

# Breast Skin-Line Estimation and Breast Segmentation in Mammograms using Fast-Marching Method

Roshan Dharshana Yapa, Koichi Harada

**Abstract**—Breast skin-line estimation and breast segmentation is an important pre-process in mammogram image processing and computer-aided diagnosis of breast cancer. Limiting the area to be processed into a specific target region in an image would increase the accuracy and efficiency of processing algorithms. In this paper we are presenting a new algorithm for estimating skin-line and breast segmentation using fast marching algorithm. Fast marching is a partial-differential equation based numerical technique to track evolution of interfaces. We have introduced some modifications to the traditional fast marching method, specifically to improve the accuracy of skin-line estimation and breast tissue segmentation. Proposed modifications ensure that the evolving front stops near the desired boundary. We have evaluated the performance of the algorithm by using 100 mammogram images taken from mini-MIAS database. The results obtained from the experimental evaluation indicate that this algorithm explains 98.6% of the ground truth breast region and accuracy of the segmentation is 99.1%. Also this algorithm is capable of partially-extracting nipple when it is available in the profile.

**Keywords**—Mammogram, Fast Marching Method, Mathematical Morphology

## I. INTRODUCTION

ACCORDING to the world health organization, more than 1.2 million women were diagnosed with breast cancer and more than 700,000 women die of breast cancer each year worldwide. It is the second leading (after lung cancer) cause of cancer deaths in women today. It is also estimated that eight to twelve percent of women will develop breast cancer in their life time.

Early detection of breast cancer can be 100% successfully treated. Hence every effort should be taken to detect breast cancers in earlier stages. Therefore many computer vision techniques applied to analysis of digital mammograms have been proposed. These techniques can be divided into two stages as pre-processing stage and computer-aided detection or computer-aided diagnosis stage. In pre-processing stage

image is splitted into interested regions such as breast tissue, pectoral muscle, markers and background area. Among tasks that belong to pre-processing stage, accurate skin-line estimation is an important pre-requisite for enhancement and analysis of digital mammograms for computer aided diagnosis of breast cancer. Appropriate display of benign and malignant lesions inside the breast region in mammogram has a direct relationship to the skin line of the breast. Failure to detect breast skin-line accurately, would lead to overlook a lesion located near the skin line. Skin-line region in mammograms is normally very low in gray-level contrast. This is due to the fact that breast tissue in the skin-line zone is less dense compared to the other neighbouring tissues and due to the noise in low contrast mammograms.

This paper attempts to introduce a breast segmentation and breast skin-line detection using a combination of an improved fast-marching method and mathematical morphological operators such as area morphology, alternating sequential filter (ASF), openings and closings. Section two presents an outline of traditional fast marching method. In section three we discuss improvements introduced to traditional fast marching methods to achieve desired result. Section four of this paper discusses some results obtained in combination of this improved fast marching method and mathematical morphological operators. Finally, the conclusions are drawn in section five.

## II. CHALLENGES ON ACCURATE BREAST SKIN-LINE ESTIMATION AND SEGMENTATION

Breast skin-line estimation and breast tissue segmentation is a challenging task in many aspects. Intuitively, the nature of mammogram (X-ray) itself has limitations of presenting accurate information due to 3D to 2D projection of information. Each pixel on mammogram represents two or more tissues superimposed on each other; Intensity of each pixel is proportional to x-ray components remained after attenuation by different tissues in breast. Superimposition makes it difficult to differentiate between different regions. Also there is a region of decreasing contrast near the breast edge where the breast tapers off, caused by lack of uniform compression of breast, near the breast edge region [8]. This effect decreases the visibility along the peripheral region of the mammogram and makes it difficult to preserve the breast skin-line and identify the nipple position. Figure 1 shows an

Manuscript received November 19, 2007.

Roshan Dharshana Yapa is a Graduate Student in Department of Information Engineering of the Graduate School of Engineering in Hiroshima University, Japan. (email: dharshana@hiroshima-u.ac.jp)

Koichi Harada is a Professor in the Department of Information Engineering of the Graduate School of Engineering in Hiroshima University, Japan. (email: hrd@hiroshima-u.ac.jp).

intensity profile of a mammogram along the line drawn on the breast boundary region.

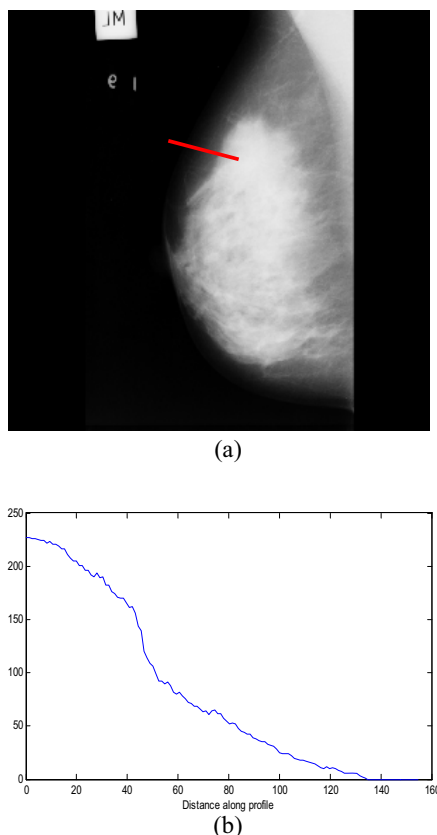


Fig. 1 Intensity profile along a breast edge region

The Intensity profile itself, doesn't provide useful information to distinct the breast boundary from its background. Moreover, there are other challenges such as noise present in low contrast mammograms. The noise can be derived from mammogram image acquisition process. Figure 2(a) shows a mammogram image and its Log-attenuated image showing noise exist around breast-edge region. Further, accurate skin-line estimation is hindered by external structures such as makers for patient identification and other related information (figure 2(b)).

