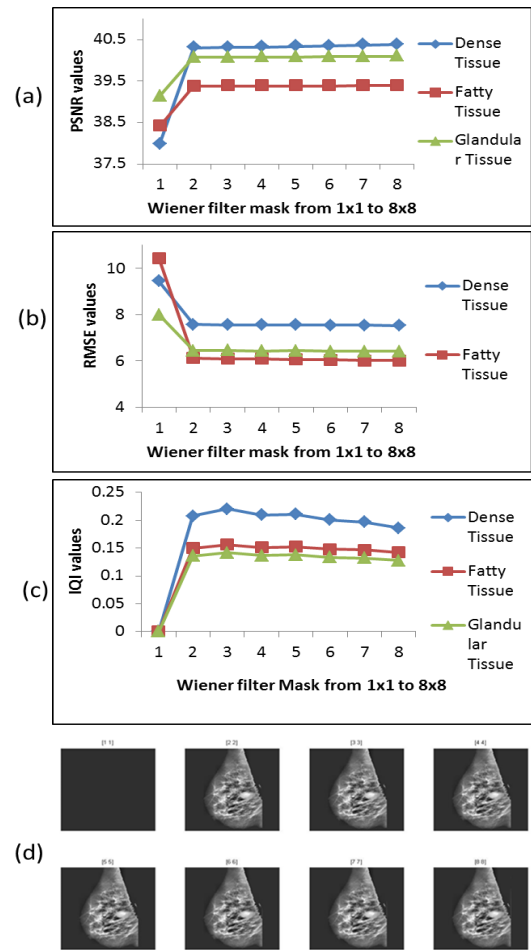


Fig 5. Selection of significant filter mask for wiener filter (a)PSNR (b)RMSE and (c)IQI values of different filter masks from [1 1] to [8 8] for dense, fatty and fatty glandular tissues.

As mask of wiener filter increased the PSNR value (fig5a) is increased for mast [1 1], [2 2], and [3 3] whereas the RMSE (fig 5b), IQI (fig 5c) values were reduced. Hence, [3 3] mask selected as significant filter mask for wiener filter. However, mask increased beyond significant, PSNR increased but image gets blurred (fig 5d). Similarly for the contrast index (CI) values, the default CI is suitable as compared with our results. For CI 0.2, PSNR increased and continued with slight increase (fig6a) RMSE (fig 6b) and IQI (fig6c) reduced from the CI 0.2. Hence, contrast index 0.2 selected as significant value.

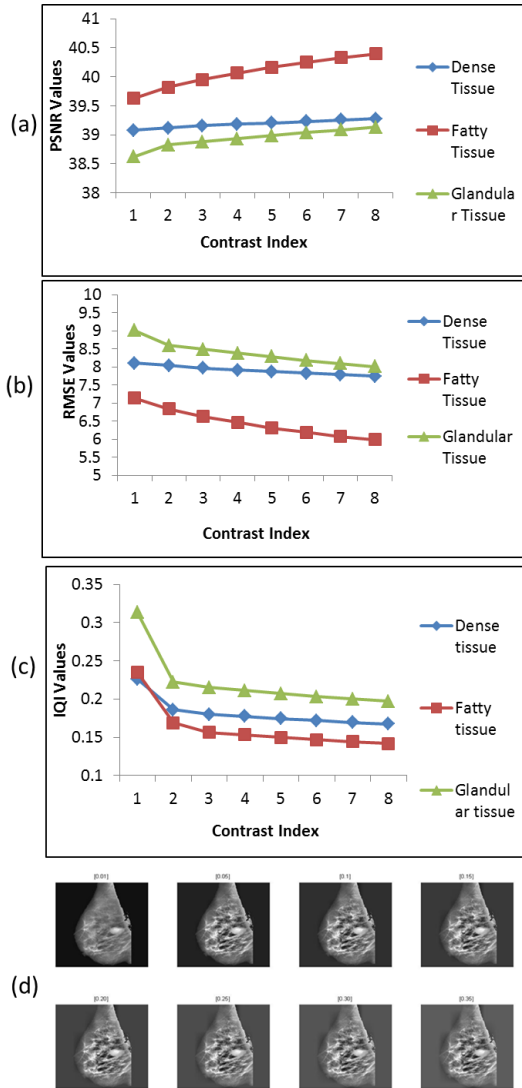


Fig 6. Selection of significant CI for CLAHE (a)PSNR (b)RMSE and (c)IQI values of different CI from 0.1 to 0.8 for dense, fatty and fatty glandular tissues.

Efficiency of proposed algorithm is measured in terms of correctness and completeness by true positive (TP), False Positive (FP) and False Negative (FN). TP is assigned to proper segmentation, FP is assigned to over segmentation and FN is assigned to under segmentation using equation 4 and 5.

$$correctness = \frac{TP}{TP+FP} \quad (4)$$

$$completeness = \frac{TP}{TP+FN} \quad (5)$$

An experimental result shows that among 100 images 97 are properly segmented and over segmented and 2 images are under segmented. Hence the correctness is 0.9897 and completeness is 0.9797 accuracy of the algorithm is 98% and 97% respectively.

