

Accurate Breast Contour Detection Algorithms in Digital Mammogram

Indra Kanta Maitra
Research Fellow
Department of Computer
Science & Engineering,
University of Calcutta, India

Sanjay Nag
Research Scholar
Department of Computer
Science & Engineering,
University of Calcutta, India

Samir K Bandyopadhyay
Professor
Department of Computer
Science & Engineering,
University of Calcutta, India

ABSTRACT

Computer Aided Diagnosis (CAD) systems have improved diagnosis of abnormalities in mammogram images. The principal feature within the breast region is the breast contour. Extraction of the breast region and delineation of the breast contour allows the search for abnormalities to be limited to the region of the breast without undue influence from the background of the mammogram. After performing an essential pre-processing step to suppress artifacts and accentuate the breast region, the exact breast region as the region of interest (ROI), has to be segmented. In this paper we present a fully automated segmentation and boundary detection method for mammographic images.

In this research paper we have proposed a new homogeneity enhancement process namely Binary Homogeneity Enhancement Algorithm (BHEA) for digital mammogram. This is followed by a novel approach for edge detection (EDA) and finally obtaining the breast boundary by using our proposed Breast Border Boundary Enhancement Algorithm. This composite method have been implemented and applied to mini-MIAS, one of the most well-known mammographic database consisting of 322 medio-lateral oblique (MLO) view obtained via a digitization procedure. To demonstrate the capability of our segmentation algorithm it was extensively tested on mammograms using ground truth images and quantitative metrics to evaluate its performance characteristics. The experimental results indicate that the breast boundary regions were extracted accurately characterize the corresponding ground truth images. The algorithm is fully autonomous, and is able to preserve, skin and nipple (if in profile), a task very few existing mammogram segmentation algorithms can claim.

Keywords

Mammogram Segmentation, Contrast Limited Adaptive Histogram Equalization (CLAHE), Binary Homogeneity Enhancement Algorithm (BHEA), Edge Detection Algorithm (EDA), Breast Boundary Detection Algorithm (BBDA).

1. INTRODUCTION

Breast cancer is a leading cause of cancer deaths among women. For women in US and other developed countries, it is the most frequently diagnosed cancer. About 2100 new cases of breast cancer and 800 deaths are registered each year in Norway. In India, a death rate of one in eight women has been reported due to breast cancer [1]. Breast cancer begins in breast tissue, which

is made up of glands for milk production, called lobules, and the ducts that connect lobules to the nipple. The remainder of the breast is made up of fatty, connective and lymphatic tissue [2].

Mammogram is one of popular technique to identify breast cancer. Mammography is the most effective and economical breast imaging modality due to its simplicity, portability and cost effectiveness. Digital mammography is a technique for recording x-ray images in computer code. The images are displayed on a computer monitor and can be enhanced before they are printed on film. Images can also be manipulated; the radiologist can magnify or zoom in on an area. Digital mammography may have some more advantages over conventional mammography like the images can be stored and retrieved electronically.

Recent studies showed that the interpretation of the mammogram by the radiologists give high rates of false positive cases. Indeed 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 tools. They could help the radiologists in the interpretation of the mammograms and could be useful for an accurate diagnosis [3, 4, 5].

Imaging techniques play an important role, especially of abnormal areas that cannot be felt but can be seen on a conventional mammogram [6]. Mammogram segmentation technique is mostly used process to detect the abnormalities. Mammogram segmentation consists of so many methods out of which breast boundary contour extraction is most important one. Before applying any image-processing algorithm on mammogram, pre-processing steps are very important in order to limit the abnormalities without undue influence from background of the mammogram. These steps are needed only on digitized screen film mammography (SFM) images because digital mammography devices perform this step automatically during the image storing process. On images obtained directly from the digital mammography devices segmentation process is much easier. Previous work from many authors used mammography image databases including this paper, especially MiniMIAS [7] and DDSM [8], both comprised of scanned and digitized SFM images.

As stated earlier, breast boundary detection as well as breast segmentation is a very important feature in image analysis. Breast boundary detection is important factor, which can

improve clinical diagnosis of mammographic diseases. The accurate segmentation of the breast region in mammograms is an essential step in the computerized analysis of mammograms. It allows the search for abnormalities to be limited to the region of the breast without undue influence from the background of the mammogram. It also facilitates enhancements for techniques such as comparative analysis, which includes the automated comparison of corresponding mammograms. The breast boundary contains significant information relating to the symmetry and deformation between two mammograms. Hence in this work, we have proposed novel step-by-step algorithms for mammogram segmentation, determine the edge of segmented region, extract the breast boundary and improve the breast boundary using boundary enhancement. 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, these methods are more accurate and reliable compared with other common methods.

2. PREVIOUS WORKS

Numerous promising approaches are coming up for isolating the breast region in mammograms. Some of these have focused on using thresholding [9, 10], gradients [11], modelling of the non-breast region of a mammogram using a polynomial [12], or active contours [13]. Few of them studied by us have shown promising results.

One of the earliest approaches to segmentation of the breast contour, was presented by Semmlow et al. [14]. They used a spatial filter and Sobel edge detector to locate the breast boundary on xeromammograms. The most obvious approach would seem to be the use of thresholding [9,10], however since there is usually a certain amount of overlap between the breast region and background, such partitions will inevitably result in the misclassification of some background pixels as breast region and vice versa. Recent efforts, such as that of Masek et al. [15] using local thresholding have shown more promise.

Abdel-Mottaleb et al. [16] use a system of masking images with different thresholds to find the breast edge. Gradients are calculated from the images as to where the skin should be located and a union of two images can find the entire breast area.

Méndez et al. [11] find the breast contour using a gradient based method. They first use a two-level thresholding technique to isolate the breast region of the mammogram. The mammogram, oriented with the mammogram facing upwards and smoothed, is then divided into three regions using a number of automatically determined reference points and a tracking algorithm is applied to the mammogram to detect border.

The global segmentation approach cited by Bick et al. [9] incorporates aspects of thresholding, region growing and morphological filtering. The mammogram is initially filtered to reduce noise and then features are extracted using a texture operator which calculates the range, which is the difference between the maximum and minimum intensity values in a neighbourhood, for each candidate pixel. A histogram is then constructed for all pixels whose local range was minimal. This histogram was then used to classify pixels as belonging to either the breast or non-breast regions. Region growing is then used to label the different regions, morphological filtering is used to

eliminate irregularities along the breast contour and contour tracing extracts the breast contour.

An interesting algorithm was described by Lou et al. [17]. It is based on the assumption that the trace of intensity values from the breast region to the air-background is a monotonic decreasing function. The algorithm first searches for an initial boundary using a clustered image. For each initial boundary point a corresponding point is estimated with an extrapolation method. Through a refinement process, a contour point is derived from the extrapolated point and by linking all the boundary points, the breast contour is defined.

