Analysis of Tumor Characteristics Based on MCA Decomposition and Watershed Segmentation

Narain Ponraj.D
Asst.Prof/ECE,
Karunya University,
Coimbatore, 641114
India

Evangelin Jenifer.M
PG Student/ECE,
Karunya University,
Coimbatore, 641114
India

Dr.P.Poongodi
Professor/ECE,
PPG Institute of Technology,
Coimbatore, 641035
India

Samuel Manoharan.J
Asst.Prof/ECE,
Karunya University,
Coimbatore, 641114
India.

ABSTRACT

An accurate and standardized technique for breast tumor segmentation is a critical step for monitoring and quantifying breast cancer. The fully automated tumor segmentation in mammograms presents many challenges related to characteristics of an image. In this paper, two different methods for mass detection are applied. First method uses morphological component analysis and multiple layer thresholding. Second method uses watershed segmentation. Features are extracted and the best one is found out for efficient identification of breast cancer.

Keywords

Breast Cancer, Morphological Component Analysis, Undecimated Wavelet Transform, Watershed Segmentation.

1. INTRODUCTION

One of the most frequent leading causes of cancer deaths in the world is the breast cancer. Most of the women are suffering from breast cancer. Early prediction of breast cancer is the only primary solution for preventing breast cancer.

X-ray has been used as the diagnostic and therapeutic tool in medical field. Mammography is a low dose x-ray used for visualizing the internal structure of the breast. Mammography has been proven to be the most reliable method and it is the key screening tool for the early detection of breast cancer[1].Mammography diagnosis of breast cancer in its early stage. Breast cancer mortality rates have been reduced due to the mammographic screening procedure. The screening programs have reduced the mortality rates by 30-70%.

Sometimes mammogram is difficult to due to its sensitivity. It may also be affected by the image quality and radiologist's experiences. More than one radiologist examines the mammogram image in order to predict whether the cancer is present or not. Therefore Computer- aided detection (CAD) schemes have been developed to assist the radiologists in improving the accuracy of diagnosis [2]. The CAD methods are low cost tool for the interpretation of mammographic films. It can be used to assist radiologists for the detection of suspicious lesions and provides a second reading which has been found to reduce the misdiagnosis.CAD systems requires a well-trained computer to become practical in clinical settings. The results of the CAD systems are compared with the expert radiologist values so that the location of cancerous portion may be accurate. This may reduce unnecessary biopsies.

There are two signs of breast cancer. They are microcalcification and mass. Microcalcifications are the calcium deposits which can be visible in the mammogram Image. Masses are difficult to be interpreted. There are two types of masses. They are the Benign mass and the Malignant mass. A mass can be benign if it is non cancerous or malignant if it is cancerous. A benign mass do not spread to other parts of the body but still may need to be removed because the local tissue may be damaged. On the other hand, a malignant mass can destroy neighboring tissues and spread to other parts of organ or body [3]. Masses are more difficult to detect than micro calcifications because the features of a mass bear semblance to those of the normal breast parenchyma. A mass with regular shape has a higher probability of being benign where as a mass with an irregular shape has a high probability of being malignant [4].

2. IMAGE DATASET

The most easily accessed and the most commonly used databases are the Mammographic Image Analysis Society (MIAS) database and the Digital Database for Screening Mammography (DDSM).When benchmarking an algorithm it is recommendable to use a standard test database (data set) for researchers to be able to directly compare the results. Most of the mammographic databases are not publicly available. In this paper the most easily accessed databases which belong to three types such as Normal, benign and malignant are taken from the Mammographic Image Analysis Society (MIAS) database.

3. MCA BASED METHODOLOGY

Mass detection in mammogram plays an important task for the early diagnosis of breast cancer. However, masses appear same as the normal regions in terms of their abundant morphological characteristics and their unclear border [5]. So it is difficult to interpret the difference. Preprocessing stage is very important to improve the mass detection performance, It is necessary because it preserves both the intensity distribution and morphological characteristics of regions. To maintain the local contrast of an original mammogram, to preserve the margin characteristics and to suppress the structural noises, blood vessels, and glandular tissues Morphological component analysis(MCA) is introduced as a preprocessing step in this scheme.MCA technique uses undecimated wavelet transform to form the piecewise component. Features such as area, eccentricity, solidity etc are extracted based on the concentric multiple layers.