### III. FAST MARCHING METHOD

Extracting shapes accurately from 2D and 3D medical images becomes an important task. Many researchers have used deformable models on these tasks [1], [2]. But it has been shown that these models have severe limitations such as they are unable to handle complex geometry and topological

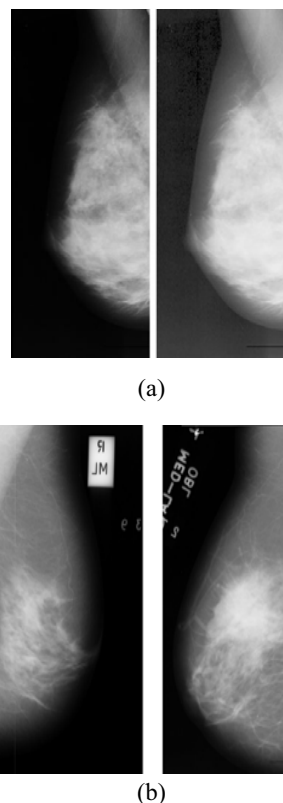


Fig. 2 (a) A mammogram image and its power-transformed image showing noise around the breast tissue. (b) Markers placed near the breast region would give fault segmentation and fault skin-line estimation.

changes without additional supporting techniques. To overcome these difficulties an alternative implicit surface evolution models have been introduced in Melladi et al. [3] and Caselles et al. [4]. Those curves and surface models evolve under image dependent implicit speed function. This method is known as Fast Marching method. In this paper we mainly concern about 2D representation of fast marching method. As explained by Sethian [5] the fast marching method can be briefed as follows.

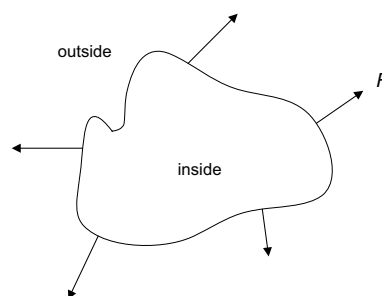


Fig. 3 A boundary (closed curve) that separating one region from another. This curve propagates with speed  $F$  in normal direction.

Consider a boundary (can be a curve in 2D or surface in 3D), moving to a direction normal to (normal direction is oriented with respect to an inside and an outside) itself with a known speed function  $F$  (figure 3). The speed function depends on local image qualities such as curvature, normal direction etc., global qualities such as shape of the curve, position of the front etc., and independent properties such as fluid velocity that passively transport the front. This method exploits a strong link between computational fluid equations in physics. The main idea here is to develop numerical methods to track the motion of this moving surface (boundary) by exploiting the link with computational fluid equations. The central idea here is to represent a curve as zero level set of a higher dimensional function  $\phi$  (figure 4); the motion of the curve is then embedded within the higher dimensional surface. This approach gives several advantages. First, although the higher dimensional function remains as a function, the zero level set can change topology and form sharp corners. Second, a discrete grid can be used together with finite differences to devise a numerical scheme to approximate the solution. Third, intrinsic geometric quantities like normal and the curvature of the curve can be easily extracted from the higher dimensional function. Generally fast marching method for boundary detection starts from a seed point or an initial region which is given manually or estimated from other detected boundaries. Let  $\Gamma$  be this initial position of the hyper-surface and let  $F$  be the speed of this evolving surface in the normal direction. As explained above, this evolving front can be considered as the zero level set ( $\phi(x,y,t)=0$ ) of a time-dependent level set function  $\phi(x,y,t)$ .

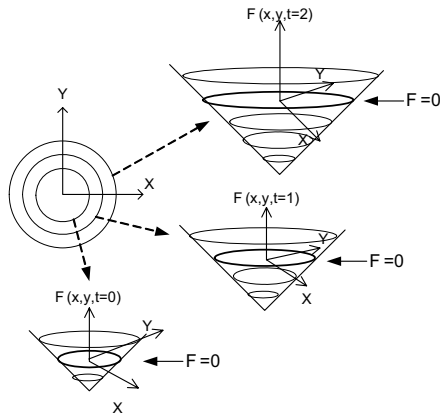


Fig. 4 Boundary at any given time will be represented by a zero level set of a higher dimensional function.

Then, by the chain rule, an evolution equation for the moving hyper-surface can be represented by;

$$\phi_t + F(x,y,z)|\nabla\phi| = 0 \quad (1)$$

Consider the special case of a surface moving with speed  $F(x,y) > 0$ . Now, let  $T(x,y)$  be the time at which the surface crosses a given point  $(x,y)$ . The function  $T(x,y)$  then satisfies the equation;

$$|\nabla T|F = 1 \quad (2)$$

According to above two equations, there are two ways of approximating the position of the moving surface. One method is numerically approximating the derivatives in equation 1 by iteration towards the solution. The other method is explicitly solving the equation 2 for  $T(x,y)$ . Fast marching algorithm that we discuss on this paper relies on the latter approach.

