AUTOMATED DIGITAL MAMMOGRAM SEGMENTATION FOR DETECTION OF ABNORMAL MASSES USING BINARY HOMOGENEITY ENHANCEMENT ALGORITHM

Indra Kanta Maitra, Sanjay Nag and Prof. Samir K. Bandyopadhyay

Dept. of Computer Science & Engineering, University of Calcutta, Kolkata, India

Prof. Samir K. Bandyopadhyay, Professor, Dept. of Computer Science & Engineering, University of Calcutta, 92 A.P.C. Road, Kolkata – 700009, West Bengal, India, ikm.1975@yahoo.com

Abstract- Many image processing techniques have been developed over the past two decades to help radiologists in diagnosing breast cancer. At the same time, many studies proven that an early diagnosis of breast cancer can increase the survival rate, thus making screening programmes a mandatory step for females. Radiologists have to examine a large number of images. Digital Mammogram has emerged as the most popular screening technique for early detection of Breast Cancer and other abnormalities. Raw digital mammograms are medical images that are difficult to interpret so we need to develop Computer Aided Diagnosis (CAD) systems that will improve detection of abnormalities in mammogram images. Extraction of the breast region by delineation of the breast contour and pectoral muscle allows the search for abnormalities to be limited to the region of the breast without undue influence from the background of the mammogram. We need to perform essential pre-processing steps to suppress artifacts, enhance the breast region and then extract breast region by the process of segmentation.

In this paper we present a fully automated scheme for detection of abnormal masses by anatomical segmentation of Breast Region of Interest (ROI). We are using medio-lateral oblique (MLO) view of mammograms. We have proposed a new homogeneity enhancement process namely Binary Homogeneity Enhancement Algorithm (BHEA), followed by an innovative approach for edge detection (EDA). Then obtain the breast boundary by using our proposed Breast Boundary Detection Algorithm (BBDA). After we use our proposed Pectoral Muscle Detection Algorithm (PMDA) to suppress the pectoral muscle thus obtaining the breast ROI, we use our proposed Anatomical Segmentation of Breast ROI (ASB) algorithm to differentiate various regions within the breast. After segregating the different breast regions we use our proposed Seeded Region Growing Algorithm (SRGA) to isolate normal and abnormal regions in the breast tissue. If any abnormalities are present it gets accurately highlighted by this algorithm thus helping the radiologists to further investigate these regions. This composite method have been implemented and applied on all mammograms with abnormalities in mini-MIAS database. The algorithms proposed are fully autonomous, and are able to isolate different types of abnormalities, if present, a task very few existing mammogram segmentation algorithms can claim.

Keywords: Mammogram, Gaussian Smoothening, Binary Homogeneity Enhancement Algorithm (BHEA), Edge Detection Algorithm (EDA), Pectoral Muscle Detection Algorithm (PMDA), Breast Boundary Detection Algorithm (BBDA), Anatomical Segmentation of Breast ROI (ASB),Seeded Region Growing Algorithm (SRGA).

I. INTRODUCTION

Breast cancer is the most common type of cancer in women in both developed and developing countries, accounting for 1 of every 3 cancers diagnosed. The incidence of breast cancer in India is on the rise and is rapidly becoming the number one cancer in females pushing the cervical cancer to the second spot. It is reported that one in 22 women in India is likely to suffer from breast cancer during her lifetime, while the figure is definitely more in America with one in eight being a victim of this deadly cancer [1].

In India, the death toll due to the breast cancer is increasing at a rapid pace [2]. Since the cause of breast cancer remains unknown, primary prevention becomes impossible. Only the early detection and diagnosis is the way of control but it is a major challenge in India due to lack of awareness and lethargy of the Indian women

towards the health care and regular checkup. But the major obstacle is the expensive health care system, especially in breast cancer treatment.

Mammography screening associated with clinical breast examination is the only viable and effective method at present for mass screening to detect breast cancer. A digital mammogram is created when a conventional mammogram is digitized, so, it can be used by a computer, without the loss of information from the original mammogram. Digitization can be performed through the use of a specific mammogram digitizer [3] or a camera [4], [5].

It is estimated that current breast screening programs fail to detect approximately 25% of the cancers that are visible on retrospective review [6]. The screening programs involve problems such as high costs, inexperienced radiologists, and visual fatigue. Detection methods are based on clinical examination, mammography, ultrasound, and core biopsy. An early detection gives the patient a good chance of survival, whereas a late detection can be fatal and often ends in the death of the patient. A breast cancer once detected can be classified as benign or malignant.

Digital mammography refers to the application of digital system techniques on digital mammograms. Digital systems have the capacity to bring revolutionary advantages to breast cancer detection [7]. Radiologists turn to digital mammography for an alternative diagnostic method due to the problems created by conventional screening programs. An automated system can overcome these problems by reducing the number of false positive and false negative readings from radiologists and increase the chance of detecting abnormalities early. Up to this moment mass detection from digitized mammograms is still very challenging. The main reason is that the abnormal masses usually mix with the homogeneous tissues in the breast. The images provided by different patients have different dynamics of intensity and present a weak contrast. Moreover the size of the significant details can be very small. Several research works have tried to develop computer aided diagnosis (CAD) tools that could help the radiologists in the interpretation of the mammograms and could be useful for an accurate diagnosis [8] [9] [10].

Detection of abnormal masses within the breast as well as breast image segmentation is a very important feature in image analysis. Detection of abnormal masses within the breast is an important factor which can improve clinical diagnosis of mammographic diseases. The accurate segmentation of the abnormal masses in mammogram is an essential step in the computerized analysis of mammograms. It allows the search for abnormalities to be limited to the anatomical region of the breast without undue influence from the background and other regions of the mammogram. It also facilitates enhancements for techniques such as comparative analysis, which includes the automated comparison of corresponding mammograms. Hence in this research work, we have proposed novel step-by-step algorithms for mammogram segmentation to detect the abnormal masses. Before stepping into the main part of our algorithm, a district set of pre-processing techniques are used for better results. The proposed methods have presented significant results and shown that the proposed methods are more accurate and reliable compared with other common methods.

II. LITERATURE REVIEW

Being an interdisciplinary area of research, the mammograms and their analysis by the way of image processing, encompass many disciplines such as statistics, mathematics, computer and medical professionals. A large number of studies have been conducted in an attempt to improve the performance of CAD. One approach focuses on improving segmentation accuracy and identifying effective image features. Previous studies have suggested that improving the accuracy of mass and non-mass region segmentation could also significantly improve the performance of CAD in abnormal masses detection and characterization.

Several automated and semi-automated methods focusing on this problem have been developed. These include using a density-weighted contrast enhancement algorithm that combines adaptive filtering and edge detection (11), an adaptive multilayer topographic regional growth algorithm (12), a gray-level–based iterative and linear segmentation algorithm (13), a dynamic programming approach (14), dynamic contour modeling (15) etc. to segment mass lesions from surrounding breast tissue. One important feature in automated mass detection and classification is mass boundary spiculation level (16). A number of previous studies have been conducted in an attempt to automatically detect spiculated masses and incorporate this feature into CAD methods.