3.1 Undecimated Wavelet Transform

Undecimated wavelet transform is also called as stationary wavelet transform. It overcomes the disadvantages of discrete

wavelet transform. It forms the basis of biorthogonal wavelet transform.OWT leads to a successful implementation in image compression and the results finds its application best in areas such as filtering, deconvolution, detection or generally data analysis [6].

Decimation of the wavelet coefficients is an important property of the discrete wavelet transform (DWT).Every other of the coefficients of the current level are removed by the decimation step. Thus the computation of the wavelet transform is faster and more compact in terms of storage space. More importantly, original image can be reconstructed from the remaining coefficients. Discrete wavelet transform has the property of shift variance which is due to decimation of wavelet coefficients. Shift invariance is important for the perfect reconstruction and to preserve the image information. In order to achieve shift- invariance, researches from different invented several wavelet transform algorithms. This transforms is known under the common name undecimated wavelet transform (UWT).Fig.1 shows the MCA decomposition scheme.
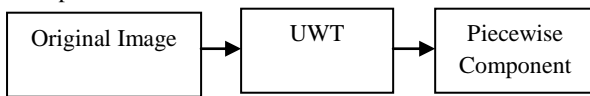

Fig. 1: MCA Decomposition Scheme

In addition to the shift-invariance the undecimated wavelet transform gives increased amount of information about the transformed signal compared to the DWT. The number of the wavelet coefficients does not shrink between the transform levels. For the better analysis and understanding of the signal properties additional information obtained from the coefficients can be very useful. For instance, in the image denoising applications the discrimination between the noise and the real data can be improved [7].

For analysing the wavelet coefficients larger amount of information is especially important. UWT preserve most of the information while reconstructing. The main drawback of the UWT is bigger computational and memory cost as well as redundancy in the coefficients. Undecimated version of biorthogonal wavelet transforms can effectively represent piecewise-smooth component which shows the intensity distributions in mammograms and the shift invariance property of UWT could make the mass location process more accurate.

3.2 Steps Involved in Mass detection Algorithm

The following steps are involved in morphology component analysis and concentric layer model,
Step 1: Otsu's method which performs histogram shape-based image thresholding is used to segment the breast region.
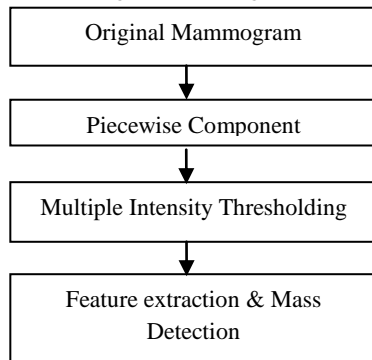

Fig. 2: Block Diagram for Mass Detection Process

Then, undecimated wavelet transform is used to decompose the mammogram image into its piecewise-smooth component.

Step 2: Using multiple intensity thresholds, piecewise-smooth parts of mammograms are separated into different intensity layers. The thresholds start from the highest gray level to the lowest one with a predefined fixed step. When the area of the thresholding regions is larger than half of the size of the mammograms the intensity layer formation is stopped.

Step 3: There are different morphological features in calculating the suspicious focal regions. Particularly, area of the tumor portion plays an important role. Segmented image which shows the detected mass portion is identified.Fig.2 shows the block diagram for the mass detection process.