As we focus on 2D digital mammograms we limit our concern to two-dimensional problem. Using two-dimensional approximation the equation 2, we are looking for a solution to the following quadratic equation:

$$\left[ \max(D_{i,j}^- T, D_{i,j}^+ T, 0)^2 + \max(D_{i,j}^- T, D_{i,j}^+ T, 0)^2 \right]^{1/2} = 1/F_{i,j} \quad (3)$$

where  $D^-$  and  $D^+$  are backward and forward difference operators.

A challenge in fast marching interface problems is to produce adequate model for the speed function  $F$ . Main goal here is to extract the skin-line boundary of a given mammogram image which is represented by the boundary of the evolving front. Hence the evolving front should be forced to stop in the vicinity of the desired object boundaries under the influence of an image-based halting criterion. For a given image with intensity values  $I(x,y)$ , often this speed function  $F(x,y)$  is defined as a decreasing function based on the image local gradient  $|\nabla I(x,y)|$ . Frequently used main forms of  $F(x,y)$  are:

$$F(x,y) = \frac{1}{1 + \alpha |\nabla(G_\sigma * I(x,y))|}, \quad \alpha > 0 \quad (4)$$

and

$$F(x,y) = e^{-\alpha |\nabla(G_\sigma * I(x,y))|}, \quad \alpha > 0 \quad (5)$$

where  $\alpha$  is a constant and the expression  $G_\sigma * I$  denotes the image convoluted with a Gaussian smoothing filter whose standard deviation is  $\sigma$ . The term  $|\nabla(G_\sigma * I(x,y))|$  is essentially zero except where the image gradient changes rapidly, in which case the value becomes large. Thus the speed  $F(x,y)$  is close to unity away from the boundary and drops to zero near the sharp changes in the image gradient.

Thus the fast marching algorithm can be briefed as follows. A small front (typically a single seed point) is initialized inside the desired region, grows outwards and stops at the sharp boundary as the speed function  $F$  reduces to near zero.

Sethian [5] has proved this algorithm works extremely fast for 2D and 3D images and its time complexity is  $O(N \log N)$  where  $N$  is the total number of points in the domain. For further information regarding traditional fast marching algorithm see ref [5].

#### IV. TRADITIONAL FAST MARCHING ON MAMMOGRAMS

In this section we discuss a way of applying fast marching algorithm on gray-scale digital mammograms for segmenting breast region from its background and to estimate position of the skin-line. Initially, our segmentation algorithm was developed based on the traditional fast marching method. Intensity gradient information (equation 5) is used as the

stopping criteria for the evolving front. An original image and the corresponding resulting images are given in figure 5.

Generally gray-scale mammograms are very low in contrast. Specially, near the skin-line region contrast falls extremely low and hence it is very hard to distinguish boundary points separating breast tissue area and background. This is attributed to the facts that the breast tissue in the skin-line zone is less dense compared to other tissues in the neighbourhood and lack of uniform compression over the breast tissue by the mammogram image acquisition system. Another fact that, stroma edge area (inside margin of the skin-line) and some other internal regions takes comparatively higher intensity gradient values than the skin-line region. Hence the traditional fast marching algorithm stops evolution of the front on some internal, higher intensity gradient points while evolving the front faster on other lower gradient regions. This causes a “boundary leak” problem as shown in figure 5(b). Thus the intensity gradient information itself is not adequate to accurately estimate the skin-line.

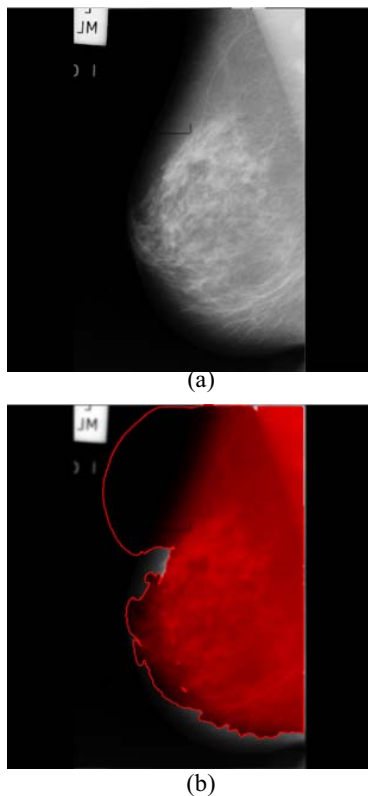


Fig. 5 Application of traditional fast marching algorithm on a mammogram. (a). Image of a mammogram. (b). “Boundary leaking” problem.

#### V. MODIFICATION OF FAST MARCHING METHOD ON MAMMOGRAMS

As we discussed above, traditional fast marching method is based on the local image gradient information. Due to the lower contrast in mammogram images, local gradient information itself is not adequate to get an accurate result for skin-line estimation. Additional image features must be

incorporated with the intensity gradient based speed function  $F$ . Assuming the Gaussian distribution for intensity values inside the breast region, we calculated intensity likelihood between breast region and each pixel in the image as:

$$P(x, y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(I(x, y) - \mu)^2}{2\sigma^2}\right) \quad (6)$$

where  $I(x, y)$  is the intensity values of each pixel at  $(x, y)$ .  $\mu$  and  $\sigma$  are mean and standard deviation of pixels inside the breast region, respectively.  $\mu$  and  $\sigma$  can be estimated using pixels inside the initial seed region defined for fast marching method.