Kegelmeyer et al (17) detected spiculated masses using local edge orientation and Laws texture features. In their algorithm, a statistical classifier is applied to labeling each pixel with its probability of being located on an abnormality, but it is not applicable for the detection of nonspiculated masses. Karssemeijer and Te Brake (18) used a method based on the statistical analysis of a map of pixel orientations to detect stellate distortions on mammograms. Comer et al.[19] and Li et al. [20] used Markov random fields to classify the different regions in a mammogram based on texture.

Huo et al (21) developed a new spiculation-sensitive pattern recognition technique to quantify the degree of speculation of a lesion and classified masses as malignant or benign. They reported that the computer-extracted speculation measure was shown to yield higher classification accuracy than the speculation rating of an experienced radiologist. Nakayama et al. [22] used a filter bank for the detection of nodular and linear patterns.

The filter bank is designed so that the sub-images generated the elements of a Hessian matrix at each resolution level. By calculating the small and large eigenvalues, a new filter bank has the following three properties. (I) Nodular patterns of various sizes can be enhanced. (II) Both nodular and linear patterns of various sizes can be enhanced. (III) The original image can be reconstructed with these patterns removed. The filter bank is applied in mammograms.
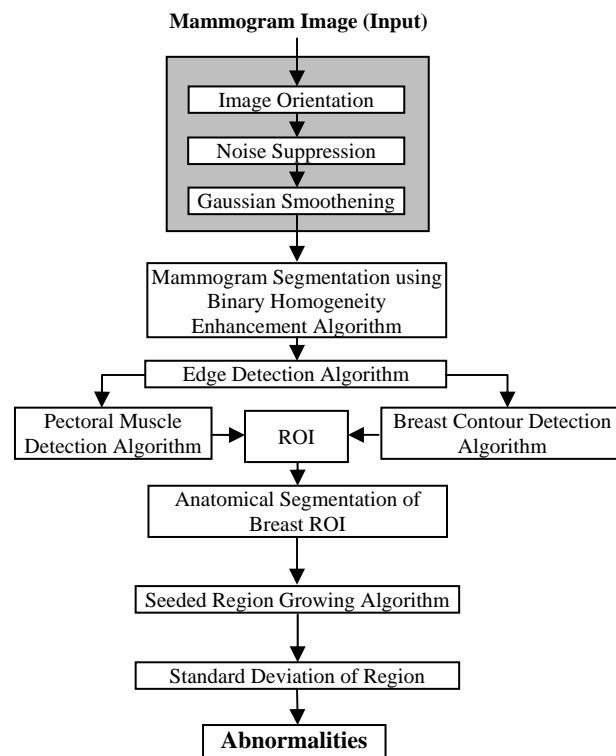
A lesion segmentation algorithm was developed by Sameti et al. [23] used fuzzy sets to partition the mammographic image data. Despite the difficulty and importance of this step in many computerized mass-detection schemes, few have attempted to analyze the performance of these segmentation algorithms alone, choosing instead to collectively analyze all components of a scheme.

Sahiner et al (24) presented a fully automated and three-stage segmentation method that included clustering, active contour, and spiculation detection stages to detect spiculated and nonspiculated masses. The results indicated that combining texture features with morphologic features extracted from automatically segmented mass regions was an effective approach for the automated characterization of mammographic masses. Recent studies of interactive CAD methods have also suggested that the accurate detection and classification of mass boundary spiculation levels plays an important role in improving the visual similarity of similar reference mass regions selected by interactive CAD methods (25).

There are significant research effort have gone into developing computerized methods to detect abnormal masses of various types and classify them. In our study, we developed and tested a new fully automated computerized method whose aim is to obtain the abnormal masses within breast ROI and delimiting them with boundaries using a hybrid of different techniques and that is comparable with research paper by Sahiner et al.

III.   PROPOSED METHODS

Digital Mammograms are medical images requires a preparation phase in order to improve the image. Our objective during this process is to improve the quality of the image to make it ready for further processing by removing the irrelevant and unwanted parts in the background of the mammogram.

Mammogram Image (Input)

Image Orientation

Noise Suppression

Gaussian Smoothening

Mammogram Segmentation using Binary Homogeneity Enhancement Algorithm

Edge Detection Algorithm

Pectoral Muscle Detection Algorithm

ROI

Breast Contour Detection Algorithm

Anatomical Segmentation of Breast ROI

Seeded Region Growing Algorithm

Standard Deviation of Region

Abnormalities

A. Mammogram Preprocessing

1) Image Orientation

The mammogram image is transformed, so that the chest wall location, i.e., the side of the image containing the pectoral muscle, is on the upper left corner of the image. In order to determine the chest wall location, the decreasing pixel intensity of the breast tissue near the skin-air interface is used. This tissue is estimated by employing the Minimum cross-entropy threshold selection technique [26] twice in the original image. After determining the chest wall location we determine the top of the image. We extract the vertical centroid of the image and then we assume that the asymmetric region closest to the right side of the vertical centroid is the tip of

the breast. The image is flipped horizontally if needed to place the asymmetric region below the vertical centroid, resulting in an image that is right way up.

2) Noise Suppression

There are different types of noises, which appear in miniMIAS images. High intensity noise is characterized by high values of optical densities, such as labels or scanning artifacts. Tape artifacts are markings left by tapes, or other shadows presenting themselves as horizontal running strips. Such noise must be replaced by black pixels.

3) Gaussian Smoothening

The Gaussian smoothing operator is a 2-D convolution operator that is used to `blur' images and remove detail and noise. It uses a kernel that represents the shape of a Gaussian (`bell-shaped') hump. In 2-D, an isotropic (i.e. circularly symmetric) Gaussian has the form:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

The idea of Gaussian smoothing is to use this 2-D distribution as a `point-spread' function, and this is achieved by convolution. Since the image is stored as a collection of discrete pixels we need to produce a discrete approximation to the Gaussian function before we can perform the convolution. In theory, the Gaussian distribution is non-zero everywhere, which would require an infinitely large convolution kernel, but in practice it is effectively zero more than about three standard deviations from the mean, and so we can truncate the kernel at this point. Once we obtain a suitable kernel, then the Gaussian smoothing can be performed using standard convolution methods.

In our proposed method we have taken 7×7 kernel as a convolution filter. The mammogram images in the MIAS database are categorized broadly into 3 categories namely Fatty, Fatty-Glandular and Dense-Glandular depending on the density of fatty tissues and abundance of glands in the breast. Close observation of all the mammograms reveal that each category of mammogram displayed a varied intensity value which is distinct from each category. This property of the mammograms has helped us in our choice of the value of deviation (Ω) for each category, thus able to adjust the level of smoothening for each category.