4. METHODOLOGY FOR WATERSHED SEGMENTATION

For the purpose of contrast enhancement Morphology is used. Morphology in image processing is based on shapes. Structuring element is used to process the image [8]. The size and shape of the structuring element plays an important role in adding or removing the pixels from the objects. In the structuring element, the centre pixel corresponds to the origin of the pixel and the ones correspond to the neighborhood. For each input pixel the structuring element is superimposed on top of the input image so that the origin of the structuring element coincides with the input pixel coordinates [9]. As a result corresponding morphological operations are carried out. The general block diagram for mass detection by watershed segmentation is shown in Fig.3.
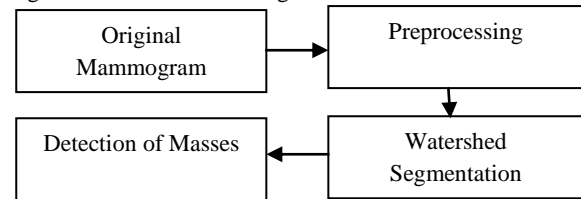

Fig. 3: Block Diagram for Watershed Mass Detection Process

4.1 Morphological Operations

Morphology has various operations such as dilation, erosion, opening, closing, top-hat transform, bottom-hat transform etc. Dilation and erosion are basic operations for all other morphological operations [10].

In dilation the value of the pixel of interest is the minimum value of all the pixels in the input pixel's neighborhood. In erosion, the value of the pixel of interest is the maximum value of all the pixels in the input pixel's neighborhood. Morphological opening of an image is erosion followed by dilation. Morphological closing of an image is the reverse consists of dilation followed by erosion with the same structuring element [11].

The top-hat transform is defined as the difference between the original image and its opening. The bottom-hat transform is said to be the difference between the closing of the original image and the original image. Top-hat transform depends on the collection of the foreground parts of an image. Bottom-hat transform is the collection of the background parts of an image that fit a particular structuring element.

Initially the original image and top-hat transform are added and then the bottom hat image is subtracted from the obtained image to get the contrast high image. As a result, the contrast of the mammogram image is increased which can be further used for segmenting the masses [12].

Enhancement is aimed at realizing improvement in the quality of a given image. It can be accomplished by removing noise, enhancing contrast, and enhancing edges. Applying contrast enhancement filters improve the readability of areas with subtle changes in contrast. However, they will also destroy areas of the image where the intensity of the pixels is outside the range of intensities being enhanced [13].

4.2 Watershed Segmentation

Watershed Segmentation is based on immersion algorithm proposed by Vincent soille. Watershed is an extent in which the rain or ice or melting snow converges in a single point in an area of land. Watershed segmentation gets its name from the approach in which the algorithm segments regions into watershed line and volume of catchment basins[14]. Image is considered as a landscape with holes. Landscape with holes is immersed in water and water rises through the holes fills the basins. When a particular point is reached, dam is built. Different basins are said to be catchment basins. The point where different basins meet are said to be Watershed.

Grayscale images include various altitudes of pixels with values from 0 to 255. The output of watershed segmentation is a binary image [16-17], black pixels are assigned to dams, or watersheds, and white pixels are assigned to regions surrounded by dams i.e. for the catchment basins.

Watershed segmentation consists of two approaches, in first approach basins are identified, and then watersheds are obtained by taking a set complement or a complete partition of the image into basins is taken and subsequently the watersheds are found by boundary detection. Second method is to use labelling method for watershed segmentation. Catchment basins have same label value and a special label is assigned for the watershed lines [18].

The watershed transform has been extensively used in many areas of image processing, including segmentation in medical fields. Some of the advantages it possesses are it is a uncomplicated method, innate method, it is prompt and can be parallelized and it highlights the suspicious region and it produces a complete division of the image [19].

4.3 Steps Involved in Watershed Segmentation

The Watershed algorithm consists of following steps:

1. Mammogram images are read using the Matlab function and stored in the form of matrix array.

2. Minimum and maximum pixel values of the input grayscale image are found in order to sort pixel values.

3. Output image with the same matrix as that of the original image is initialized, each pixel in the output image is assigned to the constant INIT (-1).

4. Sort the pixels based on maximum and minimum altitude.

5. Four and eight connectivity are generally used in finding the neighborhood .Based on the neighborhood new discrete label is assigned starting with the minimum altitude. A label is any integer greater than 0.