According to the Gaussian distribution, pixels with intensity values closer to the mean value will give higher likelihood to be belonging to breast region. Thus the calculated likelihood value can be used to distinct pixels between region of interest and background. This likelihood function is combined with intensity gradient based speed function to form a new speed function as:

$$F_{new}(x, y) = F(x, y) * P(x, y) \quad (7)$$

Now, the speed function in equation 7 is based on both intensity likelihood value and intensity gradient information and hence, we can expect to have accurate estimation of skin-line by stopping the evolving front within the expected region.

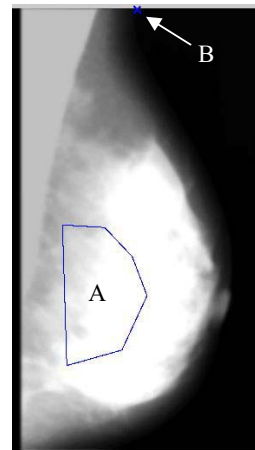


Fig. 6 Seed region (A) and end-point selection (B)

Further, according to the traditional fast marching method, the evolving front is expected to stop at intended image boundaries based on the speed function. But, it does not work well in low contrast, noisy images like mammograms. Thus, we introduced another constraint to the fast marching method to ensure that the evolving front stops near the desired boundary. For this we set an “end point” on the desired boundary line (Figure 6). When the evolving front reached this end point it will stop moving. Above mentioned two techniques can be used to obtain an accurate result on breast segmentation and estimation of breast skin-line.

## VI. PROPOSED SEGMENTATION AND SKIN-LINE ESTIMATION ALGORITHM

The proposed algorithm was specifically developed for segmenting breast tissue area from the background and accurately estimating the breast skin-line. This algorithm mainly consists with three steps as pre-processing, segmentation and skin-line estimation and finally, post-processing. Pre-processing step is used to reduce noise and remove external articles such as identification labels, markers and wedges from the mammogram and to improve area homogeneity. Gaussian smoothing algorithm and Attribute morphological operators [9], [10] are used in this case. Then the fast-marching algorithm proposed above is applied on mammograms to obtain breast tissue segmentation and skin-line estimation. Finally, morphological processing is used to remove spurious pixels from the skin-line and obtained a smoothed estimate.

### A. Pre-Processing

External articles, specially, placed near the breast boundary will decrease the accuracy of breast skin-line estimation and segmentation results. Removing these external articles from mammograms will increase the precision of final estimates. Attribute morphological based article suppression algorithm [11] was used to remove those external articles from the background of mammograms. Attribute morphology doesn't need a pre-defined structuring element such as disk, square, hexagonal etc., as in standard morphological operators. Morphological filters using standard structuring elements always do not guarantee that it will reconstruct the original structure to be preserved, as it is. Attribute morphology will be a good solution for problems that we need to remove unwanted articles on an image without damaging the shape of the original object. Attribute morphological operators, for example, area-opening operator can be used to remove such articles less than a given size. It can be seen as transformations with a structuring element which locally adopts its shape to image structures, and therefore has nice filtering capabilities while preserving original shapes.

After removing external articles from a mammogram image, again we used another attribute morphological operator called *Alternating Sequential Filter (ASF)* (area closing followed by area opening). In mathematical morphology, connected regions of constant intensity are known as "Flat Zones" [12]. *ASF* is used to improve the homogeneity among neighbouring regions on the mammogram by increasing the area of flat zones while reducing the total number of flat zones in the image. This filter is very good in filtering dark and bright impulse noises equally well while preserving desired image structures [12]. This improves the precision of final estimates and robustness of the algorithm. An example of applying these pre-processing algorithms on mammograms is shown in figure 7.

### B. Breast Segmentation and Skin-Line Estimation

Then the proposed Fast-Marching algorithm is applied on pre-processed mammogram images. As explained in section 5, before execute the fast-marching algorithm on a mammogram,

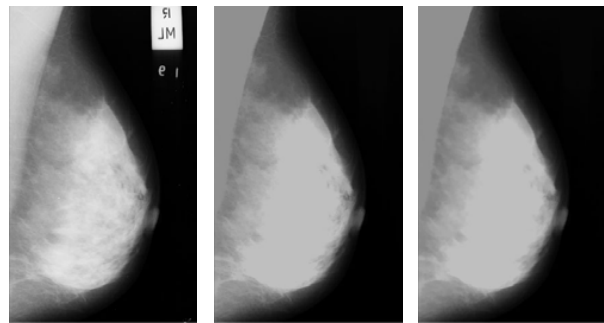


Fig. 7 (a) Original Mammogram with identification labels. (b) After applying attribute morphological article suppression algorithm. (c) After applying *Alternating Sequential Filter (ASF)*.

user needs to define an initial seed region inside the breast tissue and an end point on the breast skin-line. Then the algorithm starts to grow the initial region until it reaches the desired boundary. Figure 8(a) shows a segmented breast tissue region and figure 8(b) shows the estimated breast skin-line superimposed on the corresponding original mammogram image.

