[Expert Systems with Applications 39 \(2012\) 4505–4510](http://dx.doi.org/10.1016/j.eswa.2011.09.136)

Contents lists available at [SciVerse ScienceDirect](http://www.sciencedirect.com/science/journal/09574174)

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

An improved GVF snake based breast region extrapolation scheme for digital mammograms

Chen-Chung Liu^a, Chung-Yen Tsai^b, Ta-Shan Tsui^{b,c}, Shyr-Shen Yu^{b,*}

a Department of Electronic Engineering, National Chin-Yi University of Technology, 35, Lane 215, Sec. 1, Chung-Shan Rd., Taiping, Taichung 411, Taiwan ^b Department of Computer Science and Engineering, National Chung-Hsing University, 250, Kuo-Kuang Rd., Taichung 402, Taiwan ^cDepartment of Applied Mathematics, National Chung-Hsing University, 250, Kuo-Kuang Rd., Taichung 402, Taiwan

article info

Keywords: Breast Mammogram Extrapolation GVF snake Edge map adjusting

ABSTRACT

To accurately extrapolate the breast region from a mammogram is a crucial stage of breast mass analysis. It significantly influences the overall analysis accuracy and processing speed of the whole breast mass analysis. In this paper, a novel edge map adjusting gradient vector flow snake (EMA GVF snake) algorithm for extrapolation of breast region from mammograms is proposed. In the proposed algorithm, the median filter is used to filter out the noise in a mammogram, the scale down stage is used to resize down the mammogram size (hence speeding up the extrapolation). The binarization processing stage and the morphological erosion processing stage are used to find a rough breast border. Then a novel gradient adjusting stage is applied to get a modified edge map and the gradient vector flow snake (GVF snake) is used to get the accurate breast border from the rough breast border. The proposed algorithm is tested on 322 digital mammograms from the Mammogram Image Analysis Society database. The mean error function, misclassification error function and the relative foreground area error function are conducted to evaluate the results of the detected breast border and the extracted breast region. Experimental results show that the breast border extrapolated by the proposed algorithm approximately follows the breast border extrapolated by an expert radiologist. Experimental results also show that the proposed algorithm is more robust and precise than the traditional GVF snake scheme for the breast extrapolation on mammograms.

- 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Breast cancer is one of the common cancers to cause death ([Bick](#page-5-0) [et al., 1995; Karnan & Thangavel, 2007; Kom, Tiedeu, & Kom, 2007;](#page-5-0) [Mendez et al., 1996; Ojala, Nappi, & Nevalainen, 2001](#page-5-0)) and is usually occurred in women. Therefore, the early detection and diagnosis of breast cancer is a very important procedure for mortality reduction. Many studies about the mass detection on mammogram have been published [\(Delogu, Fantacci, Kasae, & Reticoa, 2007; Kar](#page-5-0)[nan & Thangavel, 2007; Kom et al., 2007; Rojas Domi](#page-5-0)ı[nguez & Nan](#page-5-0)[di, 2008](#page-5-0)). However, the position and shape of tumor is not easy to be detected. Due to the detections of breast features are frequently affected by noises and artifacts ([Delogu et al., 2007; Karnan &](#page-5-0) [Thangavel, 2007; Kom et al., 2007; Rojas Domi](#page-5-0)ı[nguez & Nandi,](#page-5-0) [2008; Subashini, Ramalingam, & Palanivel, 2010](#page-5-0)) such that one cannot extract the true features effectively. And, the breast border and nipple position are usually used to discover the suspicious regions of tumor or microcalcification in digital mammograms [\(Kar](#page-5-0)[nan & Thangavel, 2007](#page-5-0)). So, to accurately extrapolate the breast

region from a mammogram is a crucial stage of breast mass analysis.

There are two main problems that affect the accuracy of the breast region extrapolation for a mammogram. One is the noises appearing in mammograms, and the other is that the gradient on the breast boundary is too small ([Ojala et al., 2001\)](#page-5-0). In other words, the gray level value on the breast boundary is almost the same as the gray level value of the background. On the other hand, the non-uniform background may have bright regions, such as the information plates and the unexposed film regions, that also low down the performance of an extrapolation scheme of breast region.