Step 1. Calculate Gaussian Kernel values to create a convolution filter.
Step 2. Convolution of the mammogram image starting from Left Topmost pixel to Bottom Rightmost pixel.
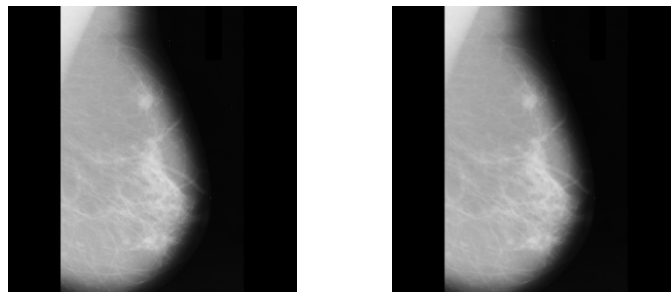Step 3. Perform matrix multiplication between calculated Gaussian filter values and pixel intensity.

Figure 1.   Mammogram before and after Gaussian Smoothening (MIAS 023.R)

B. Anatomical Segmentation of Digital Mammogram and Detection of Abnormalities

This paper is based on the image segmentation method, the inputs are images and, outputs are the attributes extracted from those images. Segmentation divides image into its constituent regions or objects. Segmentation plays an important role in image analysis. The goal of segmentation is to isolate the regions of interest (ROI) depending on the problem and its characters. The approach is to partition an image based on abrupt changes in intensity, such as edges in an image and partitioning image into regions that are similar according to a set of predefined criteria.

A mammogram contains two distinctive regions, the exposed breast region and the unexposed air-background (non-breast) region. The principal feature on a mammogram is the breast contour, otherwise known as the skin-air interface, or breast boundary and the pectoral muscle. The breast region can be obtained by partitioning the mammogram into breast, non –breast region and pectoral muscle. On isolation of both non-breast and pectoral muscle region we can obtain the breast region for further studies. After obtaining the breast ROI, we need to differentiate and partition the anatomical regions within the breast ROI. During this process, we identify each region or segment as a closed object and determine the arithmetic Mode value for the pixel intensity, in that region. This paper proposes a new algorithm to identify and isolate different regions within the breast ROI and detect abnormalities, if present, within the breast region.

1) Binary Homogeneity Enhancement Algorithm (BHEA)

In this paper we have proposed a new homogeneity enhancement process namely Binary Homogeneity Enhancement Algorithm (BHEA) for digital mammogram. According to this method we take a maximum difference threshold (MDT) value, which is constant threshold determine by observation. We start checking this value with the image data by horizontally scanning from left of the array to the right from the topmost row of the mammogram image. If result of any subtraction is greater than the MDT, the array will be divided into two equal subsets along middle position and the first and last positions of the two subsets will be pushed to stack. Otherwise, the mode value of subset will be propagated to all other position after modifying value using uniform color quantization technique in color space breaking it into sixteen level scales. The process will be continued recursively, checking the stack, till the rightmost pixel in the last row of the mammogram image is reached. The same process will be repeated by scanning the image vertically from top to bottom followed by uniform color quantization. Algorithm for Horizontal Processing:

Step 1. Scan Image from first row from left most pixel to right most pixel.
Step 2. Check if the value of pixel is : MDT-val <pixel intensity<MDT+val .
Step 3. If value differs then divide the array from start to end through the midpoint. Push the Start and End Position of both halves into the Stack and Goto Step2.
Step 4. If value of consecutive pixels satisfy the condition then continue till the value differs. Then find Mode value of all the pixels of similar intensity and replace them in the image after performing uniform color quantization for the set of pixels.
Step 5. Check Stack. If not empty, Pop value from Stack and Goto Step2.
Step 6. Continue till rightmost pixel of the Last row is reached.

Algorithm for Vertical Processing: The same process, as mentioned above, by replacing row wise scanning with column wise scanning starting from first column from top most pixel to bottom most pixel and continuing till bottom most pixel of the last column is reached.
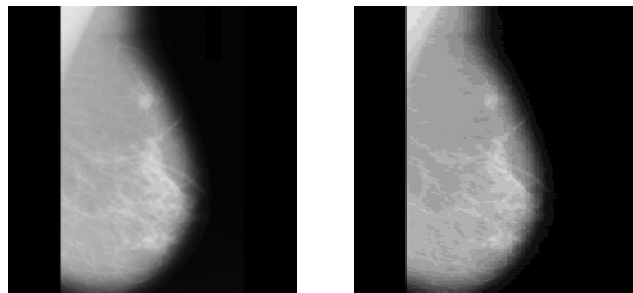

Figure 2.   Mammogram before and after BHEA (MIAS 023.R)

2) Edge Detection Algorithm (EDA)

In this research paper we have proposed a new method for detecting edge rather than using one of the known methods of edge detection. In the first step of our proposed method we perform horizontal scanning. If any change of pixel intensity is observed it is marked by a black pixel indicating a horizontal edge point. We continue this process for all rows of pixel data to obtain a Horizontal Edge Map. In the next step, we scan the image vertically. Continuing the process for all the columns we obtain a Vertical Edge Map image. Finally, we merge the Horizontal Edge Map with Vertical Edge Map by performing a logical OR operation on the two image files, to obtain the Edge map of mammogram image.

Algorithm for Horizontal Image Map:

Step 1. Scan the Image Array Horizontally from left-most pixel to right-most pixel from first row to last row.
Step 2. Take the first pixel intensity value as a reference value.
Step 3. Compare intensity of subsequent pixels with the reference value. If the value is same continue to next pixel.
Step 4. If the value differs, change the value of reference value to the pixel intensity value and mark the pixel black.
Step 5. If the last row and column pixel is not reached then Goto Step3.

Algorithm for Horizontal Image Map: The same process, as mentioned above, by replacing row wise scanning with column wise scanning starting from top-most pixel to bottom-most pixel from first column to last column and continue untill the last column and row pixel is not reached.
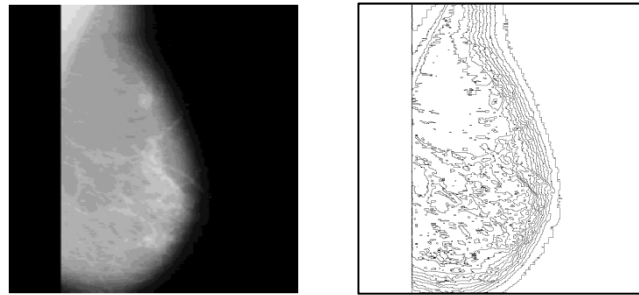
Figure 3.   Mammogram before and after EDA (MIAS 023.R)

3) *Breast Boundary Detection Algorithm (BBDA)*

The breast region contains the outline edges of various breast constituents bounded by the breast boundary that defines the breast region from the image background. For further processing it is of utmost importance to extract the boundary of the breast. In this research paper we now proposed a new method for extraction of the breast boundary.