Since the contrast is very low near the breast –edge region, the speed function used here in the fast-marching algorithm must be highly sensitive to intensity gradient values near the breast-boundary region. Hence, the equation 5 is used as the intensity gradient component of the speed function. Parameter  $\alpha$  of the intensity gradient component is considered as 1 throughout this experiment. Then the parameters of the speed function  $\mu$  and  $\sigma$  are calculated using sample intensity values obtained from the initial seed region.

Another important fact is that, selecting the "end-point" on the breast skin-line. Since the noise is high in upper and lower parts of the mammograms, these regions are likely to boundary-leaking. Hence selecting the end-point either on top-outer edge or bottom-outer edge of the breast will increase the precision of final estimates.

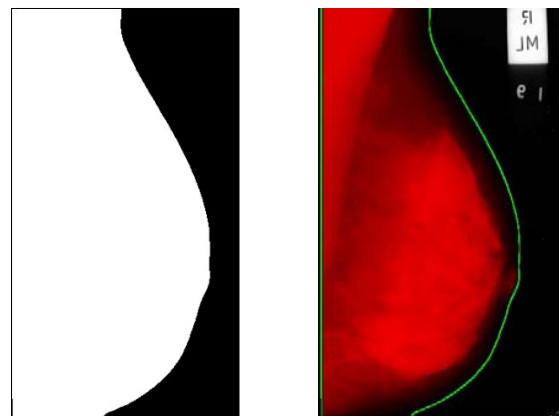


Fig. 8 (a) Segmented breast tissue region. (b) Breast contour superimposed on Original mammogram image. Red colour region represents segmented breast tissue area.

### C. Post-Processing

After obtaining estimation for breast skin-line, morphological filter (closing followed by opening operator with a “disk” shaped structuring element [10]) is applied to suppress inconsistencies on the estimated skin-line. Generally, lower and upper area of mammograms contains noise besides the low contrast at breast-edge. This could sometimes introduce undesirable artifacts on lower and upper area of estimated skin-line and segmented breast region. Standard morphological closing-opening (*ASF*) operator is used in this case to remove small false-positive and false-negative artifacts from the segmented mammogram image. Figure 9 shows an example of removing a false-positive region by using morphological *ASF* operator. As a summary, flow-chart of the proposed algorithm is shown in figure 10.

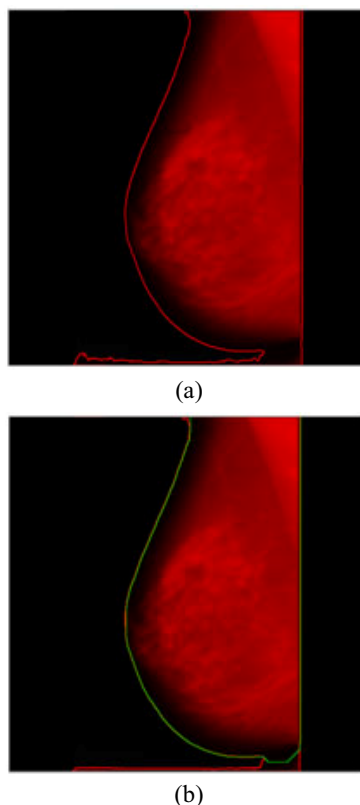


Fig. 9 (a) Estimated skin-line after applying fast-marching algorithm (red line). (b) Smoothed skin-line estimation (green line) after applying morphological *ASF* operator.

## VII. PERFORMANCE EVALUATION

In the literature, breast segmentation algorithms can be categorized into four methods as histogram based methods [13] [14] [15] [16], gradient based methods [17] [18] [19], polynomial modelling [20], active contour based methods [21] and classifiers [8]. According to the literature, there is very few segmentation algorithms have been tested extensively. Among those algorithms, Abdel-Mottaleb et. al. [17] have tested their algorithm on 500 mammograms and achieved

“acceptable” boundary in 98% of the images. Mandez et. al. [18] have tested their contour based algorithm on 156 mammograms and claim that breast contour is “accurate” or “nearly accurate” in 89% of the mammogram images. Bick et.al. have test their algorithm on 740 mammogram images and they have visually rated that their algorithm is accepted on 97% of mammogram images. Also Chandrasekhar et. al. [20] claim that their algorithm estimates breast-skin line with 94% success on mammogram images in MIAS [7] database. But all of these algorithms have been evaluated in a qualitative approach based on comments made by radiologists whether, each of the skin-line estimates is visually accepted or not. They don’t provide any quantitative measure on statistically sound way. Meanwhile, Wirth [8] has proposed a quantitative performance evaluation criterion. According to his criterion he claims that overall success of his algorithm is 98%. In our research, we too used the same quantitative evaluation criterion to measure the performance of our proposed algorithm.

Wirth’s [8] evaluation method is based on comparison of segmented mammogram image with a corresponding ground-truth image. Then it generates quantitative statistics based on comparison of the candidate image and the ground-truth image. In this research we used the same method to evaluate performance of the proposed algorithm. Ground-truth (GT) images of breast boundary can be extracted manually by using Log-attenuated images of mammograms. Then the quantitative measures are derived to evaluate the precision of the segmentation of the algorithm. The region extracted by the segmentation algorithm (mask) which matches the GT is denoted by true positive (TP). This emphasise the region that the algorithms has extracted from the ground truth region. The region shown in the mask but not shown in the GT is denoted as false positive (FP) pixels which represent the misleadingly classified region by the algorithm. Also the region shown in GT but not in mask is defined as false negative (FN) classifications which represent missing pixels in the breast region by the segmentation algorithm. From these measures, two performance evaluation metrics can be defined as *completeness* (CM) and *correctness* (CR). The *completeness* can be defined as the percentage of the true breast region (GT) extracted by the segmentation algorithm.