There have also been various attempts to use active contours, such as that of McLoughlin and Bones [18]. They first derive an approximate separation of the breast region and background using a global threshold. Pixels below this threshold are used to obtain a model of the background of the mammogram using Poisson approximation. The threshold found by the Poisson model is used to form a binary mask from which an initial contour can be extracted and is smoothed using the greedy snake algorithm. The algorithm reports acceptable results when tested on 40 mammograms.

A semi-automated method based on the concept of "united snakes"[19] is described by Ojala et al. [20]. It uses an interactive boundary tracing technique called livewire to initialize the snake. The united snake compactly unifies the most significant snake variants, allowing the user to choose the most appropriate snake. The algorithm is tested on mammograms from two databases using the "basic" snake as defined by Kass et al. [21], with the results being smoothed breast contours. Ojala et al. [22] later describe an active contour method for smoothing breast contours in mammograms as part of a comparison with two other methods, namely Fourier smoothing and BSpline approximation. A more recent attempt at using snakes [13] uses preliminary thresholding and noise removal algorithm in combination with a modified "greedy" snake algorithm [23].

One of the inherent limitations of these methods is the fact that very few of them preserve the skin or nipple in profile. One of the more promising methods in this respect focuses on modelling the non-breast region of a mammogram using a polynomial. The method, as described by Chandrasekhar and Attikiouzel [24, 12], involves modelling the non-breast region (background) of a mammogram as a polynomial and subtracting it from the original mammogram. It is based on the assumption that the intensities comprising the background are low in value and lie within a closed interval. An initial threshold is used to approximate the breast region. This region includes the whole breast region and a small portion of the breast contiguous with the non-breast region is included in the region being modelled. This modelled background is then subtracted from the original mammogram, yielding a difference image which, when thresholded, results in a binary mammogram. A connected components algorithm is then used to identify and merge related regions, followed by morphological operations to smooth irregularities to yield a labelled binary mammogram representing the breast/non-breast association. The boundary between the two regions represents the contour of the breast.

M. Wirth et al. explored a new algorithm for breast region segmentation using fuzzy reasoning. The algorithm uses

morphological pre-processing to suppress artefacts and accentuate the breast region, followed by a fuzzy rule-based algorithm to classify the breast tissue region. To demonstrate the capability of segmentation algorithm that is extensively tested on mammograms from two databases using ground truth images and quantitative metrics to evaluate its performance characteristics [25].

3. PROPOSED METHODS

Digital Mammograms are medical images that are difficult to interpret, thus a preparation phase is needed in order to improve the image quality and make the segmentation results more accurate. 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. After obtaining the processed mammogram image, we propose the following steps:

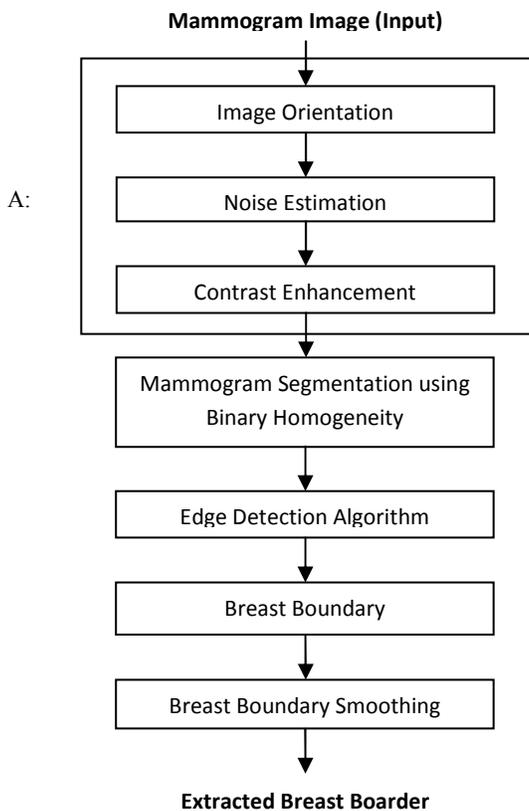


Figure 1. Proposed Algorithm Flowchart (A: Preprocessing)

3.1 Mammogram Preprocessing

3.1.1 Image Orientation

The Orientation of the Mammogram is determined according to Hierarchical Segmentation of Mammograms Based on Pixel Intensity [26]. 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 [27] twice in the original

image. By estimating the first derivatives of these pixels using the appropriate masks, we can determine the chest wall location. 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.

3.1.2 Noise Estimation

There are different types of noises, which appear in miniMIAS images. The algorithm should estimate these regions and exclude them from the remaining process.

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.

Such regions can be identified and removed from the image. The property of a tape artifact, which is used for the detection and exclusion from the further processing, is the horizontal line corresponding to their edges. Hierarchical Segmentation of Mammograms Based on Pixel Intensity [26] provides detail algorithms. Using the 2-level Minimum cross-entropy threshold selection technique [27], and combining with logical and morphological operations we can obtain a noise free image.

3.1.3 Contrast Enhancement

The contrast enhancement phase is done using the Contrast Limited Adaptive Histogram Equalization (CLAHE) technique, which is a special case of the histogram equalization technique [28] that functions adaptively on the image to be enhanced. The pixel's intensity is thus transformed to a value within the display range proportional to the pixel intensity's rank in the local intensity histogram. CLAHE [29] is a refinement of Adaptive Histogram Equalization (AHE) where the enhancement calculation is modified by imposing a user-defined maximum, i.e. clip level, to height of the local histogram and thus on the maximum contrast enhancement factor. The enhancement is there by reduced in very uniform areas of the image, which prevent over enhancement of noise and reduces the edge-shadowing effect of unlimited AHE [30].

The CLAHE method seeks to reduce the noise and edge-shadowing effect produced in homogeneous areas and was originally developed for medical imaging [31]. This method has been used for enhancement to remove the noise and reduces the edge-shadowing effect in the pre-processing of digital mammogram [32].

The CLAHE operates on small regions in the image called tiles rather than the entire image. Each tile's contrast is enhanced, so that the histogram of the output region approximately matches the uniform distribution or Rayleigh distribution or exponential distribution. Distribution is the desired histogram shape for the image tiles. The neighbouring tiles are then combined using bilinear interpolation to eliminate artificially induced boundaries. The contrast, especially in homogeneous areas, can be limited to avoid amplifying any noise and reduce edge-shadowing effect that might be present in the image; the CLAHE technique is described below [33]:

Step 1. Mammogram was divided into a number of non-overlapping contextual regions of equal sizes, experimentally set to be 8x8, which corresponds to approximately 64 pixels.

Step 2. The histogram of each contextual region was calculated.

Step 3. A clip limit, for clipping histograms, was set ($t=0.001$). The clip limit was a threshold parameter by which the contrast of the image could be effectively altered; a higher clip limit increased mammogram contrast.

Step 4. Each histogram was redistributed in such a way that its height did not exceed the clip limit.

