

Threshold Based Segmentation Technique for Mass Detection in Mammography

Aziz Makandar*, Bhagirathi Halalli

Department of Computer Science, Karnataka State Women's University, Vijayapura, Karnataka, India.

* Corresponding author. Tel.: 8762616111; email: azizkswu@gmail.com

Manuscript submitted July 29, 2015; accepted December 18, 2015.

doi: 10.17706/jcp.11.6.472-478

Abstract: Breast cancer is second leading cause of death among women. Mass of the cancer is initially originates from a single cell but slowly increases in size by rapid multiplication of cells to produces symptoms. Most of the time cancer symptoms are identified at the late stage, when the tumor becomes bigger in size and treatment becomes invasive. Early detection of the cancer before the development of the symptoms may help in less number of modalities for the treatment. Screening is the basic procedure for identification of breast cancer at an earliest and mammography is an efficient screening method, in which abnormalities can be detected. However, it is difficult to identify the tumor in the breast tissue because tumors possesses equal intensity in the breast tissue and appears poor in contrast. Hence, the computer aided detection helps for physicians and radiologist to find abnormality at an earliest in the absence of any symptoms. In this study, we used segmentation algorithm to develop an efficient system to find abnormality at the earliest stage. The proposed segmentation algorithm detects clearly defined region of mass using morphological threshold based segmentation technique. The efficiency of the algorithm is measured with 55 images of Mini-MIAS database. These results showed satisfactory segmentation and the accuracy of the algorithm is 94.54% in identification of mass in mammography and false identification rate 5.45%. Thus, the proposed method is compared with traditional Otsu thresholding method, which is more effective comparing to Otsu thresholding segmentation results.

Key words: Breast cancer, mammography, mass segmentation, Otsu thresholding, morphological operations.

1. Introduction

Breast cancer is the most common cause of the death in women, according to a survey conducted by WHO in 2012; most of the younger women's are affected by the breast cancer. Maximum deaths are registered in India followed by China and USA [1]. In India, Mumbai is one of the leading cities of breast cancer deaths [2]. This implies that one-fourth among all cancer of women is breast cancer. Early detection plays a very important role in the diagnosis of breast cancer. 50% of cases could be solved if the patient undergoes for screening regularly, if failed to diagnoses at screening level, then it might lead to spreading the cause. Hence detecting abnormality without any symptoms and earliest may help to cure cancer.

Many works have been conducted on segmentation of mass in mammography, but it is difficult to extract the poor contrast masses from the mammography [3], [4]. The common mass segmentation approaches are based on region and edges. In region-based methods, regions are grown to segment the mass by adding nearest neighbor if it satisfies the thresholding criteria [5]-[8]. Edge based techniques identifies the masses

on sudden changes in the breast tissues. The region-based techniques such as region growing and clustering which uses k-means clustering [9] or wavelet-based techniques [10]-[12]. The most of the edge based techniques are using watershed transform and contour based segmentation techniques [13], [14]. The proposed method considered the threshold-based segmentation technique. To detect an abnormality, first we segment the mammography in several regions using threshold-based segmentation technique with morphological operations. Each region indicates its characteristics on which mass can be identified.

2. Image Segmentation

Segmentation is a process of partitioning the image into sub-regions [15]. Each sub-region helps to represent the information that it belongs and to get structuring elements of the target region [16]. Image segmentation is used to extract Region of Interest (ROI) and its boundaries from the images. Mammogram images it is important to distinguish the suspicious region from its surroundings. The shape, margin and intensity of tissue are the important criteria's to distinguish normality and abnormality in the breast tissue. Usually a high intensity and circular based objects are more likely to be ill-defined masses [17]. It consists of volume and shapemass is of convex kind of border shape and consist of volume and shape. A mass found at mammogram will normally describe by its intensity, shape, density and margin as shown in Fig. 1.

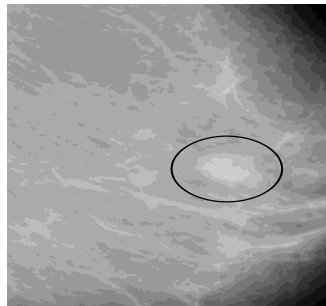


Fig. 1. Breast mass in mammography.