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

Values of CM can be ranged from 0 to 1. The value 0 represents that none of the pixels belongs to GT region is properly segmented by the algorithm, and 1 represent that all the pixels belongs to the GT region are segmented. The *correctness* can be defined as the percentage of correctly extracted breast region.

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

Same as for the CM, values of CR also ranged from 0 to 1. The value 1 implies that the region extracted by the segmentation algorithm is 100% correct. CR values less than 1, denotes that the algorithm has under-segmented the breast region by (100-CR) %.



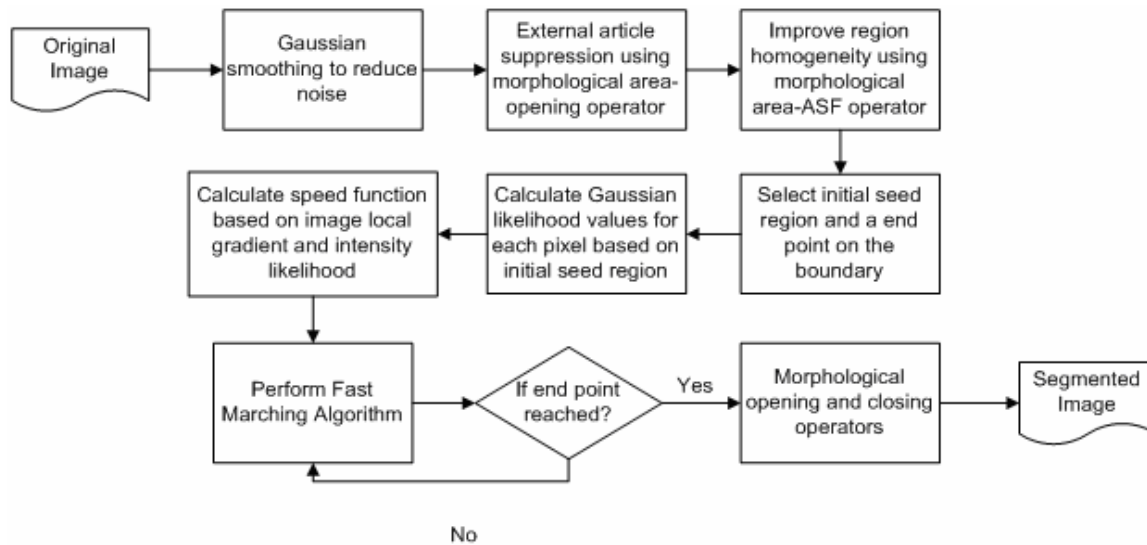


Fig. 10 Flow-chart of the proposed algorithm

In the context of CM, values less than 1 denotes that the algorithm has over-segmented the breast region by  $(100 - \text{CM})\%$ . According to Wirth [8], an algorithm can be accepted as accurate if the percentage of both CM and CR are greater than 95%. Further, an overall measurement of the algorithm can be obtained by combining both measurements CM and CR together in to a single measure with an optimum value 1.

$$\text{quality} = \frac{TP}{TP + FP + FN}$$

### VIII. EXPERIMENT RESULTS

This experiment was done using 100 mammograms taken from mini-MIAS database [7]. Mini-MIAS database is a public database of mammogram images. This is a popular database among mammogram research community. It contains 320 mammograms digitized at 200 micron pixel edge. Also these images have been clipped or padded so that every image consists of 1024x1024 pixels.

The 100 images used for the experiment, consisted mammograms belong to fatty, fatty-glandular, and dense-glandular breast tissues which give different levels of visibility due to different levels of X-ray attenuations [6]. Among the 100 mammogram images used for the experiment, 28, 48 and 24 images were from fatty, fatty-glandular and dense-glandular tissue types respectively. 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 all the 100 mammogram images used for the experiment were 98.6% and 99.1% respectively which means only 1.4% of pixels from breast region were missed and 0.9% of pixels were recognized incorrectly from the background. Corresponding standard deviations were 1.22% and 0.59% respectively. This shows that the segmentation algorithm maintains a higher precision level. Also the algorithm is highly robust with

respect to different breast density types. Further the table 1 shows mean value of *completeness* and *correctness* measures for each type of tissue densities. By combining CM and CR measures, our proposed algorithm gives an average quality measure of 97.9%.

Tissue Type	Mean CM	Mean CR
Fatty	0.988	0.991
Fatty-Glandular	0.989	0.991
Dense-Glandular	0.981	0.990

Results obtained by applying the proposed segmentation algorithm on Fatty, Fatty-Glandular and Dense-Glandular images are shown in figures 11, 12 and 13 respectively. Figure 11 shows the result obtained for MIAS image “mdb070” with CM = 0.996 and CR=0.999 which means that the segmented image describes 99.6% of GT and it has correctly extracted 99.9% from GT region.