We now proceed with the output image of the edge extraction process and our objective is to identify the outermost edge line that constitutes the edge of the breast. We scan the image from the right side of the image to locate the rightmost point at the first row of the image, which is the starting point of the processing. We then consider all the surrounding pixels in a clockwise priority and consider the pixel with the highest priority. We store the pixel traversed in a Plotting List to be used later for drawing the breast boundary. The pixels that surrounded the starting pixel, but are of lower priority are stored in a Backtrack Stack to be used only if the traversal process reaches a dead end. If a dead end is reached, we pop out from the Backtrack stack a lesser priority pixel and continue with the traversal process. The traversal continues to the next pixel and continues till it reaches the baseline or the bottom of the image, indicating the end of the breast region. The Plotting list contains the breast boundary pixels that are plotted after smoothening for further processing.

Algorithm to Detect Breast Boundary:

Step 1.   Scan the image from the Left side of the image to locate the leftmost pixel of the breast region.
Step 2.   Draw a vertical line along this pixel from top to bottom representing the Left baseline or boundary.
Step 3.   Scan the edge map from the right side to left, from the first row.
Step 4.   Obtain a pixel that is black indicating an edge path, traverse the pixel path by considering all the surrounding pixels in a clockwise priority and consider the pixel with the highest priority.
Step 5.   The pixels that surrounded the edge pixel, but are of lower priority are stored in a Backtrack Stack to be used only if the traversal process reaches a dead end.
Step 6.   If a dead end is reached, pop out from the Backtrack stack a lesser priority pixel and continue with the traversal process.
Step 7.   Store the pixels traversed in a Plotting List to be used later for drawing the Breast boundary.
Step 8.   Traversal continues to the next pixel till it reaches the left baseline or the bottom of the image.
Step 9.   If the bottom of the image or the left baseline is not reached the path is discarded, the plot list is erased and continue from Step4.
            Else Get pixels from the Plotting List at a discrete interval.
Step 10. Plot the first pixel.
Step 11. Draw a simple curve between two consecutive pixel position separated by the discrete interval.
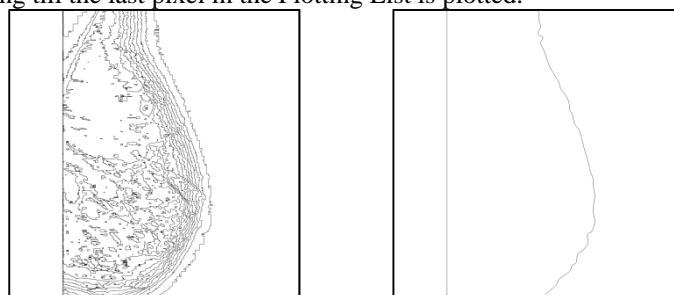Step 12. Continue plotting till the last pixel in the Plotting List is plotted.

Figure 4.   Mammogram Edge Map along with derived Breast Contour (MIAS 023.R)

4) *Pectoral Muscle Detection Algorithm (PMDA)*

After we detect the edges of the mammogram by our proposed method, we get an image that contains the breast region clearly differentiated from the background. For further processing it is of utmost importance to extract the breast region by eliminating the pectoral muscle region. In this research paper we now proposed a new

method for detection, extraction and isolation of Pectoral Muscle from the breast region. A careful and close study of all the mammogram images in the MIAS database, it is observed that the pectoral muscle have a distinct intensity value that differs from the rest of the fatty tissues on the mammogram image. Similarly, it is also observed that the pectoral muscle lies at the top left corner of the breast region (Right pair flipped horizontally during preprocessing) and it traverses from the top margin downward towards the left baseline of the mammogram image forming an inverted triangle. Our objective is to identify this triangle from the breast region.

We now need to identify the outermost edge line that constitutes the edge of the Pectoral muscle. It is observed that the largest inverted triangle in the edge map starting from the top margin and ending on the vertical baseline, of the breast region on the left side, is the pectoral region. We first identify the rightmost pixel of the breast region on the right side and draw a vertical line from top margin to bottom margin, parallel to the left vertical baseline. We draw another line parallel to the top and bottom margin passing through the rightmost pixel of the breast region on the right vertical line and ending on the left vertical baseline. We limit our search to the obtained rectangle thus reducing the processing time that would have incurred if we considered the entire mammogram. It is also observed that almost all, barring a few mammograms, have the pectoral region confined within this rectangle.

Algorithm for Pectoral Muscle Isolation:

Step 1. Scan the image from the right side of the image to locate the rightmost pixel of the breast region.
Step 2. Draw a vertical line along this pixel from top to bottom.
Step 3. Draw a horizontal line parallel to the top margin passing through the rightmost pixel on the right vertical line to the left base line.
Step 4. Scan the edge map from the right side to left, on the obtained rectangle, from the first row.
Step 5. Obtain a pixel that is black indicating an edge path, traverse the pixel path by considering all the surrounding pixels in a clockwise priority and consider the pixel with the highest priority.
Step 6. The pixels that surrounded the edge pixel, but are of lower priority are stored in a Backtrack Stack to be used only if the traversal process reaches a dead end.
Step 7. If a dead end is reached, pop out from the Backtrack stack a lesser priority pixel and continue with the traversal process.
Step 8. Store the pixels traversed in a Plotting List to be used later for drawing the pectoral boundary.
Step 9. Traversal continues to the next pixel till it reaches the left baseline or the bottom of the rectangle.
Step 10. If the bottom of the rectangle is reached the path is discarded, the plot list is erased and continue from Step5.
Else Get pixels from the Plotting List at a discrete interval.
Step 11. Plot the first pixel.
Step 12. Draw a simple curve between two consecutive pixel position separated by the discrete interval.
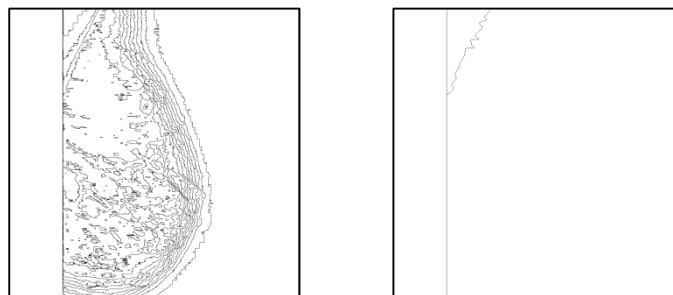Step 13. Continue plotting till the last pixel in the Plotting List is plotted.