Step 5. All histograms were modified by the transformation function

$$T(r_k) = \sum_{j=0}^k P_r(r_j) \quad (1)$$

$$\text{Where } P_r(r_j) = \frac{n_j}{n} \quad (2)$$

is the probability density function of the input mammogram image grayscale value j , n is the total number of pixels in the input mammogram image and n_j is the input pixel number of grayscale value j .

Step 6. The neighboring tiles were combined using bilinear interpolation and the mammogram image grayscale values were altered according to the modified histograms.

In our experiment, we defined tiles size i.e. the rectangular contextual regions to 8X8, which is chosen from best result from trial. Contrast factor that prevents over-saturation of the image specifically in homogeneous areas is restricted to 0.01 here to get the optimized output. The number of Bins for the histogram building for contrast enhancing transformation is restricted to 64 and the distribution of histogram is 'Rayleigh' i.e. Bell-shaped for our experimentation. The range is not specified in the experiment to get the full range of output image.

3.2 Mammogram Segmentation

The paper is based on the image segmentation method, which refers to the major step in image processing, the inputs are images and, outputs are the attributes extracted from those images. Segmentation divides image into its constituent regions or objects. Image segmentation techniques can be broadly classified as into five main classes threshold based, Cluster based, Edge based, Region based, Watershed based segmentation.

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. Many applications of image analysis need to obtain the regions of interest before the analysis can start. Therefore, the need of an efficient segmentation method has always been there. A gray level image consists of two main features, namely region and edge. Segmentation algorithms for gray images are generally based on two basic properties of image intensity values, discontinuity and similarity. In the first category, the approach is to partition an image based on abrupt changes in intensity, such as edges in an image. The principle approaches in the second category are based on partitioning image into regions that are similar

according to a set of predefined criteria. Thresholding, region growing and region splitting and merging are examples of the methods in this category.

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. The breast contour can be obtained by partitioning the mammogram into breast and non-breast regions. The extracted breast contour should adequately model the soft-tissue/air interface and preserve the nipple in profile.

3.2.1 Binary Homogeneity Enhancement Algorithm (BHEA)

In this research paper we have proposed a new homogeneity enhancement process namely Binary Homogeneity Enhancement Algorithm (BHEA) for digital mammogram. In this process, the Mammogram image is treated as an array of pixel data. First step of the process is to determine the dimension of the image and determine the middle position of image array. We then 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. 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 in sixteen level scales. The process will be continued recursively, popping the start and end position subset array from the stack and repeat the aforesaid process. The process will be continued until the stack is empty.

The same process will be repeated by scanning the image vertically from top to bottom followed by uniform color quantization technique in color space breaking in sixteen level scales.

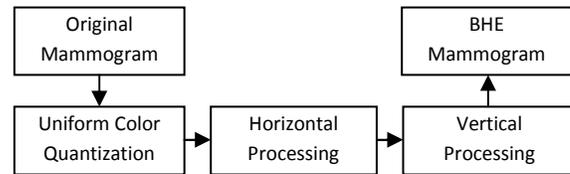


Figure 2. Process logic of BHEA

Algorithm: Binary Homogeneity Enhancement Algorithm

Input: Raw Mammogram (orgM)

Output: Horizontal Processed Mammogram (horM), Binary Homogeneity Enhanced Mammogram (resM)

Begin:

Step1. Open orgM

Step2. Create Stack

Step3. Loop row=0 to orgM.Height

col=0

Stack.push=row,col,orgM.Width

Call Procedure hCheck()

row=row+1

End Loop

Step4. Loop col=0 to horM.Width

row=0

```

Stack.push=col,row,horM.Height
Call Procedure vCheck()
col=col+1
End Loop
Step5. End

Procedure: hCheck()
Begin:
Step1. If Stack=Empty
Return
Else
row=Stack.Pop
Start=Stack.Pop
Last=Stack.Pop
If Start=Last
Call Procedure hCalc(row,Start,Last)
Call Procedure hCheck()
Else
j= Start
gval= orgM.Pixel[row,j]
Loop k=j+1 to Last
kval= orgM.Pixel[row,k]
If kval < gval-MDT OR kval > gval+MDT
midval= (Start + Last) / 2
Stack.push=row,Start,midval
Stack.push=row,midval+1,Last
Break
EndIf
k=k+1
End Loop
If k > Last
Call Procedure hCalc(row,Start,Last)
Call Procedure hCheck()
Else
Call Procedure hCheck()
EndIf
EndIf
Step2. Return

```

```

Procedure: hCalc(row,Start,Last)
Begin:
Step1. If Start <> Last
Len=0
Loop l=Start to Last
Array[Len]=(orgM.Pixel[row,l]/16)*16
Len=Len+1
l=l+1
End Loop
Loop l=Start to Last
horM.Pixel[row,l]=MODE(Array)
l=l+1
End Loop
Else
horM.Pixel[row,l]=(orgM.Pixel[row,l]/16)*16
EndIf
Step2. Return

```

```

Procedure: vCheck()
Begin:
Step1. If Stack=Empty
Return
Else

```

```

col=Stack.Pop
Start=Stack.Pop
Last=Stack.Pop
If Start=Last
Call Procedure vCalc(col,Start,Last)
Call Procedure vCheck()
Else
j= Start
gval= horM.Pixel[j,col]
Loop k=j+1 to Last
kval= horM.Pixel[k,col]
If kval < gval-MDT OR kval > gval+MDT
midval= (Start + Last) / 2
Stack.push=col,Start,midval
Stack.push= col,midval+1,Last
Break
EndIf
k=k+1
End Loop
If k > Last
Call Procedure vCalc(col,Start,Last)
Call Procedure vCheck()
Else
Call Procedure vCheck()
EndIf
EndIf
Step2. Return

Procedure: vCalc(col,Start,Last)
Begin:
Step1. If Start <> Last
Len=0
Loop l=Start to Last
Array[Len]=(horM.Pixel[l,col]/16)*16
Len=Len+1
l=l+1
End Loop
Loop l=Start,Last
resM.Pixel[l,col]=MODE(Array)
l=l+1
End Loop
Else
resM.Pixel[l,col]=(horM.Pixel[l,col]/16)*16
EndIf
Step2. Return

```

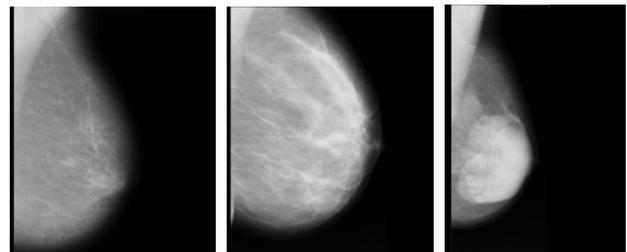


Figure 3. The Original Mammograms (Fatty, Fatty-glandular, Dense-glandular from left to right)