For the extrapolation scheme of breast region from a mammogram, Bick et al. analyze the local intensity range and the modified histogram of a mammogram to classify pixels into breast region and non-breast region ([Bick et al., 1995\)](#page-5-0). First, they only take small part of pixels that their pixel values are in the local intensity range as the breast region, and then they employ the region growing scheme and the morphological filtering scheme to extrapolate the final breast region. Their algorithm was tested on 740 mammograms and their extrapolation results were evaluated via vision, and 97% of the testing cases were rated as ''acceptable''. [Ojala](#page-5-0) [et al. \(2001\)](#page-5-0) compute the local variance for each gray level bin located between the left peak (background) and the middle peak

[⇑] Corresponding author.

E-mail addresses: ccl@ncut.edu.tw (C.-C. Liu), James1022jk@gamail.com (C.-Y. Tsai), tstsui@amath.nchu.edu.tw (T.-S. Tsui), pyu@nchu.edu.tw (S.-S. Yu).

(breast region) in the histogram of a mammogram. The maximum local variance of gray level bins is selected as the threshold to extrapolate an initial extrapolation of breast region for a mammogram. And then Ojala et al. use Fourier transform [\(Gonzalez &](#page-5-0) [Woods, 1992\)](#page-5-0), snake ([Kass, Witkin, & Terzopoulos, 1987\)](#page-5-0) and B-splines ([Jain, 1989\)](#page-5-0) to obtain the final extrapolation of breast region for a mammogram. Their algorithm was tested on 20 mammograms and was evaluated by using a mean error (ME) function. ME function computes the average absolute difference between the manual extrapolation breast region and the extrapolation breast region extrapolated by an automatic scheme. Their algorithm took 96% of the 20 testing cases to be rated as ''acceptable''. [Karnan and Thangavel \(2007\)](#page-5-0) used Bick et al. scheme to extract the breast boundary for a mammogram, then utilized gene algorithm to approximate the boundary of the breast.

The snake (or called active contour model) proposed by [Kass et](#page-5-0) [al. \(1987\)](#page-5-0) is to obtain a deformable curve in the spatial domain of an image by minimizing an energy function. The shape of a snake is changed by the internal forces that attempt to maintain the smoothness of the shape of the snake, and is also changed by the external forces which attract the snake to toward the salient features of the image. The traditional snake model has two drawbacks: one is that the initial snake must be quite close to the true boundary, and the other is that the active contour is not easy to deform into concave boundary points. Therefore, [Xu and Prince](#page-5-0) [\(1998\)](#page-5-0) proposed an improved snake, gradient vector flow (GVF) snake, to obtain a better performance for image segmentations. The basic idea of the GVF snake is to enlarge the influence range of each external force by creating a gradient vector flow field to replace the external force field of the traditional snake. The GVF snake improves the restricted capture range and the weak convergence of the traditional snake. But it is still dependent on the gradient distribution of the image.

The scheme proposed in this paper combines the median filter, the scale down stage, the binarization processing stage, the morphological erosion processing stage, a novel gradient adjusting stage, and the GVF snake to obtain the extrapolation of breast region from mammograms. The gradient adjusting stage can overcome the problem that the gray level value on the breast boundary is too close to that of background. This can let the GVF snake converge to the real breast boundary very well. The remainder of this paper is organized as follows: Section 2 introduces the proposed extrapolation scheme of breast region from mammograms. Section [3](#page-3-0) presents the experimental results. Finally, the conclusions of this paper are presented in Section [4.](#page-5-0)

2. Breast segmentation

The edge map adjusting GVF snake (EMA GVF snake) breast region extrapolation scheme presented in this paper improves the scheme proposed by [Ojala et al. \(2001\).](#page-5-0) There are two main steps in the proposed algorithm: First, the proposed algorithm uses the binarization scheme, the morphological opening operation, and the morphological erosion operation to obtain an initial extrapolation of the breast region. Then a gradient adjusting stage is employed and GVF snake is used to obtain the final extrapolation of breast region. [Fig. 1](#page-2-0) shows a flow chart of the proposed scheme for extrapolation of breast region from mammograms. Details of the presented extrapolation scheme are described in the following subsections.