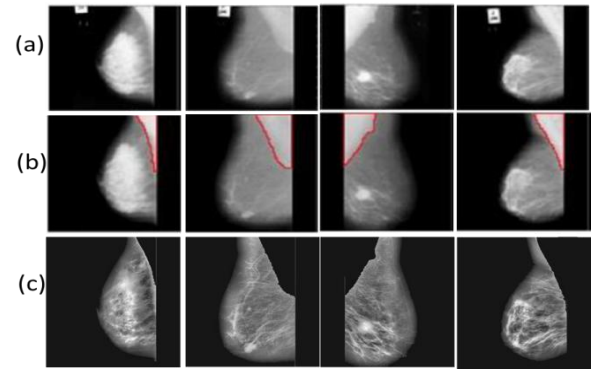


Fig 7. Comparison of result (a) Original Images (b) Farhan et al. results (c) Results of our proposed Method

In contrast to Farhan et.al [22], the proposed technique gives more accurate results. Figure 7 illustrates the comparison results between existing work (fig7b) and proposed work (fig7c). These results suggest that, our current study has convincingly enhanced the quality of the image with better in contrast.

5. CONCLUSIONS

Breast cancer is leading cause of death for women. Timely screening may help to detect the abnormality but current screening method, the mammography is low dose x-ray image. It may miss the small changes in the breast because of low dose x-ray it generates poor contrast image. The missed changes may lead for serious problem. Hence, CAD system helps to identify the diminutive changes in the breast. This early detection may save the life of the patient. The proposed method of pre-processing presented with removal of background artefacts, pectoral muscle suppression and image quality enrichment helps much in early detection. Our results show that, significant filter mask [3, 3] for wiener filter and CI is 0.2 for CLAHE are influencing factors for enhancement of mammography. The complete pectoral muscle was reduced by modified region growing techniques. The proposed method tested on images of Mini-MIAS database, ROI extracted from all the images accurately and proved to be suitable for CAD system of early detection of breast cancer. Collectively, these results prove that effective and convenient assistance for medical diagnosis. Hence, the proposed method definitely can be considered for automated detection of abnormality like benign, malignant and micro calcifications.

6. REFERENCES

- [1] National Cancer Institute (NCI) Web site, <http://www.cancernet.gov>
- [2] Union for International Cancer Control, <http://timesofindia.indiatimes.com/city/indore/37-pros-face-risk-of-cancer-Survey/articleshow/29830667.cms>
- [3] Indra Kanta Maitra, Sanjay Nag, Samir Kumar Bandyopadhyay, Technique for preprocessing of digital mammogram, computational methods and programs in biomedicine 107 (2012), pp. 175–188.
- [4] Aziz Makandar and Bhagirathi Halalli, A Review on Preprocessing Techniques for Digital Mammography images, International Journal of Computer Applications (IJCA) National conference on Digital Image and Signal Processing, DISP 2015, pp.23-27.
- [5] Jwad Nagi, Automated Breast Profile Segmentation for ROI Detection Using Digital Mammograms, IEEE

- EMBS Conference on Biomedical Engineering & Sciences (IECBES 2010), Kuala Lumpur.
- [6] Samir Kumar Bandyopadhyay, Pre-processing of Mammogram Images, *International Journal of Engineering Science and Technology*, Vol. 2(11), 2010, pp. 6753-6758.
- [7] Barghout, Lauren, and Lawrence W. Lee, Perceptual information processing system, Paravue Inc. U.S. Patent Application 10/618,543.
- [8] Ferrari RJ, Rangayyan RM, Desautels J E, Borges RA, Frere AF, Automatic identification of the pectoral muscle in mammograms, *IEEE Transactions on Medical Imaging* 2004;23(2), pp.232-245.
- [9] Liu CC, Tsai CY, Liu J, Yu CY, Yu SS. A pectoral muscle segmentation algorithm for digital mammograms using Otsu thresholding and multiple regression analysis. *Computer and Mathematics with Applications* 2012;64(5), pp.100-1107.
- [10] D.NarainPonraj, 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, December 2011, pp.656-664.
- [11] 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).
- [12] K. MeenakshiSundaram, D. Sasikala, P. Aarthi Rani, A Study On Preprocessing A Mammogram Image Using Adaptive Median Filter, *IJIRSET*, Vol. 3, Issue 3, March 2014.
- [13] J. Suckling, J. Parker, D.R. Dance, S. Astley, I. Hutt, C.R.M. Boggis, I. Ricketts, E. Stamatakis, N. Cernaes, 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, 1994, pp. 375–378.
- [14] R.C., Gonzalez, R.E. Woods, *Digital image processing*, 2007. Third edition, pp.118-529
- [15] Linda G. Shapiro and George C. Stockman (2001), *Computer Vision*, New Jersey, Prentice-Hall, pp 279-325.
- [16] Aziz Makandar and Bhagirathi Halalli, “Breast Cancer Image Enhancement using Median Filter and CLAHE,” *International Journal of Scientific & Engineering Research*, Volume 6, Issue 4, pp. 462-465, 2015.
- [17] Makandar, Aziz Ur Rahaman, and K. Karibasappa, Wavelet Based Medical Image Compression Using SPHIT, *Journal of Compute and Mathematical Sciences* Vol 1.7(2010): pp. 769-924.
- [18] B. Senthilkumar, G. Umamaheswari, A Novel Edge Detection Algorithm for the Detection of Breast Cancer, *European Journal of Scientific Research*, ISSN 1450-216X Vol.53 No.1 (2011), pp.51-55.
- [19] Makandar, Aziz, and Bhagirathi Halalli, Image Enhancement Techniques using Highpasand Lowpass Filters, *International Journal of Computer Applications*109.14 (2015): 21-27.
- [20] 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, 2007.
- [21] Samir Kumar Bandyopadhyay, Pre-processing of Mammogram Images, *International Journal of Engineering Science and Technology*, Vol. 2(11), 2010, pp. 6753-6758.
- [22] 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 2013, pp: 1-13.