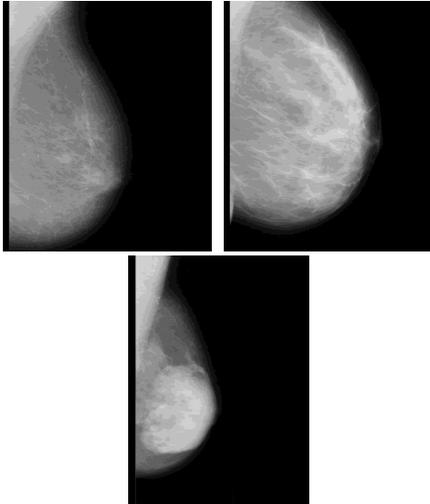


Figure 4. The Mammograms after Uniform Color Quantization and Horizontal Processing of BHE (Fatty, Fatty-glandular, Dense-glandular from left to right)

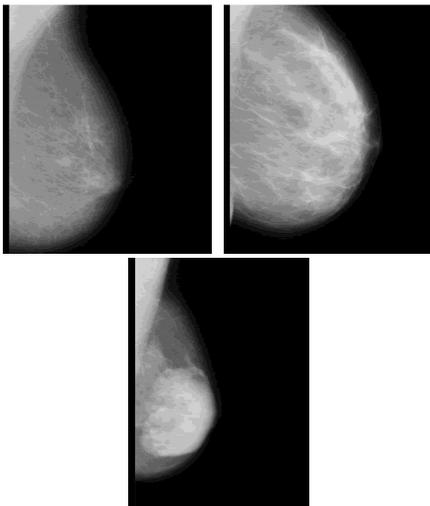


Figure 5. The Mammograms after Vertical Processing of BHE (Fatty, Fatty-glandular, Dense-glandular from left to right)

3.2.2 Edge Detection Algorithm (EDA)

After the mammogram image was processed by the proposed Binary Homogeneity Enhancement Algorithm (BHEA) we obtain an image that is homogenous and have pixels in grayscale of sixteen level scales. This image forms the input for the next level of processing for detection of the edge, i.e. breast boundary and the breast region. In this research paper we have proposed a new method for detecting edge rather than using one of the known methods of edge detection using a convolution filter.

It can be observed that all the mammogram images in the mini MIAS database, has the image of the breast at the center (left to right or right to left), and is bounded by non-breast image on either side. The mammogram image have skin-air interface on the side of the nipple where the air interface have black pixels.

Similarly, the other side behind the breast has a well-defined baseline that separates the breast image from the body and contains black pixels. We consider these features in our proposed method of edge detection. In the first step of our proposed method we start scanning pixels horizontally from left most pixel (Start Pixel) to right most pixel (Last Pixel) of the processed image. If any change of pixel intensity observed is marked by a black pixel indicating a horizontal edge point. We continue this process for all rows of pixel data of the image to obtain a Horizontal Edge Map image. In the next step, we scan the image vertically from top most pixel (Start Pixel) to the bottom most pixel (Last Pixel). Continuing the process for all the columns we obtain a Vertical Edge Map image. Finally, we merge the Horizontal Edge Map image with Vertical Edge Map image by performing a logical OR operation on the two image files, to obtain the Edge map of mammogram image.

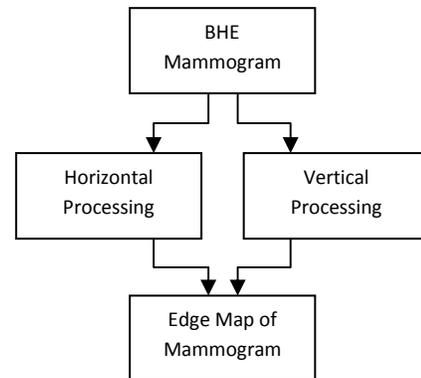


Figure 6. Process Logic to Generate Edge Map of Mammogram

Algorithm: Edge Detection Algorithm

Input: Raw Mammogram(orgM)

Output: Horizontal Edge Mammogram (horM), Vertical Edge Mammogram (verM), Edge Map of Mammogram (resM)

Begin:

Step1. Open orgM

Step2. Loop row=0 to orgM.Height

Loop col=0 to orgM.Width

val= orgM.Pixel[row,col] AND 000000ff

k= col+1

Loop k to orgM.Width

kval = orgM.Pixel[row,k] AND 000000ff

If kval <> val

Break

EndIf

k=k+1

End Loop

horM.Pixel[row,k-1]=ff000000

col=k

End Loop

row=row+1

End Loop

Step3. Loop col=0 to orgM.Width

Loop row=0 to orgM.Height

val= orgM.Pixel[row,col] AND 000000ff

```

k= row+1
Loop k to orgM.Height
    kval = orgM.Pixel[k,col] AND 000000ff
    If kval <> val
        Break
    EndIf
    k=k+1
End Loop
verM.Pixel[k-1,col]=ff000000
row=k
End Loop
col=col+1
End Loop
Step4. Loop row=0 to orgM.Height
    Loop col=0 to orgM.Width
        Bval=horM.Pixel[row,col] AND 000000ff
        Cval=verM.Pixel[row,col] AND 000000ff
        If Bval = 255 AND Cval = 255
            resM.Pixel[row,col]= ffffffff
        Else
            resM.Pixel[row,col]= ff000000
        EndIf
        col=col+1
    End Loop
    row=row+1
End Loop
Step5. End

```



Figure 7. The Edge Map of Mammograms using EDA (Fatty, Fatty-glandular, Dense-glandular from left to right)

3.2.3 Breast Boundary Detection Algorithm (BBDA)
 After the mammogram image is processed by the proposed Binary Homogeneity Enhancement Algorithm (BHEA) we obtain an image that is homogenous and have pixels in grayscale of sixteen level scales. This image forms the input for the next level of processing for detection of the edge, i.e. breast boundary and the breast region. 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. 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.

In the previous process we have already extracted the edges of the breast region from the background of mammogram image. We now proceed with the output image of the edge extraction process. We now need to identify the outermost edge line that constitutes the edge of the breast. We output the breast boundary edge line on another image. We scan the image from the right side of the image to locate the rightmost point at the topmost 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, where there are no pixels that have already been traversed, we pop out from the Backtrack stack a lesser priority pixel and continue with the traversal process. The traversal continues to the next pixel 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 on a blank image to obtain the breast boundary for further processing.