Figure 5.   Mammogram Edge Map along with Isolated Pectoral Boundary (MIAS 023.R)

5) Anatomical Segmentation of Breast ROI (ASB)

After merging the images obtained from Pectoral Muscle Detection Algorithm (PMDA) and Breast Boundary Detection Algorithm (BBDA) we get the ROI of the Breast. We now consider the edge map that corresponds only to the ROI of the breast. The edge map indicates various closed structures within the breast region that corresponds to the different anatomical regions of the breast namely, the fatty tissues, ducts, lobules, glands, irregular mass and calcifications. Our objective at this stage is to identify these regions on the mammogram image specially the presence of any abnormalities on the breast region.

In this research paper we propose a new segmentation scheme for Anatomical Segmentation of breast. We start by identifying the left baseline of the breast image from the edge map. We then draw a line vertically from top to bottom identifying the left boundary of the breast. We then scan the breast boundary on the right side to locate the rightmost pixel on the breast contour. After the pixel is located we draw another vertical line from top to bottom passing through the rightmost pixel thus partitioning the image only to the breast ROI. This process optimizes the algorithm and increases the processing efficiency. At this stage we try to locate all edge paths that

are circular or terminate either on the left base line or the bottom of the image, forming a closed structure. This process removes all noise and discrete objects from the edge map that are inconsequential to the image.

We start by locating all the edge paths that originates from the top margin line namely the first row of the image. The algorithm travels each individual path and stores them on the plotting list. This list is plotted on another image if the edge path is circular or end on the last row of the image or the vertical line representing the left boundary of the breast. After all pixel paths on the first row are traversed the algorithm repeats similar scanning and traversal of all pixel paths from the row that is indicated by dividing the image vertically into sixteen segments, like a ladder. The first row of each segment of the ladder is used for locating pixel paths for traversal. All complete paths are then plotted on another image thus providing the anatomical regions of the breast.

Algorithm to detect the Anatomical regions of Breast:

Step 1. Scan the image from the Left side of the image to locate the leftmost pixel of the breast region.
Step 2. Draw a vertical line along this pixel from top to bottom representing the Left baseline or boundary.
Step 3. Scan the image from the right side of the image to locate the rightmost pixel of the breast region.
Step 4. Draw a vertical line along this pixel from top to bottom.
Step 5. Partition the obtained rectangle horizontally into sixteen segments and start with the first row of the first segment.
Step 6. Scan the enclosed rectangle from the right side to left, from the first row of the segment.
Step 7. Obtain a pixel that is black indicating an edge path, traverse the pixel path by considering all the surrounding pixels in a clockwise priority and consider the pixel with the highest priority.
Step 8. The pixels that surrounded the edge pixel, but are of lower priority are stored in a Backtrack Stack to be used only if the traversal process reaches a dead end.
Step 9. If a dead end is reached, pop out from the Backtrack stack a lesser priority pixel and continue with the traversal process.
Step 10. Store the pixels traversed in a Plotting List for plotting.
Step 11. Traversal continues to the next pixel till it reaches the left baseline or the bottom of the image or the start position is reached.
Step 12. If the traversal is terminated, the plotting list is erased and continues from Step7.
Else plot pixels from the Plotting List.
Step 13. Continue to Step6 till all black pixels, indicating an edge path, is traversed.
Step 14. Move to the first row of the next segment and continue from Step6 to Step11.
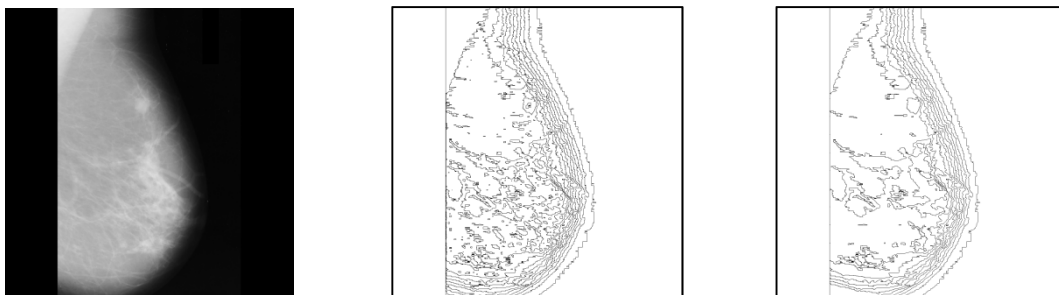

Figure 6.   Original Mammogram along with Edge Map showing Breast Region of Interest and Anatomical Regions after ASB (MIAS 023.R)

6)   Seeded Region Growing Algorithm (SRGA)

The segmentation process performed on the edge map differentiates various regions on the breast depending on their intensity values. Each region has a different intensity value. The fatty tissues, glands, lobules and the ducts display different intensity values and thus can be segregated into different regions. An abnormality such as a mass, tumors or calcifications may be present within the breast has distinctly higher intensity values than the normal tissues of the breast. So in this step we need to categorize all the obtained closed structures to their intensity values. The distribution of pixels intensities also vary within each segmented region but the majority of the pixels have similar intensity values. So for each region we calculate the arithmetic Mode value for the intensities from the original mammogram and replace those pixels in the region with the computed mode values.

To identify, isolate each closed structure and then performing coloring of each structure with their respective Mode values we proposed a new algorithm by modifying the seeded region growing method for coloring bounded objects. Each region, within the mammogram, is bounded by a single pixel boundary as obtained during the edge detection process followed by the process of anatomical segmentation. During this process we scan the segmented image as obtained after performing the Anatomical Segmentation of Breast (ASB), to locate region that is yet to be colored. On finding the seed for a region we start coloring the region by first comparing the pixel intensity of that pixel location on the original mammogram image. For each pixel the four boundary pixels

located north, east, west and south of the pixel is also checked to find out whether they are colored or they form the boundary pixel. If the pixels are not colored and are not boundary pixels they again form the seed for further searching. A stack is used to store the seeds to be investigated, while we use a List to store the pixels of the regions that have been included to the region. All the pixel positions within the List are then searched on the original image to get their intensity values for calculating the Mode value. The pixel locations of each region are then replaced by the computed Mode value intensity.

Algorithm for Seeded Region Growing:

Step 1. Scan the image from the Left side of the breast ROI from first row.
Step 2. If the pixel is not colored and is not a boundary pixel consider it as a seed for coloring.
Step 3. Store Pixel in a Seed Stack.
Step 4. Consider pixels at top, bottom and two sides of the seed pixel.
Step 5. If the pixel is not colored and is not a boundary pixel Push to Seed Stack.
Step 6. Pop from Seed Stack a seed pixel and store the pixel into Color List for coloring.
Step 7. Continue till no white pixels are left in the bounded region and Seed Stack is not empty, and Goto Step2.
Step 8. After getting the complete region on the Color List Scan the Original Mammogram for the Pixel Locations.
Step 9. Calculate the Mode Value of the pixel intensities in the image for the set of Pixel Locations.
Step 10. Color the pixels representing the region with the corresponding Mode intensity value.
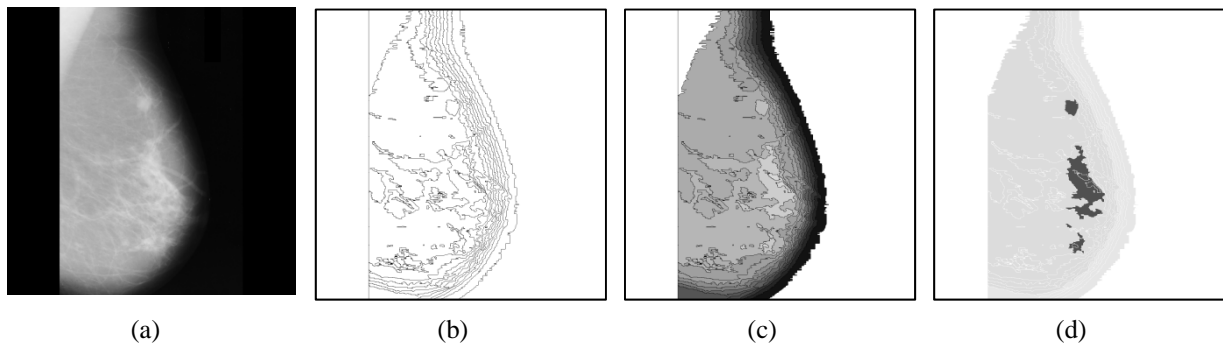Step 11. Clear Color List and continue till the Breast Regions are completely colored, Goto Step1.

| (a) | (b) | (c) | (d) |

Figure 7.   MIAS mdb023.R a) Original Mammogram, b) Anatomical Regions, c) Mammogram after SRGA, d) Isolated Abnormal Masses

IV.   EXPERIMENTAL RESULT EVALUATION

A. Experimental Results

1) Case 1: Normal Mammogram

Results obtained by applying the proposed algorithms on MIAS image mdb272 comprised of predominantly Fatty-Fibro Glandular tissue where no abnormalities is present. This image represents the control image for our experiment indicating that no mass is detected.
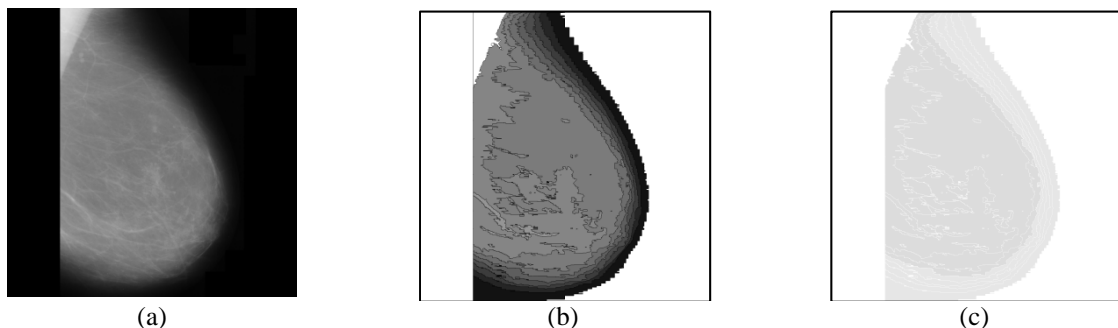

| (a) | (b) | (c) |

Figure 8.   MIAS mdb272.L: a) Original Mammogram, b) Anatomical Regions after SRGA, c) Mammogram showing absence of abnormal region

2) Case 2: Fatty Tissue with Abnormalities

Results obtained by applying the proposed algorithms on MIAS image mdb028 comprised predominantly Fatty tissues where abnormalities are present.
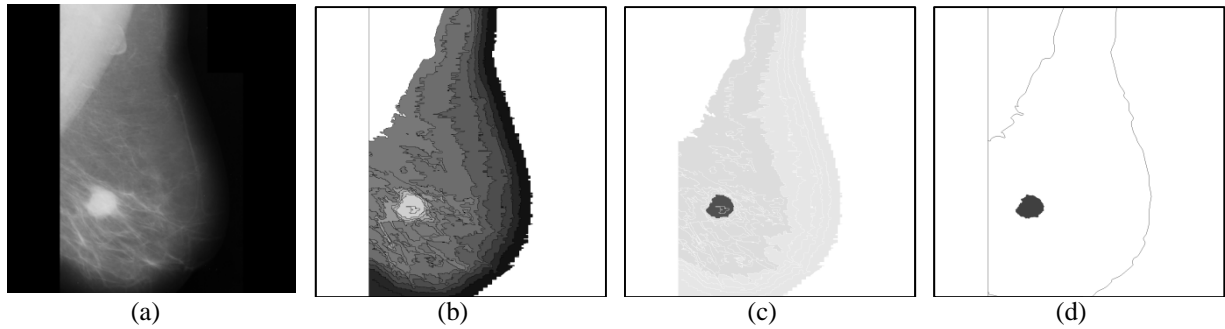
Figure 9.   MIAS mdb028.L: a) Original Mammogram, b) Anatomical Regions after SRGA, c) Mammogram showing presence of abnormal region, d) GT of Abnormal Region

3)   Case 3: Fatty-Fibro Glandular Tissue with Abnormalities

Results obtained by applying the proposed algorithms on MIAS image mdb265 comprised predominantly Fatty Fibro-Glandular tissues where abnormalities are present.
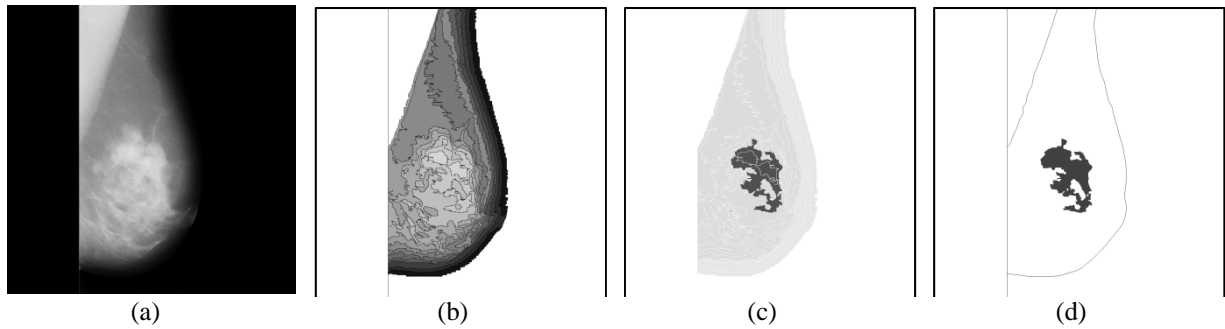

Figure 10.   MIAS mdb265.R: a) Original Mammogram, b) Anatomical Regions after SRGA, c) Mammogram showing presence of abnormal region, d) GT of Abnormal Region