2.1. Otsu Thresholding

Image segmentation is the basic approach in medical image processing. Among all the methods of segmentations, threshold-based segmentation is common technique and Otsu thresholding segmentation is one of most commonly used a method. Otsu is an effective technique, it automatically selects the threshold and it works fairly on region based segmentation technique [18]. It overcomes the drawback of global thresholding by selecting optimal threshold value. Whereas global thresholding selects the average intensity of foreground and background as a threshold value. Otsu also helps discriminate the interclass threshold values and works on the normalized histogram, Then the global mean intensity (m_G) and Global Variance intensity (σ_G^2) of the pixel is calculated. The goodness of the threshold is calculated by using mean and variance by the equation 1.

$$\eta(k) = \frac{\sigma_B^2(k)}{\sigma_G^2} \quad (1)$$

where, $\sigma_B^2(K)$ is the variance between two classes that are foreground and background intensities.

$$\sigma_B^2(K) = \frac{(m_G p_1(k) - m(k))^2}{p_1(k)(1-p_1(k))} \quad (2)$$

For the computation of optimal threshold, the following condition should satisfy

$$\sigma_B^2(K^*) = \max_{0 < k < L-1} \sigma_B^2(K) \quad (3)$$

The Otsu produces smooth lines and fairly separated lines between the regions, but it does not work in the presence of noise. The distribution of the histogram is not balanced hence some information can be missed. Hence cannot be suited for mammography segmentation. Otsu thresholding can also be extended for multiple thresholding by selecting multiple regions.

2.2. Morphology Based Thresholding Segmentation

Thresholding is yet effective and simple method of segmenting the image into different regions. In proposed algorithm before applying thresholding to the image it transformed with watershed and morphological operations [19]. Watershed was originally proposed by Digabel and Lantuejoul. It is one of the useful concepts in image segmentation. Many modifications have been carried out on the watershed algorithm. Because it gets over segmented on the gray scale image. The concept of the watershed could be illustrated by geography as the representation of a topographical representation of the image. If the image is of the landscape it start filling with the water from the minimum gray value in the region of interest [20]. When water fills up two or more regions it start merging, so we have to prevent merging by increasing the margins of basins till the high intensity. To control this more commonly creates a dam at points where water of two different regions meets. These regions are considered as catchment basins and the dams or the watershed lines which divide two different regions based on similarity satisfy the region. Most of the time it over segments the image, to control over segmentation mathematical, morphological and logical operations are used. The proposed work is a threshold-based segmentation and it is modified to extract the ROI with watershed transform and morphological operation.

2.2.1. Mathematical morphological operations

Mathematical morphological operations help to structure elements and measure the shape of the image. It also helps to refine the characteristics of the image in order to maintain the image data and characters [21]. This work considered, two morphological operations are opening and closing. Opening and closing are most commonly used operators in morphology. They have implemented by the basic operations such as erosion and dilation [22]. The morphological opening is denoted by the equation 4 and it is achieved by erosion followed by dilation.

$$A \circ B = (A \ominus B) \oplus B \quad (4)$$

Morphological opening helps to smoothen the edges and breaks the weak connections. Also helps remove the unwanted regions which do not contain structuring elements.

The morphological closing operation is denoted by equation 5 and it is achieved by dilation followed by erosion. It is union of all translations of B without overlap on A.

$$A \cdot B = (A \oplus B) \ominus B \quad (5)$$

It helps to join the weak edges and fill the breaks in the edges. Also helps filling gaps and small holes in the structuring elements.

3. Dataset and Methodology

3.1. Dataset

For measuring the efficiency of the proposed algorithm, we used the images of Mammographic Image Analysis Society (mini-MIAS) database [23]. The database consists of 322 digitized mammograms (Among which it consist 202 normal and 120 abnormal images). The digitized Images that were 50 micron pixel and represented with an 8-bit word of each pixel and all the images were in 1024×1024 size and padded to a

200 micron pixel. It also includes radiologist's mark of centroid and radius on the locations of abnormalities if present (ref) were selected for our comparative studies. Here, we have selected first 55 abnormal images from the database which consist all kind of images as listed in the Table 1.

Table 1. Images Selected for Testing from Mini-MIAS Database

Abnormality type/Tissue type	Fatty	Glandular	Dense
Benign	16	12	6
Malignant	7	7	7

3.2. Methodology

The proposed method helps to extract the suspicious region from the breast. The proposed algorithm is schematically illustrated in the Fig. 2. Initially the mammogram images are poor in contrast and consist of many artifacts. Before processing with these images it is necessary to remove artifacts and pectoral muscle because they influence a lot in detection of abnormality. Hence the preprocessing step is carried out.