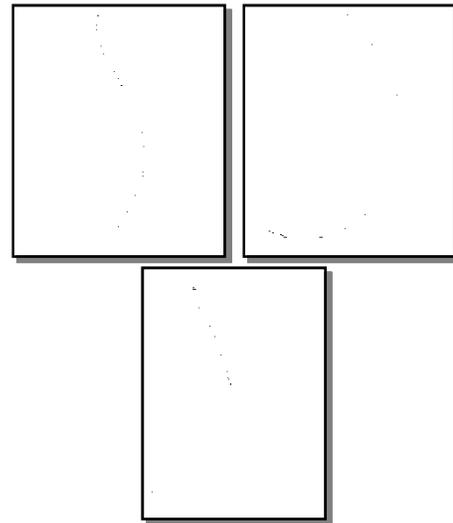


Figure 8. Breast Boundary of Mammogram using BBDA (Fatty, Fatty-glandular, Dense-glandular from left to right)

Algorithm: Breast Boundary Detection Algorithm
 Input: Edge Map Mammogram (orgM)
 Output: Breast Boundary Mammogram (resM)
 Step1. Open orgM
 Step2. Create BackTrackStack
 Step3. Create PlotList ,SmoothList
 Step4. HorzRef=0
 Step5. Loop row=0 to orgM.Height
 Loop col=0 to orgM.Width
 flag=false
 If (orgM.Pixel[row,col] AND 000000ff)=0

```

        If flag=false
            HorzRef=col
            flag=true
        EndIf
        resM.Pixel[row,col]=ff000000
    EndIf
    col=col+1
End Loop
row=row+1
End Loop
Step6. Bstart=0,listctr=0,startcol=0
Step7. Loop orgM.Width-1 to col=0
    Bstart = orgM.Pixel[0,col] AND 000000ff
    If Bstart=0
        PlotList[listctr].x=0
        PlotList[listctr].y=col
        BackTrackStack.push=0,col
        startcol=col
        Break
    EndIf
    col=col-1
End Loop
Step8. Di=0, Dj=startcol
Step9. prevrow=0, prevcol=0
Step10. Loop Di <> orgM.Height AND Dj<> HorzRef
    Create ChqStack
    If Di <> 0
        If ((orgM.Pixel[Di-1,Dj+1] AND 000000ff) <> 0) AND (prevrow <> Di-1 OR prevcol <> Dj+1)
            ChqStack.push= Di-1,Dj+1
        EndIf
        If ((orgM.Pixel[Di-1,Dj] AND 000000ff) <> 0) AND (prevrow <> Di-1 OR prevcol <> Dj)
            ChqStack.push= Di-1,Dj
        EndIf
        If ((orgM.Pixel[Di-1,Dj-1] AND 000000ff) <> 0) AND (prevrow <> Di-1 OR prevcol <> Dj-1)
            ChqStack.push= Di-1,Dj-1
        EndIf
    EndIf
EndIF
    If ((orgM.Pixel[Di,Dj-1] AND 000000ff) <> 0) AND (prevrow <> Di OR prevcol <> Dj-1)
        ChqStack.push= Di,Dj-1
    EndIf
    If ((orgM.Pixel[Di+1,Dj-1] AND 000000ff) <> 0) AND (prevrow <> Di+1 OR prevcol <> Dj-1)
        ChqStack.push= Di+1,Dj-1
    EndIf
    If ((orgM.Pixel[Di+1,Dj] AND 000000ff) <> 0) AND (prevrow <> Di+1 OR prevcol <> Dj)
        ChqStack.push= Di+1,Dj
    EndIf
    If ((orgM.Pixel[Di+1,Dj+1] AND 000000ff) <> 0) AND (prevrow <> Di+1 OR prevcol <> Dj+1)
        ChqStack.push= Di+1,Dj+1
    EndIf
    If ((orgM.Pixel[Di,Dj+1] AND 000000ff) <> 0) AND (prevrow <> Di OR prevcol <> Dj+1)
        ChqStack.push= Di,Dj+1
    EndIf

```

```

    chqBTskFlag = false
    Loop ChqStack <> Empty
        r = ChqStack.pop
        c = ChqStack.pop
        selFlag=false
        Loop ctr=0 to listCtr
            If r=PlotList[ctr].x AND c=PlotList[ctr].y
                Break
            EndIf
            ctr=ctr+1
        End Loop
        If ctr > listctr
            prevrow=r
            prevcol=c
            listctr=listctr+1
            PlotList[listctr].x=r
            PlotList[listctr].y=c
            di=r
            dj=c
            Loop chqStack <> Empty
                BackTrackStack.push=chqStack.pop
            End Loop
            selFlag=true
            chqBTskFlag=true
        EndIf
        If selFlag = true
            Break
        EndIf
    End Loop
    If chqBTskFlag = false
        di= BackTrackStack.pop
        dj= BackTrackStack.pop
    EndIf
Step11. Loop plotctr=0,listctr
    plotr=PlotList[plotctr].x
    plotc=PlotList[plotctr].y
    resM.Pixel[plotr,plotc]=ff000000
    plotctr=plotctr+1
End Loop
Step12. End

```

3.2.4 Breast Boundary Smoothing Algorithm (BBSA)

The breast boundary image we obtain after the mammogram image was processed by the proposed Binary Homogeneity Enhancement Algorithm (BHEA), followed by edge detection of the enhanced image and finally our proposed Breast Boundary Detection method, is further processed by our Breast Boundary Smoothing method to obtain the final image. The breast boundary image obtained so far is not smooth. We obtain a smooth breast boundary by taking pixels at a fixed discrete interval and joining them by drawing a simple curve between the pixel positions to obtain a smooth and enhanced breast boundary. This forms our final output image.

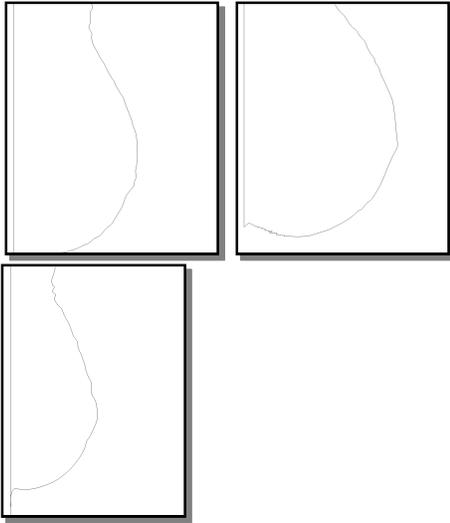


Figure 9. Smooth and Enhanced Breast Boundary (Fatty, Fatty-glandular, Dense-glandular from left to right)

4. EXPERIMENTAL RESULT ANALYSIS

4.1 Accuracy Estimation

Performance evaluation in algorithm design is an important step that is commonly neglected. Few mammogram segmentation algorithms have been tested extensively. What constitutes an “acceptable” result differs significantly, and is often based on visual subjective opinion with very little quantitative endorsement.



Figure 10. Performance indices: TP, FP, FN

The accuracy of this technique was evaluated through quantitative measures derived through the comparison of each segmented mammogram “mask” with its corresponding “gold standard”. The gold standard is generated by manually segmenting the breast region from each mammogram. A mammogram is displayed using LOG-attenuation so that the skin-air boundary can be easily identified. The boundary of the breast is then manually traced to extract the real breast region to generate a ground truth (GT) image (Figure 10).