6. Pixel in the next higher altitude is examined after analyzing the current pixel of interest.

    a. Initially all the neighbors are labeled as -1,if this condition exists ,then new label assigned to the pixel of interest (any integer > 0).

    b. If the neighbors are labeled with the same label, then label of its neighbors is assigned to the basin pixel.

    C. Watershed point is given as a label, if all its neighbors are labeled and they are not similar neighbors have different labels.

7. Above steps are repeated until the maximum altitude is reached, all pixels in the maximum altitude are labels as watershed points.

8. All the watershed points are assigned the value 0 (black) and all non-watershed points as the value 1 (white) to produce the binary image.

9. Cancerous portion is identified using the final binary output image.Fig.4 shows the flowchart for watershed segmentation.
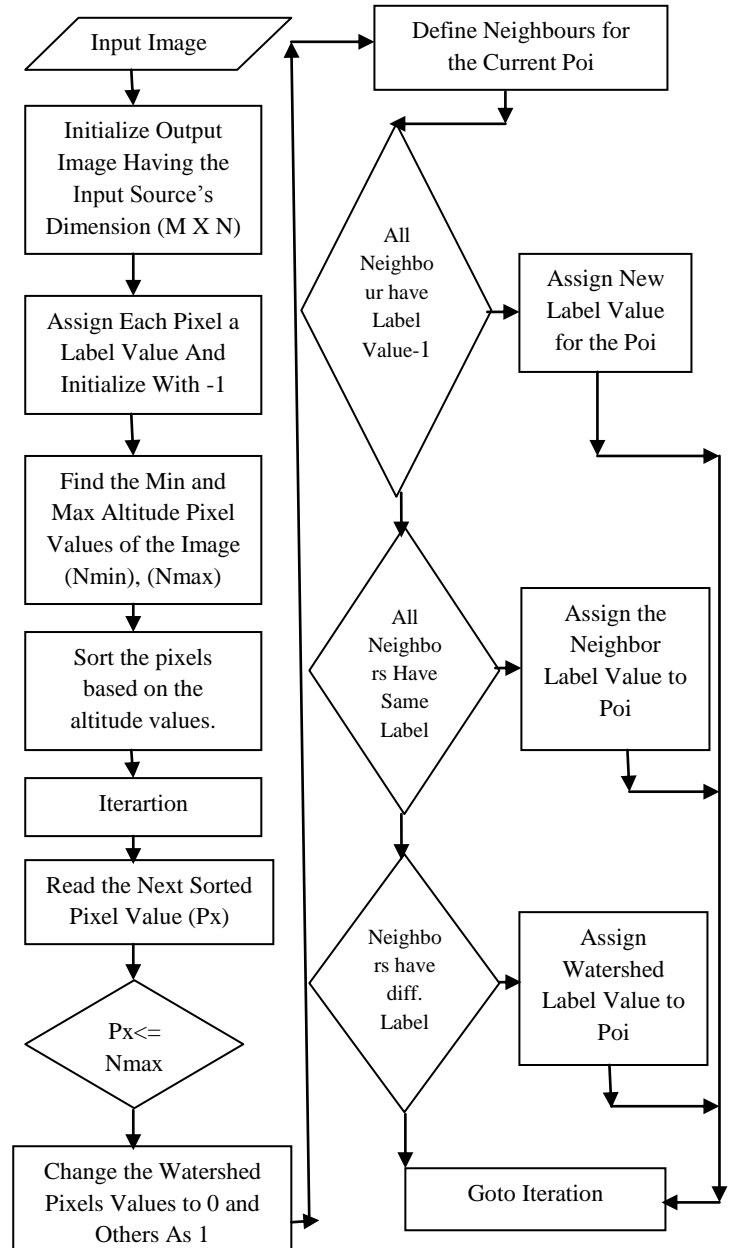