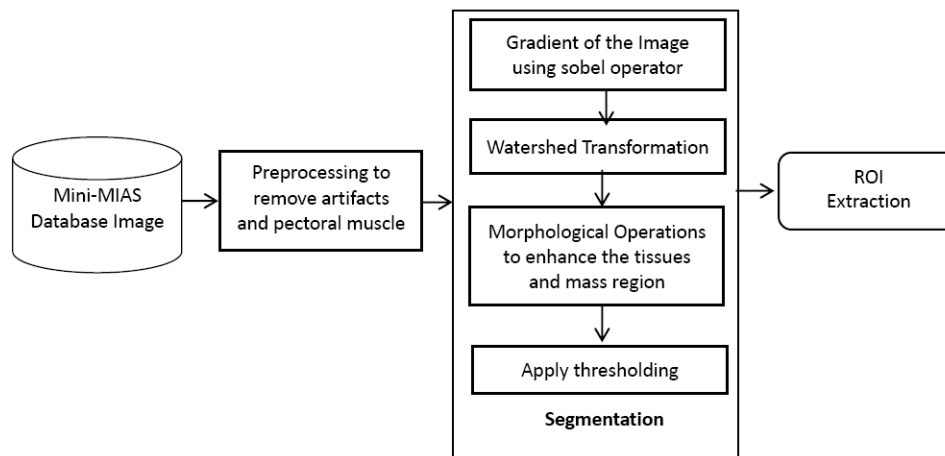


Fig. 2. Block diagram of proposed methodology.

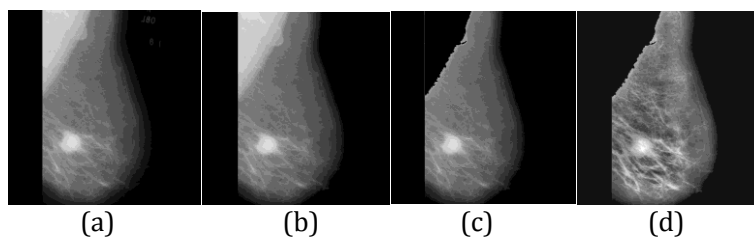


Fig. 3. (a) Original image; (b) Artifacts and background removed; (c) Pectoral muscle removed; (d) Enhanced image.

3.2.1. Preprocessing

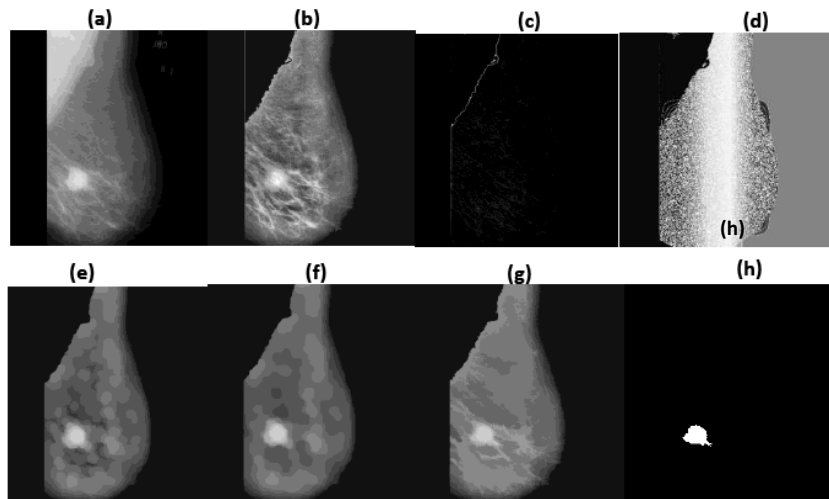
The preprocessing of mammogram is very essential because it consists of many artifacts, unwanted labels and pectoral muscle. Also it is poor in contrast; hence it is preprocessed the mammogram by removing background artifacts and pectoral muscle and enhancing the mammogram images [24]. Initially for removing the artifacts and background we converted gray image into binary with threshold value 0.1 and selected biggest blob which is the breast part without background and artifacts. Then pectoral muscle suppressed by using modified region growing technique. The proposed single seed point region growing algorithm worked with conventional method of region growing techniques with little modification of selection of seed point. The seed point is selected based on the orientation of the mammogram. In the

dataset all images are either in left oriented or in righted. After removing surpluses of the image, it is enhanced by using median filter and contrast limited adaptive histogram equalization technique is applied as shown in the Fig. 3.