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 breast region. Conversely, the pixels not in the GT, but in the mask are defined as false positive (FP) pixels. From this we can derive two metrics: completeness (CM) and correctness (CR). The completeness is the percentage of the GT region, which is explained by the segmented region:

$$Completeness = \frac{TP}{TP + FN}$$

Completeness can range from 0 to 1, with 0 indicating that none of the regions with properly partitioned, and one indicating that the regions were all segmented. For example, a value of 0.87 indicates an 87% overlap with the ground truth. The correctness represents the percentage of correctly extracted breast region.

$$Correctness = \frac{TP}{TP + FP}$$

Like CM, the optimum value for CR is 1. From these metrics we can characterize the segmentation results. A low value of completeness (CM) used in this context indicates an over-segmentation, whereby two or more regions in the examined segmented image represent a region in the GT. An under segmentation is similarly defined for correctness (two or more regions in the GT are represented by a single region in the segmented image). The algorithm will be considered accurate if the percentage of both completeness and correctness is greater than 95%. A more general measure of the final algorithm can be achieved by combining completeness and correctness into a single measure, with an optimum value of 1.

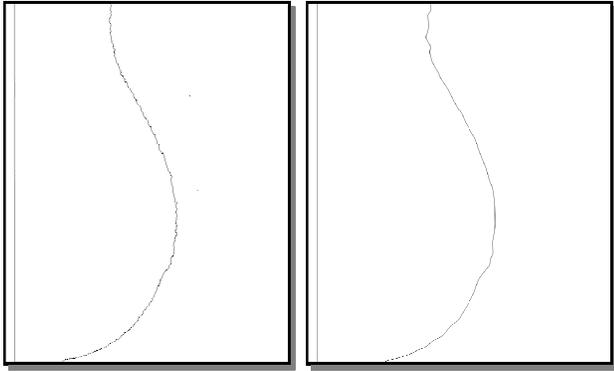
$$Quality = \frac{TP}{FN + FP + TP}$$

The pre-processing results do have some influence on the effectiveness of the algorithm, but since the artifact suppression and homogeneity accentuation algorithms are not enhancement algorithms per se, and there are no ground truth images it is non-trivial to quantitatively measure their effectiveness.

4.2 Experimental Results

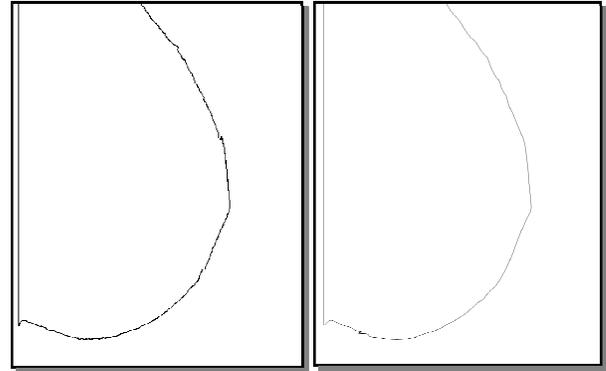
Experiment 1: Fatty Tissue

Results obtained by applying the proposed segmentation algorithm on mammogram of MIAS image 078 predominantly comprised of Fatty tissues shown in the figure 11 and figure 12 along with ground truth. The breast region contains 420,694 pixels algorithm misses 1602 pixels from the breast region and mistakes 5316 pixels from the background region as breast pixels, hence CM= 0.996, CR= 0.987 and Quality= 0.983. The quantitative measures indicate that the extracted breast region is marginally under-segmented but well within accuracy level.



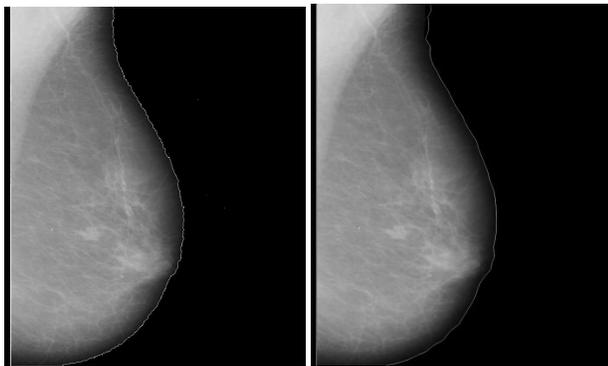
(a) (b)

Figure 11. Fatty Mammogram (078.l) (a) Ground Truth Boundary and (b) Boundary Retrieved by the Proposed Method.



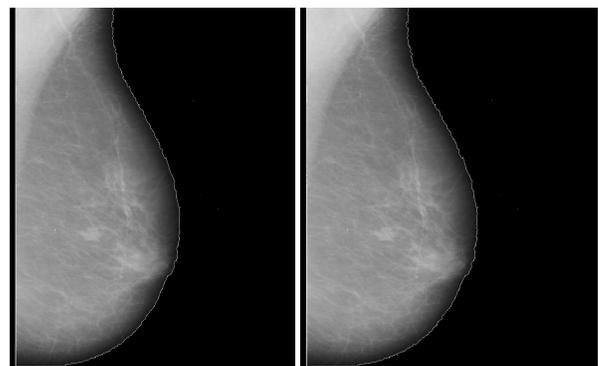
(a) (b)

Figure 13. Fatty-Fibro Grandular Mammogram (049.r) (a) Ground Truth Boundary and (b) Boundary Retrieved by the Proposed Method.



(a) (b)

Figure 12. Fatty Mammogram (078.l) (a) Ground Truth Boundary Superimposed on Original Mammogram and (b) Boundary Retrieved by the Proposed Method Superimposed on Original Mammogram.



(a) (b)

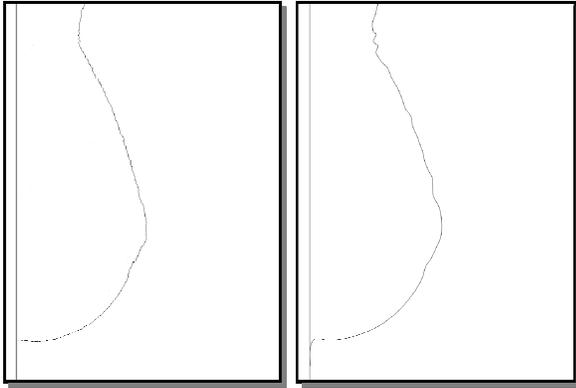
Figure 14. Fatty-Fibro Grandular Mammogram (049.r) (a) Ground Truth Boundary Superimposed on Original Mammogram and (b) Boundary Retrieved by the Proposed Method Superimposed on Original Mammogram.

Experiment 2: Fatty-Fibro Glandular Tissue

Results obtained by applying the proposed segmentation algorithm on mammogram of MIAS image 049 predominantly comprised of Fatty-Fibro Glandular tissues shown in the figure 13 and figure 14 along with ground truth. The breast region contains 500,270 pixels algorithm misses 3716 pixels from the breast region and mistakes 1202 pixels from the background region as breast pixels, hence $CM= 0.992$, $CR= 0.997$ and $Quality= 0.990$. The quantitative measures indicate that the extracted breast region is well within accuracy level.