Fig. 4: Flowchart for Watershed Segmentation

5. RESULTS AND DISCUSSION

Mammogram images with masses are collected from the MIAS database. Cancerous mammograms images were taken for processing and the results are discussed below. Original image with 1024 by 1024 matrix is taken into consideration.Fig.5 shows the original images used for MCA.
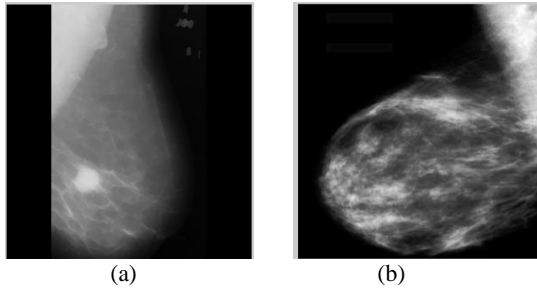
(a)   (b)

Fig. 5: Original Images used for MCA

After Morpholgical component analysis, piecewise smooth component is obtained.This component is a result after preprocessing of mammogram image.Fig.6 Shows the piecewise component.
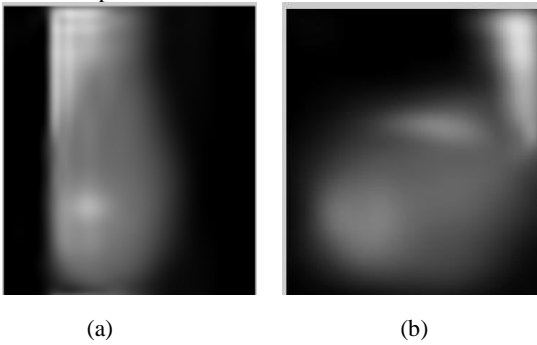

(a)   (b)

Fig. 6: Piecewise Smooth Component

The MCA scheme has been tested in input images shown in Fig.5 and the portion which is affected by cancer is detected.Fig.7 shows the detected cancerous portion of two different images.
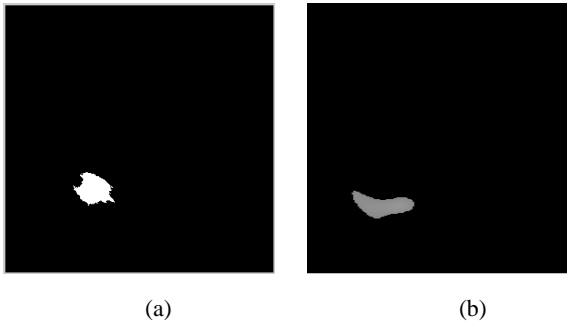

(a)   (b)

Fig. 7: Detected Cancerous Portion

Two different mammogram images with 1024 by 1024 matrix are taken into consideration for analysing watershed segmentation and they are shown in Fig.8.
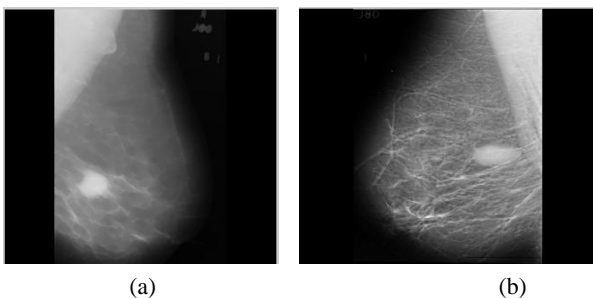

(a)   (b)

Fig. 8: Original images used for Watershed segmentation

Original images are preprocessed using morphological operations such as top hat transform, bottom hat transform etc. And the resultant enhanced images are shown in Fig.9.
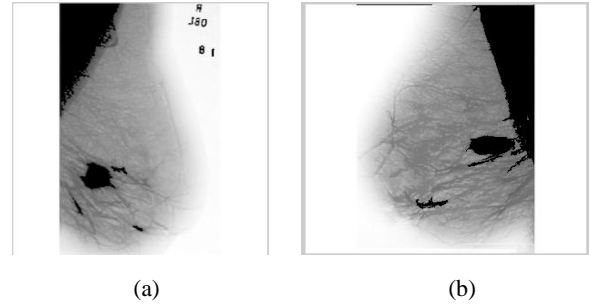