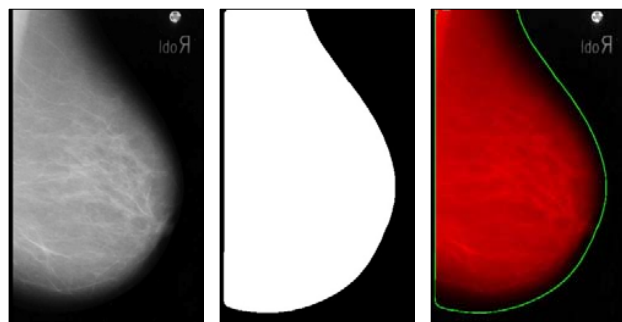


Fig. 11 Segmentation of Fatty mammogram “mdb070”. (a) Original Image (b) Segmented mask (c) Estimated skin-line and segmented region superimposed on original image.

Figure 12 shows results obtained by applying the algorithm on Fatty-Glandular mammogram “mdb016”. On this

experiment, the breast region was extracted with  $CM = 0.994$  and  $CR = 0.987$ .

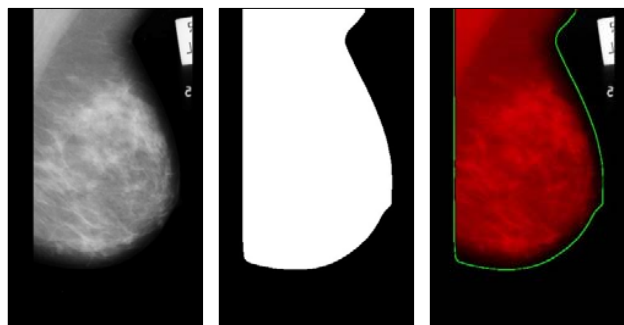


Fig. 12 Segmentation of Fatty-Glandular mammogram "mdb016". (a) Original Image. (b) Segmented mask. (c) Estimated skin-line and segmented region superimposed on the original image.

Finally the figure 13 shows results obtained from a mammogram composed of Dense-Glandular tissue, "mdb003". On this experiment, breast region was extracted with  $CM = 0.994$  and  $CR = 0.993$ .

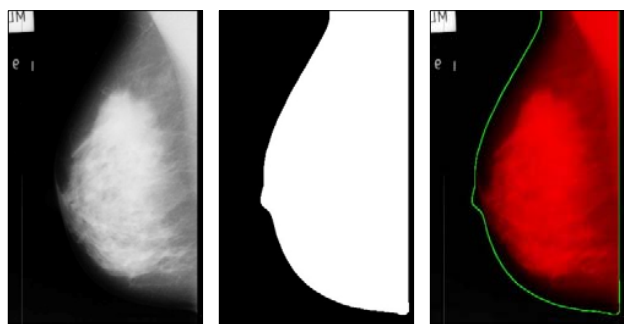


Fig. 13 Segmentation of Fatty-Glandular mammogram "mdb003". (a) Original Image. (b) Segmented mask. (c) Estimated skin-line and segmented region superimposed on the original image.

This algorithm is also capable of partially-extracting nipple from mammograms when it exists in the profile. This can be demonstrated by comparing the results shown in figures 12 and 13 with corresponding GT images, figure 14(a) and 14(b) respectively.



Fig. 14 (a) GT image of "mdb016". (b) GT image of "mdb004".

But, two mammograms among 100 mammograms processed in this experiment were marginally below the 95% accuracy indicator specified. Figure 15 shows the results obtained by applying the algorithm on MIAS mammogram image "mdb066". For the given image, the algorithm segmented the breast region with  $CM = 0.944$  and  $CR = 0.995$  which implies that the segmented region entirely contained within the GT region. Resulted low *completeness* was due to very low contrast near the breast-edge and due to high non-uniformity on both background and breast region. This non-uniformity can be happened due to problems with the image acquisition such as dusts induced by the scanner, background noise and dust artifacts. Figure 15(c) shows the Log-attenuated image of the original mammogram 15(a) which demonstrates non-uniformity, low contrast and excessive noise exist in the background of the image. But the *completeness* value obtained from our algorithm is comparatively higher than the value obtained by Wirth [8], for the same image ( $CM = 0.87$ , and  $CR = 0.99$ ).

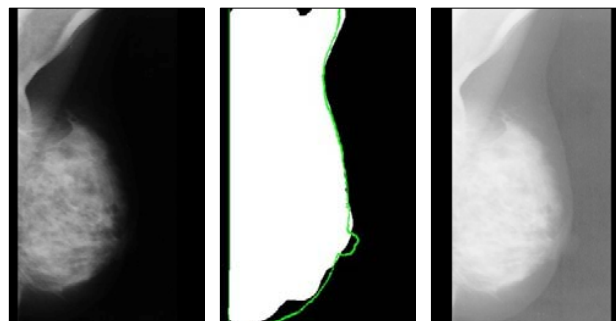


Fig. 15 False segmentation. (a) Original Image. (b) Segmented region superimposed with contour of the ground-truth mask. (c). Log-attenuated image showing noisy background.

## IX. CONCLUSIONS AND FUTURE WORK

Accurate segmentation of breast tissue region is a very important pre-processing step in digitized mammogram analysis for diagnosing abnormalities in breast tissue. But this is not an easy task due to inherent low contrast and extensive noise exists in mammograms. This segmentation process is further interfered by external articles such as information labels, markers, wedges and articles induced by the digitization process.