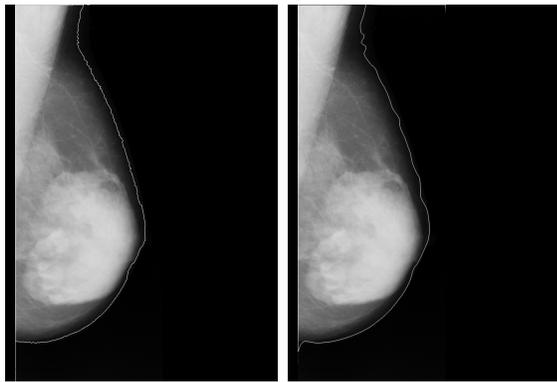
Experiment 3: Dense-Fibro Glandular Tissue

Results obtained by applying the proposed segmentation algorithm on mammogram of MIAS image 037 predominantly comprised of Dense-Fibro Glandular tissues shown in the figure 15 and figure 16 along with ground truth. The breast region contains 243,641 pixels algorithm misses 1150 pixels from the breast region and mistakes 754 pixels from the background region as breast pixels, hence $CM= 0.995$, $CR= 0.997$ and $Quality= 0.992$. The quantitative measures indicate that the extracted breast region is well within accuracy level.



(a) (b)

Figure 15. Dense-Fibro Glandular Mammogram (037.r) (a) Ground Truth Boundary and (b) Boundary Retrieved by the Proposed Method.



(a) (b)

Figure 16. Dense-Fibro Glandular Mammogram (037.r) (a) Ground Truth Boundary Superimposed on Original Mammogram and (b) Boundary Retrieved by the Proposed Method Superimposed on Original Mammogram.

4.3 Comparative Analysis

This experiment was done using all 322 mammograms taken from mini-MIAS database consisting of mammograms belong to fatty, fatty-fibro glandular and dense-fibro glandular breast tissues which give different levels of visibility due to different levels of X-ray attenuations. Usually, mammograms from breasts dominated by fatty tissues are darker and low in contrast while dense-glandular tissues are brighter as a consequence of different levels of X-ray attenuation by different types of tissues.

Mean values for completeness and correctness for the mammogram images belong to fatty, fatty-fibro glandular and dense-fibro glandular breast tissues used for the experiment were 99.4% and 98.9%, 99.3% and 99.5%, 99.6% and 99.5% respectively and overall 99.4% and 99.1% which means only 0.6% of pixels from breast region were missed and 0.9% of pixels were recognized incorrectly from the background. The average quality factor of the all mammogram image is

approximately 99.2. This shows that the segmentation algorithm accurate and maintains a higher precision level.

We have mainly compared our proposed algorithm with the Rule-Based Fuzzy Reasoning Algorithm of M. Wirth, D. Nikitenko, J. Lyon [25] partially because of two key contributions of work, the first of these is a fully-automated segmentation algorithm which in most circumstances provides a breast contour representative of the actual contour including the nipple in profile, a feat realized by very few algorithms and the second and main cause relates to the algorithm’s performance evaluation. The algorithm has been extensively tested using mammograms of differing densities from multiple databases. We have also emphasized on algorithms performance evaluation part and our results are comparable with their algorithm.

We have also compared our work with Fast-Marching Method of Roshan Dharshana Yapa, Koichi Harada [35] and Local Adaptive Thresholding Method of Maysam Shahedi B K, Rassoul Amirfattahi, Farah Torkamani Azar and Saeed Sadri [34]. These two research papers are also very systematic and algorithm’s performance evaluation part is well organized. Here also we got slightly better comparable results.

4.4 Failure Assessment

In totality algorithm’s performance out of 322 different mammograms of mini-MIAS database three to four mammograms marginally below the 95% accuracy indicator specified. Figure 17 shows the results obtained by applying our algorithm on MIAS mammogram image “mdb066”. For the mammogram image, the algorithm segmented the breast boundary with CM = 0.988, CR = 0.971 and Quality = .960, resulted slightly low completeness was due to very low contrast near the breast-edge and due to high non-uniformity on both background and breast region. But the completeness value obtained from our algorithm is comparatively higher than the value obtained by Wirth [25] and Roshan Dharshana Yapa [35] for the same image CM = 0.87 and CM = 0.944 respectively.

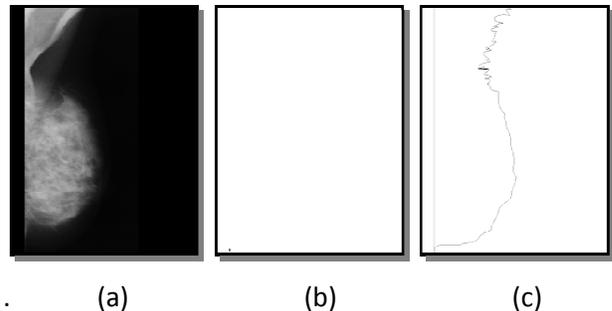


Figure 17. Mammogram 066 of mini-MIAS database (a) Original Image, (b) Ground Truth and (c) Boundary Retrieved by the Proposed Method from original image.

5. CONCLUSION

In our proposed method, we have projected four distinct steps. Each of the steps is representing a noble algorithm proposed by us. The ultimate smooth breast boundary image we obtain by the proposed Binary Homogeneity Enhancement Algorithm (BHEA), followed by Edge Detection Algorithm (EDA) and

Breast Boundary Detection Algorithm (BBDA), and finally proposed Breast Boundary Smoothing Algorithm. Results of the proposed method show a very bright reliable detection rate of breast boundary, 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 322 mammograms from mini-MIAS mammogram database. To determine the robustness of the algorithm our method has been tested on mammograms with differing breast tissue densities: fatty, fatty-fibro glandular and dense-fibro glandular and proven a comparative rate of robustness with other recognized methods. This method can be used for other medical image processing applications with minor modification that need to automatically segment an exact part of the images for post-processing stages. The last part of the method that is Breast Boundary Smoothing Algorithm can further extended in future. In conclusion, it can be mentioned, our proposed noble method is acceptably accurate, promising and comparable with any other standard methods.