(a)   (b)

Fig. 9: Preprocessed Images

Based on immersion algorithm watershed Image is obtained which results in efficient segmentation of region of interest. Fig.10 shows the watershed images.
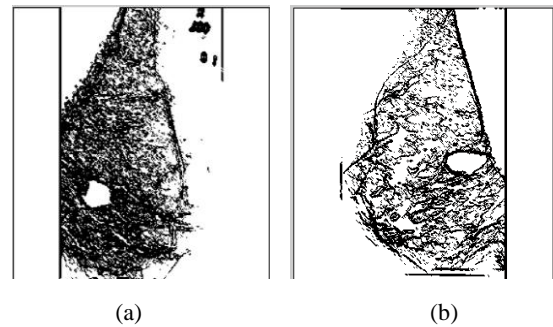

(a)   (b)

Fig. 10: Watershed images

Watershed image is further processed to extract the region with mass.Fig.11 shows the resultant segmented images.

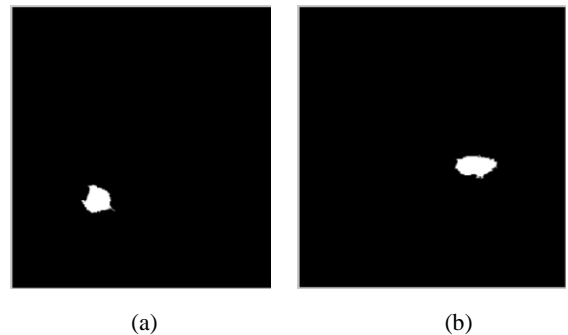(a)   (b)

Fig. 11: Segmented images

Comparison of Resultant segmented images from watershed segmentation and the mass detection process using MCA decomposition is shown below in Fig 12.
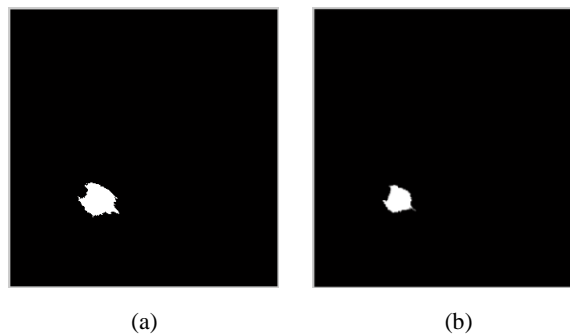

(a)   (b)

Fig. 12: (a) Segmented Image after MCA Decomposition and Thresholding (b) Segmented Image after Watershed Segmentation

Tumor characteristics after the mass detection scheme and watershed segmentation are given in Table 1 and Table 2.

Table.1.Tumor Characteristics after Mass Detection Process Using MCA

Solidity	0.7234
Eccentricity	0.3985
Major Axis Length	28.959
Minor Axis Length	23.234

Table.2.Tumor Characteristics after Watershed Segmentation

Solidity	0.8134
Eccentricity	0.4061
Major Axis Length	32.3167
Minor Axis Length	29.5324

Tumor area of the mammogram is found out. By using the tumor area relative error is obtained. Relative error is the percentage of relative error specifies the accuracy of segmentation. Relative error is a measure of the uncertainty of measurement compared to the size of the measurement. Relative error is also Known as relative uncertainty. Tumor Area and relative error comparison are given in Table 3.

Table.3 Comparison of Tumor Area with an Expert Radiologist Results And the Relative Error

Features	Expert Radiologist	Watershed Segmentation	Mass Detection process using MCA
Tumor area	710	706	850
Relative Error		0.563	17.63

Tumor area obtained using watershed segmentation is almost similar to the expert radiologist tumor area. It is seen that the relative error of watershed segmentation shows a good result compared to the mass detection process using MCA.