4)   Case 4: Dense-Fibro Glandular Tissue with Abnormalities

Results obtained by applying the proposed algorithms on MIAS image mdb202 comprised predominantly Dense Fibro-Glandular tissues where abnormalities are present.
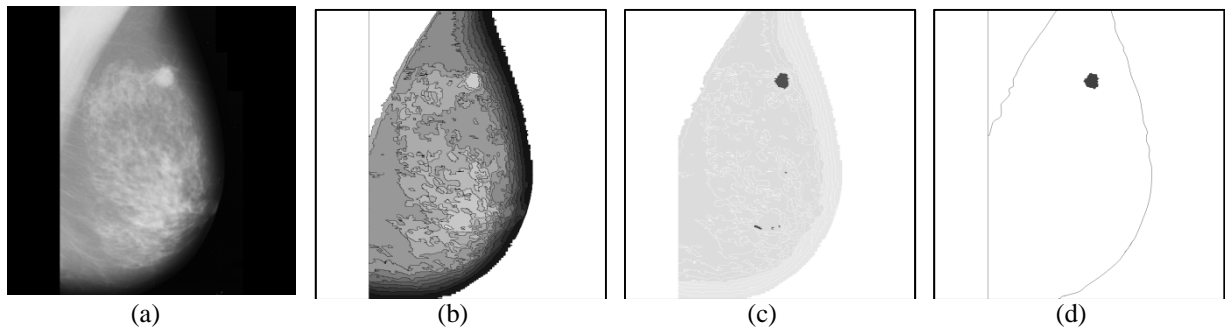

Figure 11.   MIAS mdb202.L: a) Original Mammogram, b) Anatomical Regions after SRGA, c) Mammogram showing presence of abnormal regions, d) GT of Abnormal Region excluding insignificantly small abnormal regions

B.   Accuracy Estimation

Performance evaluation in algorithm design is an important step that is commonly neglected. What constitutes an "acceptable" result differs significantly, and is often based on visual subjective opinion with very little quantitative endorsement. The accuracy of this technique was evaluated through quantitative measures derived through the comparison of each segmented mammogram representing an abnormal mass, as "mask" with its corresponding "gold standard". The gold standard is generated by manually segmenting the breast region representing an abnormal mass from each mammogram. The boundary of the regions is then manually traced to extract the abnormal breast region to generate a ground truth (GT) image.

A quantitative measure was then derived to describe the accuracy of the segmentation. The region extracted by the segmentation algorithm (mask), which matches the GT, is denoted as true positive (TP) emphasizing that the algorithm has indeed found a portion of the breast. Pixels shown in the GT but not shown in the mask are defined as false negative (FN) classifications. These are considered missing pixels in the abnormal mass region. Conversely, the pixels not in the GT, but in the mask are defined as false positive (FP) pixels. By using these

parameters on the output of all the mammogram images in MIAS database having abnormalities, we obtain the following result on different quality measures.

TABLE I. Common Measures Used in the Evaluation of Our Proposed Methods

Common measures	Computation	Avg Result
Accuracy (Percentage agreement)	|TN|+|TP|/|TN|+|TP|+|FP|+|FN|	0.9987
Dice similarity coefficient (DSC)	2×|TP|/2×|TP|+|FP|+|FN|	0.8796
Error rate	|FP|+|FN|/|FP|+|FN|+|TP|+|TN|	0.0012
Sensitivity (Percentage of Correct Estimation)	|TP|/|TP|+|FN|	0.9421
Specificity (True Negative Fraction/Rate)	|TN|/|TN|+|FP|	0.9299
False Positive Fraction/Rate	1 - Specificity	0.0700
Under estimation fraction (UEF)	|FN|/|TN|+|FN|	0.0005
Over estimation fraction (OEF)	|FP|/|TN|+|FN|	0.0006

C. Statistical Analysis

The results obtained during Experimental Evaluation phase on the normal and abnormal cases are further analyzed using statistical methods to indicate that the abnormal regions show distinct characteristics in comparison to normal regions. During the generation of the distinct regions within the mammogram in SRGA we have also categorized and enumerated each region along with their respective Statistical Mode value. This dataset is used for further statistical analysis. First we calculate the Arithmetic mean for the distribution to obtain the deviations of each region, so that we derive the Standard Deviation of the dataset. We then calculate the Z score to normalize the distribution. We truncate the Z values that are negative as they do not influence the abnormal regions. Using the obtained truncated dataset we have calculated the truncated mean and the Standard Deviation (σ) of the apportioned dataset. Finally, we calculate the 2σ and 3σ of the population. The results indicate that the abnormal region(s) lie on or above the 3σ mark. The control image (normal mammogram) is devoid of this region in the graph.
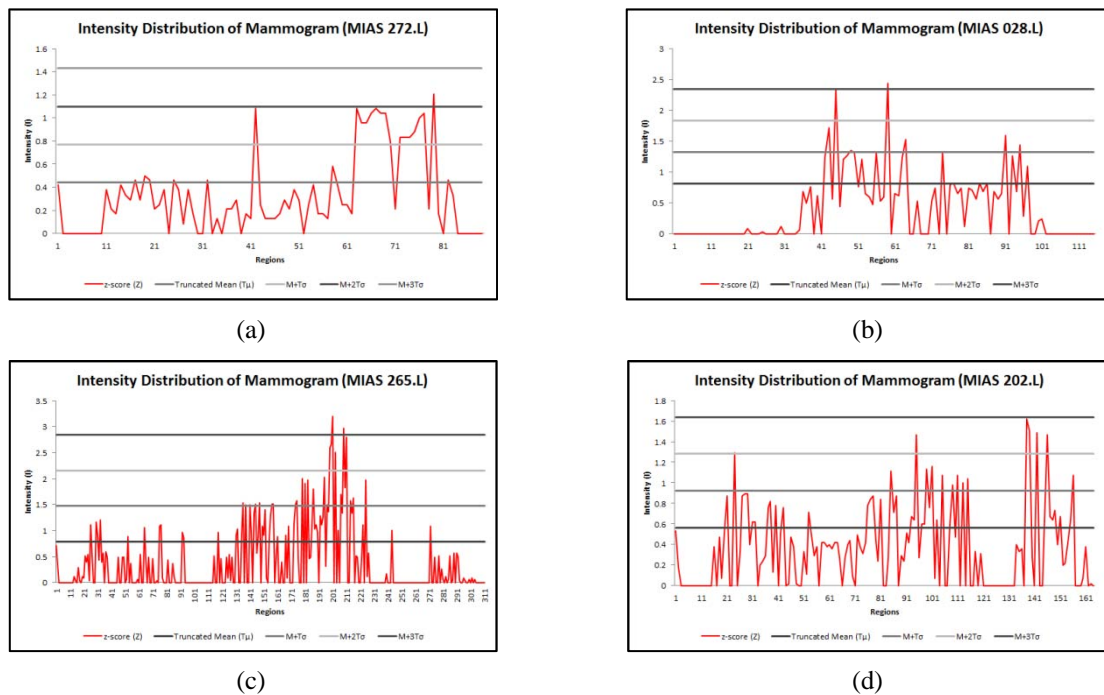

(a)

(b)

(c)

(d)

Figure 12. Intensity Distribution of Mammogram Images (a) showing Normal case where (b), (c) and (d) showing abnormal cases of Fatty, Fatty-Fibro Glandular and Dense-Fibro Glandular mammogram.