6. REFERENCES

- [1] Norum J., "Breast cancer screening by mammography in Norway. Is it cost-effective?", *Annals of Oncology*, Volume 10, Issue 2, pp. 197-203, 1999.
- [2] Breast Cancer Facts & Figures, 2009-2010, American Cancer Society.
- [3] Sterns EE, "Relation between clinical and mammographic diagnosis of breast problems and the cancer/ biopsy rate," *Can J Surg*. 1996 Apr; 39(2):128-32.
- [4] Highnam R and Brady M, "Mammographic Image Analysis", Kluwer Academic Publishers, 1999. ISBN: 0-7923-5620-9.
- [5] 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
- [6] National Cancer Institute (NCI) Web site, <http://www.cancernet.gov>
- [7] Suckling J., Parker J., Dance D.R., Astley S., Hutt I., Boggis C.R.M., Ricketts I., Stamatakis E., Cernaes N., Kok S.L., Taylor P., Betal D., Savage J., "The Mammographic Image Analysis Society Digital Mammogram Database", *Proceedings of the 2nd International Workshop on Digital Mammography*, York, England, 10–12 July 1994, Elsevier Science, Amsterdam, The Netherlands, pp. 375-378.
- [8] Heath M., Bowyer K., Kopans D., Moore R., Kegelmeyer P. Jr., "The Digital Database for Screening Mammography", *Proceedings of the 5th International Workshop on Digital Mammography*, Toronto, Canada, 11-14 June 2000, Medical Physics Publishing, 2001, pp. 212-218.
- [9] U. Bick, M.L. Giger, R.A. Schmidt, R.M. Nishikawa, D.E. Wolverton, and K. Doi. "Automated Segmentation of Digitized Mammograms", *Academic Radiology*, 2(2):1-9. 1995.
- [10] F.-F. Yin, M.L. Giger, K. Doi, C.E. Metz, C.J. Vyborny, R.A. and Schmidt, "Computerized Detection of Masses in Digital Mammograms: Analysis of Bilateral Subtraction Images", *Medical Physics*, 18(5):955-963. 1991.
- [11] A.J. Mendez, P.J. Tahoces, M.J. Lado, M. Souto, J.L. Correa, and J.J. Vidal, "Automatic Detection of Breast Border and Nipple in Digital Mammograms", *Computer Methods and Programs in Biomedicine*, 49:253-262. 1996.
- [12] R. Chandrasekhar, and Y. Attikiouzel, "Automatic Breast Border Segmentation by Background Modeling and Subtraction", In *proceedings of the 5th ICGST-GVIP Journal*, Volume 5, Issue2, Jan. 2005 International Workshop on Digital Mammography, 560-565. 2000.
- [13] M.A. Wirth, and A. Stapinski, "Segmentation of the Breast Region in Mammograms using Active Contours", In *proceedings of Visual Communications and Image Processing*, 5150:1995-2006. 2003.
- [14] J.L. Semmlow, A. Shadagopappan, L.V. Ackerman, W. Hand, and F.S. Alcorn., "A Fully Automated System for Screening Xeromammograms", *Computers and Biomedical Research*. 13:350-362. 1980.
- [15] M. Masek, Y. Attikiouzel, and C.J.S. deSilva Skinair, "Interface Extraction from Mammograms Using an Automatic Local Thresholding Algorithm", In *proceedings of the 15th Biennial International Conference Biosignal*, 204-206. 2000.
- [16] M. Abdel-Mottaleb, C.S. Carman, C.R. Hill, and S.Vafai, "Locating the Boundary between the Breast Skin Edge and the Background in Digitized Mammograms", In *proceedings of the 3rd International Workshop on Digital Mammography*, 467-470. 1996.
- [17] S.L. Lou, H.D. Lin, K.P. Lin, and D. Hoogstrate, "Automatic Breast Region Extraction from Digital Mammograms for PACS and Telemammography applications", *Computerized Medical Imaging and Graphics*, 24:205-220. 2000.
- [18] K.J. McLoughlin, and P.J. Bones, "Location of the Breast-air Boundary for a Digital Mammogram Image", In *proceedings of Image and Vision Computing* . 2000.
- [19] J. Liang, T. McInerney, and D. Terzopoulos, "United Snakes", In *proceedings of the IEEE International Conference on Computer Vision*, 933-940. 1999.
- [20] T. Ojala, J. Liang, J. Näppi, and O. Nevalainen, "Interactive Segmentation of the Breast Region from Digitized Mammograms with United Snakes", *Technical Report 315*. University of Turku, Turku, Finland. 1999.
- [21] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active Contour Models", *International Journal of Computer Vision*, 1:321-331. 1998.
- [22] T. Ojala, J. Näppi, and O. Nevalainen, "Accurate Segmentation of the Breast Region from Digitized Mammograms. Computerized", *Medical Imaging and Graphics*, 25(1):47-59. 2001.

- [23] D.J. Williams, and M. Shah, “A Fast Algorithm for Active Contours and Curvature Estimation”, *CVGIP: Image Understanding*, 55(1):14-26. 92.
- [24] R. Chandrasekhar, and Y. Attikiouzel, “Gross Segmentation of Mammograms using a Polynomial Model”, In proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society, 1056-1058. 1996.
- [25] M. Wirth, D. Nikitenko, J.Lyon, “Segmentation of the Breast Region in Mammograms using a Rule-Based Fuzzy Reasoning Algorithm”, *ICGST- GVIP Journal*, 5(2): 45-54. 2005.
- [26] M.Masek, Electronic University of Western Australia, School of Electrical, Computer Engineering and University of Western Australia Centre for Information Processing Systems. “Hierarchical Segmentation of Mammograms Based on Pixel Intensity”. PhD thesis, 2004.
- [27] A.D.Brink and N.E.Pendock, “Minimum Cross-entropy Threshold Selection, *Pattern Recognition*”, 29(1): 179-188, 1996.
- [28] Gonzalez, R.C., Woods, R.E. (1992), “Digital image processing”.
- [29] S.M. Pizer, “Psychovisual issues in the display of medical images”, Hoehne KH (ed), *Pictorial Information systems in Medicine*, Berlin, Germany, Springer-Verlag, 1985, PP 211-234.
- [30] Pisano ED, et al, “Contrast Limited Adaptive Histogram Equalization Image Processing to Improve the Detection of Simulated Spiculations in Dense Mammograms”, *Journal of Digital Imaging*, Publisher Springer New York, Issue 11(4): 193-200, 1998.
- [31] Wanga X, Wong BS, Guan TC, “Image enhancement for radiography inspection”, *International Conference on Experimental Mechanics*, 2004, pp 462-468.
- [32] Ball JE, “Digital mammogram spiculated mass detection and spicule segmentation using level sets”, *Proceedings of the 29th Annual International Conference of the IEEE EMBS*, 2007, pp. 4979-4984.
- [33] Antonis Daskalakis, et al, “An Efficient CLAHE-Based, Spot-Adaptive, Image Segmentation Technique for Improving Microarray Genes' Quantification”, 2nd International Conference on Experiments/Process/System Modelling/Simulation & Optimization, Athens, 4-7 July, 2007.
- [34] Maysam Shahedi B K, Rassoul Amirfattahi, Farah Torkamani Azar and Saeed Sadri, “Accurate Breast Region Detection in Digital Mammograms using a Local Adaptive Thresholding Method”, *Eight International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS '07)*, Santorini, Greece, pp 26-29.
- [35] Roshan Dharshana Yapa, Koichi Harada, “Breast Skin-Line Estimation and Breast Segmentation in Mammograms using Fast-Marching Method”, *International Journal of Biological and Life Sciences* 3:1 2007, pp 54-62.