3.2.2. ROI extraction

For segmentation of mass in mammogram we pass a preprocessed image as input to the proposed algorithm. Then the gradient of the image is calculated. For the proposed work we have identified “sobel” operator is most robust gradient operator that produces the smooth edges without suppressing the maximal and non-maximal elements.

After calculation of gradient, the proposed method finds the regional minima, on which watershed transformation is applied. The watershed lines are obtained by “ORing” with minimum values to get mask. Then the mask is imposed on the gradient image. But it results in over segmentation. Then the morphological operations such as opening and closing are applied to minimize the regions and fill the gap between the edges. Then the level wise thresholding is applied to select appropriate threshold point as shown in the Fig. 4.



(a) Original image from the mini-MIAS database; (b) preprocessed image; (c) Gradient image; (d) Watershed transform; (e) Opening; (f) Closing; (g) Reconstructed from opening-closing; (h) Mass identified.

Fig. 4. Experimental results of proposed method.

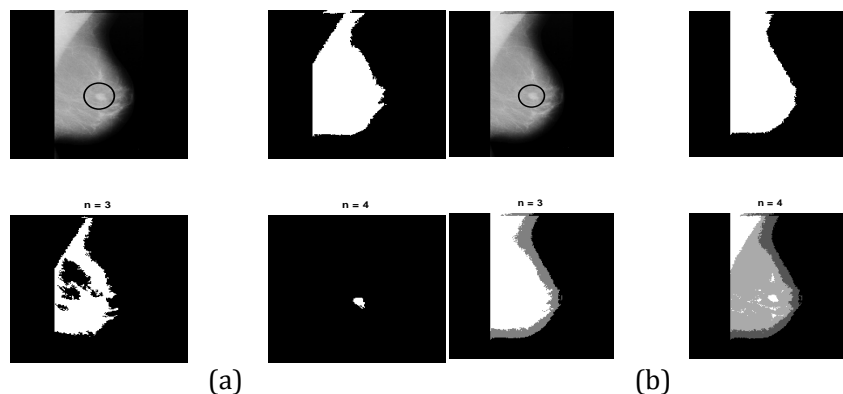


Fig. 5. Comparison of (a) proposed method and (b) Otsu thresholding method.

4. Results and Discussion

The results of proposed algorithm are showing promising results to the extracting ROI. The results are

compared with the Otsu thresholding method. The four levels of thresholding are applied to the mammogram and ROI successfully extracted at fourth level in the proposed technique as shown in the Fig 5 (a) and Otsu results in Fig. 5(b).

The proposed method tested on first 55 abnormal images of mini-MIAS database. The efficiency of the algorithm is measured by the following equations 6 and 7.

$$\text{Efficiency} = \frac{\text{Number of True Identification of mass}}{\text{Total abnormal images}} \quad (6)$$

$$\text{Error} = \frac{\text{Number of False Identification of mass}}{\text{Total abnormal images}} \quad (7)$$

The proposed method successfully identified 52 tumors among first 55 abnormal images. The 3 lesions could not be identified properly on the fourth level of threshold because of their poor intensity. They have found at higher level. Among 3 lesions 2 have missed and 1 wrongly identified. Hence, the efficiency of the algorithm is reported as 0.9454 and error rate is 0.0545. Comparing to results of Otsu thresholding our algorithm improved the results of detection of abnormality in the mammogram with simple concept.

5. Conclusion

Breast cancer is major issue facing by all the women in these days. Mammography is contemporary option for early detection and screening the breast cancer. Nevertheless, radiologist's has remarkable influence on analysis of mammography. It is difficult to identify the abnormality in mammography. But it is necessary to find the abnormality at an earliest to reduce a death rate. Hence, the new method helps to detect the tumor with efficient threshold-based segmentation approach. These results show satisfactory and the accuracy of the algorithm is 94.54% in identification of mass in mammography and false identification rate is 5.45%. Comparing with Otsu thresholding method with our finding proves to be good in identification of abnormality at fourth level. (Do not discuss limitation in research articles but in review you should write the limitation) This work can be a good technical approach for the betterment of the image quality in diagnostics and also helps to find out the relative features to classify the tumor as benign or malignant cancer.