V. CONCLUSION

In our proposed method, we have developed several distinct algorithms and the combination of these methods have provided excellent detection of mass in mammograms having abnormal regions. The accurate mass detection was achieved by the proposed Binary Homogeneity Enhancement Algorithm (BHEA), followed by Edge Detection Algorithm (EDA), Breast Boundary Detection Algorithm (BBDA), Pectoral Muscle Detection Algorithm (PMDA), Anatomical Segmentation of Breast ROI (ASB) and finally by our proposed Seeded Region Growing Algorithm(SRGA). Results of the proposed methods show a reliable detection rate of abnormal regions if present within the mammogram, furthermore, due to its simple procedure the method executes faster than other complicated methods. Performance evaluation in algorithm design is a commonly neglected concept. Few mammogram segmentation algorithms have been tested extensively. Our proposed method has been tested on all mammograms from mini-MIAS mammogram database. The results have detected abnormalities in only those mammograms which contains abnormal mass according to the MIAS database. The other mammograms that are normal have shown negative results. Further we have evaluated our findings by proven statistical methods which have also authenticated the accuracy of our detection methods. These methods can be used for other medical image processing applications with minor modification. In conclusion, it can be mentioned, our proposed methods are acceptably accurate, promising and comparable with any other standard methods.

REFERENCE

[1] Medindia, "Breast Cancer In India Rising Rapidly", January 22, 2006.
[2] Gajalakshmi Vendhan, Aleyamma Mathew, Paul Brennan, Balakrishnan Rajan, Vendhan C. Kanimozhi, Anitha Mathews, Beela S. Mathew and Paolo Boffetta "Breastfeeding and breast cancer risk in India: a multicenter case-control study", Intl J Cancer, 2009, 125 (3), 662-665.
[3] L. Shen, R. Rangayyan, and J. Desaultels, "Detection and Classification Mammographic Calcifications", International Journal of Pattern Recognition and Artificial Intelligence. Singapore: World Scientific, 1994, pp. 1403–1416.
[4] F. Aghdasi, R.Ward, and B. Palcic, "Restoration of mammographic images in the presence of signal-dependent noise", in State of the Art in Digital Mammographic Image Analysis. Singapore: World Scientific, 1994, vol. 7, pp. 42–63.
[5] Y. Chitre, A. Dhawan, and M. Moskowtz, "Artificial neural network based classification of mammographic microcalcifications using image structure features", in State of the Art of Digital Mammographic Image Analysis. Singapore: World Scientific, 1994, vol. 7, pp. 167–197.
[6] N. Karssemeijer, "Computer-Assisted Reading Mammograms", European Radiol., vol. 7, pp. 743–748, 1997.
[7] Pisano and F. Shtern,"Image processing and computer-aided diagnosis in digital mammography," in State of the Art of Digital Mammographic Image Analysis. Singapore: World Scientific, 1994, vol. 7, pp. 280–291.
[8] Sterns EE, "Relation between clinical and mammographic diagnosis of breast problems and the cancer/ biopsy rate," Can. J. Surg., vol. 39, no. 2, p 128-132, 1996
[9] Highnam R and Brady M, "Mammographic Image Analysis", Kluwer Academic Publishers, 1999. ISBN: 0-7923- 5620-9
[10] Kekre HB, Sarode Tanuja K and Gharge Saylee M, "Tumor Detection in Mammography Images using Vector Quantization Technique", International Journal of Intelligent Information Technology Application, 2009, 2(5):237-242
[11] Petrick N, Chan HP, Wei D, et al. Automated detection of breast masses on mammograms using adaptive contrast enhancement and texture classification. Med Phys 1996; 23:1685–1696.
[12] Zheng B, Chang YH, Gur D. Computer detection of masses in digitized mammograms using single-image segmentation and a multilayer topographic feature analysis. Acad Radiol 1995; 2:959–966.
[13] Catarious DM, Baydush AH, Floyd CE. Incorporation of an iterative, linear segmentation routine into a mammographic mass CAD system. Med Phys 2004; 31:1512–1520.
[14] Timp S, Karssemeijer N. A new 2D segmentation method based on dynamic programming applied to computer aided detection in mammography. Med Phys 2004; 31:958–971.
[15] Te Brake GM, Karssemeijer N. Segmentation of suspicious densities in digital mammograms. Med Phys 2001; 28:259–266.
[16] Vyborny CJ, Doi T, O'Shaughnessy KF, et al. Breast cancer: importance of spiculation in computer-aided detection. Radiology 2000; 215: 703–707.
[17] Kegelmeyer WP, Pruneda JM, Bourland PD, et al. Computer-aided mammographic screening for spiculated lesions. Radiology 1994; 191: 331–337.
[18] Karssemeijer N, Te Brake GM. Detection of stellate distortions in mammograms. IEEE Trans Med Imaging 1996; 15:611–619.
[19] M. L. Comer, S. Liu, and E. J. Delp, "Statistical segmentation of mammograms," in Digital Mammography, K. Doi, Ed., International Congress Series. Amsterdam, the Netherlands: Elsevier, 1996, pp.471–474.
[20] H. D. Li, M. Kallergi, L. P. Clarke, and V. K. Jain, "Markov random field for tumor detection in digital mammography," IEEE Trans. Med. Imag., vol. 14, pp. 565–576, June 1995.
[21] Huo Z, Giger ML, Vyborny CJ. Analysis of spiculation in the computerized classification of mammographic masses. Med Phys 1995; 22:1569–1579.
[22] Ryohei Nakayama and Yoshikazu Uchiyama "Development of New Filter Bank for Detection of Nodular Patterns and Linear Patterns in Medical Images" Systems and Computers in Japan, Vol. 36, No. 13, 2005.
[23] M. Sameti and R. K. Ward, "A fuzzy segmentation algorithm for mammogram partitioning," in Digital Mammography, K. Doi, Ed., International Congress Series. Amsterdam, the Netherlands: Elsevier, 1996, pp. 471–474.
[24] Sahiner B, Chan HP, Petrick N, et al. Improvement of mammographic mass characterization using spiculation measures and morphological features. Med Phys 2001; 28:1455–1465.
[25] Zheng B, Lu A, Hardesty LA, et al. A method to improve visual similarity of breast masses for an interactive computer-aided diagnosis environment. Med Phys 2006; 33:111–117.
[26] A.D.Brink and N.E.Pendock, "Minimum Cross-entropy Threshold Selection, Pattern Recognition", 29(1): 179-188, 1996.