6. CONCLUSIONS

Two different mass detection schemes were implemented. First a novel mass detection scheme based on MCA and new concentric layer criteria is implemented. The scheme first decomposes the mammograms into piecewise-smooth component by using MCA. Undecimated wavelet transform is used for decomposition. Based on thresholding, features are extracted. Second technique is the detection of abnormal mass of the mammograms using the Watershed segmentation. Here images are preprocessed using morphological operations. Based on the features extracted both the techniques are compared. Relative error is used to compare them both and it is proved that performing watershed segmentation yields an accurate result compared to its counter-part.

REFERENCES

[1] American Cancer Society. Breast cancer facts and figures 2007–2008.

[2] R.M. Rangayyan, F.J Ayres,et.al, july 2002. "A review of computer-aided diagnosis of breast cancer : Towards the detection of subtle signs", J.Franklin Inst..

[3] Prof. Samir Kumar Bandyopadhyay, "Survey on Segmentation Methods for Locating Masses in a Mammogram Image", International Journal of Computer Applications,November 2010, Volume 9, No.11.

[4] R.M. Rangayyan, L. shen, et.al, "Improvement of sensitivity of breast cancer diagnosis with adaptive neighbourhood contrast enhancement of mammograms", IEEE Trans. Med. Img, Sep 1997.

[5] D.Narain Ponraj, M.Evangelin Jenifer, Dr.P.Poongodi, J.Samuel Manoharan, "A Survey on the Preprocessing Techniques of Mammogram for the Detection of Breast Cancer", International Journal of Computer Science and Security (IJCSS), Volume (1), Issue (3), 2011.

[6]   Xinbo Gao, Ying Wang, Xuelong Li, "On Combining Morphological Component Analysis and Concentric Morphology Model for Mammographic Mass Detection", IEEE transactions on information technology in biomedicine, march 2010, vol. 14, no. 2.

[7] S.K.Bandyopadhyay, "pre-processing of Mammogram Images", International Journal of Engineering Science and Technology, 2010, Vol. 2(11), pp 6753-6758.

[8] D. Narain ponraj, Sweety Kunjachan, Dr.P.Poongodi, Samuel manoharan, "A Survey on Texture analysis of mammogram for the detection of breast cancer", CIIT International journal, 2011, vol.3.

[9] J. G. Schavemaker, M. J. Reinders, J. J. Gerbrands, and E. Backer, "Image sharpening by morphological filtering", Pattern Recognition, vol. 33, 2000,pp. 997–1012.

[10] Serra, J. 1982.Image Analysis and Mathematical Morphology, Academic Press, New York.

[11] Gonzalez RC, Woods R.E, 2003 Digital Image processing, 2nd ed.

[12] Shavi Gupta, Mohd. Sadiq, Mona Gupta and Naseem Rao.2011."Semi Automatic Segmentation of Breast Cancer for Mammograms Based on Watershed Segmentation", Proceedings of the 5th National Conference; INDIACom-2011.

[13] A.E.Hassanien and E.H. Tarek Abed,2003. "Digital Mammography Image Analysis System Based on Mathematical Morphology", in IEEE computer society 7th International Conference On Intelligent Engineering Systems INES ,Egypt, pg.141-147.

[14] 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 34,2010, pg.185–191.

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

[16] Vincent L, Soille P, "Watersheds in digital spaces: an efficient algorithm, based on immersion simulations", IEEE Trans Pattern Anal Mach Intell, 13(6), 1991, pg.583–598.

[17] Jaya Sharma & Sujeet Sharma, "Mammogram image segmentation using watershed", International Journal of Information Technology and Knowledge Management, July-December 2011, Volume 4, No. 2, pg. 423-425.

[18] V. Grau, A. U. J. Mewes, M. Alcañiz, R. Kikinis, and S. K. Warfield, "Improved Watershed Transform for Medical Image Segmentation Using Prior Information", IEEE transactions on medical imaging,2004,vol. 23, no.

[19] J. Freixenet, X. Munoz, D. Raba, J. Marti, and X. Cufi ,2002. "Yet another survey on image segmentation, Region and boundary information integration", Computer Vision, pg. 408-422.