References

- [1] Cancer Foundation of India. From <http://cancerfoundationofindia.org/>
- [2] Breast Cancer India. From <http://www.breastcancerindia.net/>
- [3] BharathaSreeja, G. P. R. (2012). Detection of Tumours in digital mammograms. *I. J. Modern Education and Computer Science*, 57-65.
- [4] Prasad, B. S. (2015). Detection of masses in mammogram based on non-linear filtering techniques. *Journal of Medical and Bioengineering*, 4(6), 430-435.
- [5] Nor Ashidi, M. I. T. S. (2012). Automatic Segmentation and detection of mass in digital mammograms. *Recent Researches in Communications, Signals and Information Technology*, 143-146.
- [6] VidhyaJyothi, R. G. E. (2014). FCM for malignant detection in mammogram. *Int. J. Comp. Tech. Appl*, 375-381.
- [7] Barnathan, M. (2012). Mammographic segmentation using wavecluster. *Algorithms*, 5, 318-329.
- [8] Manoj, M. R. (2012). A survey of segmentation in mass detection algorithm for mammography and thermography. *International Journal of Advanced Electrical and Electronics Engineering*, 70-77.
- [9] Lou, J.-Y., et al. (2015). A spatial shape constrained clustering method for mammography. *Hindawi Publishing Corporation, Computational and Mathematical Methods in Medicine*, 1-16.

- [10] Peter, A. I. K. O. (2015). Breast cancer risk prediction using data mining classification techniques. *Transactions on Networks and Communications*, 2-11.
- [11] Makandar, A. U. R., & Karibasappa, K. (2010). Wavelet based medical image compression using SPHIT. *Journal of Computer Science and Mathematical Science*, 1, 769-775.
- [12] AbhijitNayak, D. K. (2013). Suspicious lesion detection in mammograms using undecimated wavelet transform and adaptive thresholding. *IEEE Trans on Medical Imaging*, 1-6.
- [13] Gaikwad, V. J. (2015). Marker-controlled watershed transform in digital mammogram segmentation. *International Journal for Research in Applied Science & Engineering*, 18-21.
- [14] Salman, N. (2006). Image segmentation based on watershed and edge detection techniques. *The International Arab Journal of Information Technology*, 104-110.
- [15] Gonzalez, R. C. (2007) *Digital Image Processing* (2nd ed.).
- [16] Reddy, R. R. L. S. S. (2010). A novel image segmentation technique for detection of breast cancer. *International Journal of Information Technology and Knowledge Management*, 201-204.
- [17] Aswini, K. M. A. P. (2011). Detection of masses from mammograms using mass shape pattern. *Int. J. Comp. Tech. Appl*, 1131-1139
- [18] Ashmithakhaleel, K. N. (2014). Wavelet based automatic lesion detection using improved OTSU method. *International Journal of Computer Science & Network Solutions*, 119-127.
- [19] Nasseer, M. B. M. H. (2013, March). Segmentation of breast masses in digital mammograms using adaptive median filtering and texture analysis. *International Journal of Recent Technology and Engineering*, 2(1), 39-43.
- [20] Amel, H. A. A. A. (2015). Breast cancer image segmentation using morphological operations. *International Journal of Electronics and Communication Engineering & Technology*, 8-14.
- [21] Guzman-Cabrera, J. R.-S. R. (2013). Digital image processing technique for breast. *Int J. Thermophys*, 1519-1531.
- [22] Samir, K. B. I. K. (2015). Fully automated computer aided diagnosis (cad) system of human breast cancer using digital mammogram. *Asian Journal of Computer Science and Information Technology*, 11-20.
- [23] Suckling *et al.* (1994). The mammographic image analysis society digital mammogram database, exerpta medica. *International Congress Series 1069*, 375-378.
- [24] Makandar, A., & Halalli, B. (2015). Breast cancer image enhancement using median filter and CLAHE. *International Journal of Scientific & Engineering Research*, 6(4), 462-465.



Aziz Makandar is serving as Professor in the Department of Computer Science at the Karnataka State Women's University, Vijayapura, since 2007. So far, he has 16 years' teaching and research experience in the computer science field. He has published 19 research articles in reputed international journals and presented 21 research papers in the national and international conferences. His area of interest is in digital image processing, artificial intelligence and fuzzy.



Bhagirathi Halalli received the M.Sc. master degree in computer science with first rank from Karnataka State Women's University Vijayapura, India in 2009 and cleared Karnataka State Eligibility Test (KSET) for lectureship/Assistant Professor in 2014. Currently she is perusing the Ph.D. in the Department of Computer Science at Karnataka State Women's University, Vijayapura, India. Her area of interest is in digital image processing.