Combination of intensity information with gradient information on fast marching speed function and introducing end-point constraint ensures that the boundary expands within the intended region and stops when the boundary reached the end-point. By analyzing the results obtained from this experiment and characteristics of fast marching and level set methods, we can conclude that in most of circumstances this algorithm estimates breast skin-line which representative of the ground truth. Also this algorithm is capable of extracting nipple approximately from the mammogram when it is available in the profile. From the experimental results we have convinced that this algorithm is robust with respect to different breast tissue densities.



One negative point of this algorithm is that this needs user involvement for selecting an initial seed region and an end point on the breast edge. We are further working on eliminating this constrain from this algorithm and to develop a fully automated breast segmentation and skin-line estimation algorithm. In future we will evaluate the performance of this algorithm on all the images in MIAS database and compare the performance of this algorithm based on digital mammogram images in other available mammogram databases.

## REFERENCES

- [1] Kass M, Witkin A, Terzopoulos D., Snakes: Active Contour Models., Int. J. Computer Vision, 1987; 1(4), pp 321-31.
- [2] McInemey T, Terzopoulos D., *Deformable models in medical image analysis: a survey*, Med. Img. Anal, 1996; 1(2), pp 91-108.
- [3] R. Malladi, J. A. Sethian, and B.C. Vemuri, *Shape modeling with front propagation: A level set approach*, in IEEE Trns. On Pattern Analysis and Machine Intelligence, 1995; 17(2), pp 158-175.
- [4] V. Caselles, R. Kimmel, and G. Sapiro, *Geodisk Snakes*, Proc. of ICCV, MIT Cambridge MA, 1995.
- [5] J.A. Sethian, *Level set methods and fast marching method: evolving interfaces in computational geometry, fluid mechanics, computer vision, and material science.*, Cambridge University Press, 1999.
- [6] P.C. Johns, M.J. Yaffe, X-ray characterization of normal and neoplastic breast tissue, Phys. Med. Biol, 1987; 32, pp 675-695.
- [7] The Mammographic Image Analysis Society, *Digital Mammography Database ver 1.2*, (<http://www.wiau.man.ac.uk/services/MIAS/MIASweb.html>).
- [8] Wirth W., Nikitenko D., Lyon J., *Segmentation of Breast Region in Mammograms using a Rule-Based Fuzzy Reasoning Algorithm*. ICGST- Graphics, Vision and Image Processing Journal, 2005; 5(2); pp 45-54.
- [9] Luc Vincent. *Morphological Area Opening and Closing for Grayscale Images*, Proc. NATO Shape in Picture Workshop, Driebergen, The Netherlands, Springer-Verlag, pp. 197-208, September 1992.
- [10] Soille P., *Morphological Image Analysis: Principles and Applications*. Springer-Verlag. 2003.
- [11] Wirth M.A., Lyon J., Nikitenko N., Stapinski A. *Removing Radioopaque Artifacts from Mammograms using Area Morphology*. In Proc. of SPIE Medical Imaging: Image Processing. pp 1054-1065. 2004.
- [12] Salembier P., Serra J. *Flat Zones Filtering, Connected Operators, and Filters by Reconstruction*. IEEE Transaction on Image Processing, 4(8), pp 1153-1160. 1995.
- [13] Hoyer A., Spiesberg W., *Computerized mammogram processing*. In: Phillips Technical Review. Vol. (38), pp 347-355, 1979.
- [14] Lau T., Bischoff W., *Automated Detection of Breast Tumours Using the Asymmetry Approach*. In. Computers and Bio-Medical Research. Vol. (24), pp 273-295, 1991.
- [15] Bick U., Giger M. L., Schmidt R.A., Nishikawa R.M., Wolverson D.E., Doi K., *Automated Segmentation of Digitized Mammograms*. Academic Radiology, Vol. 2(2), pp 1-9, 1995.
- [16] Masek M., Attikiouzel Y., *Skin-Air Interface Extraction from Mammograms Using an Automatic Local Thresholding Algorithm*. In: ICB Brono CR, pp 204-206, 2000.
- [17] Abdel-Mottaleb M., Carman C.S., Hill C.R., Vafai S., *Locating the Boundary Between the Breast Skin Edge and the Background in Digitized Mammograms*. In Proc. Of the 3<sup>rd</sup> International Workshop on Digital Mammography, pp 467-470, 1996.
- [18] Mendez A.J., Tahoces P.J., Lado M.J., Souto M., Correa J.L., Vidal J.J., *Automatic Detection of Breast Boarder and Nipple in Mammograms*. Computer Methods and Programs in Bio-Medicine, 49, pp 253-262, 1996.
- [19] Karssemeijer N., te Brake G., *Combining Single View Features and Asymmetry for Detection of Mass Lesions*. In IWDM, pp 1107-1108, 1998.
- [20] Chandrasekhar R., Attikiouzel Y., *Automatic Breast Boarder Segmentation by Background Modelling and Subtraction*. In Proc. of the 5<sup>th</sup> International Workshop on Digital Mammography, pp 560-565, 2000.
- [21] Wirth M.A., Stapinski A., *Segmentation of the Breast Region in Mammograms using Active Contours*. In Proc. of Visual Communication and Image Processing, 5510, pp 1995-2006, 2003.