2.1. Initial extrapolation

In order to obtain a better initial extrapolation, a 3 \ast 3 pixels median filter is conducted to reduce the noise in the original mammograms. On the other hand, in order to save the time consumption and to reserve the extrapolation quality, the 1024 $*$ 1024 pixels original mammograms in database are resized to 256 $*$ 256 pixels. Then, the binarization processing stage uses the mean gray value of the resized mammogram as the threshold to binarize the resized mammogram into a black–white image and to obtain the rough breast region. Finally, the morphological erosion processing stage is used to find a clearer initial contour of the breast region. These stages are detailedly illustrated in the follows.

2.1.1. Binarization

The binarization step is used to obtain a rough contour. Note that the gray value of the real border of the breast is similar to that of the background. The proposed algorithm constructs the intensity histogram H of the original mammogram, and adapts a simple way to obtain a threshold T to divide the original mammogram I into background (pixel value is 0) region and objects (pixel value is 1) region. The threshold T is calculated with the following formula:

$$
T = 0.2 \cdot \frac{\sum_{n=0}^{255} n \cdot H(n)}{\sum_{n=0}^{255} H(n)},
$$
\n(1)

where *n* is the gray level value, it is between 0 and 255. $H(n)$ is the number of pixels with pixel value n . The original mammogram I is transformed into a binary image IB (see [Fig. 2](#page-2-0)) by the following formula:

$$
IB(x, y) = \begin{cases} 1, & \text{if } I(x, y) \ge T, \\ 0, & \text{otherwise.} \end{cases}
$$
 (2)

2.1.2. Morphological processing

In general, the largest area object in *IB* is the breast region and other objects may be the information plates of the original mammogram or other noises. For obtaining a better and simple binary image of breast, the presented algorithm adapts the morphological opening operation to eliminate the non-breast objects in IB. The structuring element of the morphological opening operation is a disk with radius 2 pixels.

2.1.3. Extract breast region

Many non-breast objects still survive in the binary image that has been operated by the morphological opening operation, but their sizes are smaller than the size of the object breast. The presented algorithm utilizes the characteristic of the object breast in a mammographic image to label the objects, and then to extract the largest size object as the breast region in IB. The final binary image which indicates the breast region is denoted as IB_{fin} . Those non-largest size objects will be set to be the background, and the gray levels of these objects in I are reset with a new gray level. The new gray level is the mean of gray level of the non-breast region in I (see [Fig. 3](#page-2-0)).

2.1.4. Extract breast contour

The breast contour is extracted from the binary image IB_{fin} by a morphological erosion operation and the arithmetic subtraction operation; it can be represented as the following equation:

$$
IB_{\text{edge}} = IB_{\text{fin}} - IB_{\text{ero}},\tag{3}
$$

where IB_{ero} is the image obtained by taking a morphological erosion operation on the binary image IB_{fin} , IB_{edge} is the result breast border image (see [Fig. 4\)](#page-2-0).

Fig. 1. A flow chart of the EMA GVF snake based scheme for breast region extrapolation from mammographic images.

Fig. 2. The initial binary output of the presented algorithm: (a) original mammogram and (b) after binarization.

Fig. 3. (a) Original mammogram and (b) after non-breast objects removal.

2.2. Calculation of gradient

The breast border image IB_{edge} obtained in the previous section is served as the initial breast contour in the GVF snake processing. But in mammograms, the gray levels between the real breast border and the background pixels are various around the real breast border. If a GVF snake scheme directly uses the gradient field of mammograms obtained by edge schemes such as Sobel scheme, Canny scheme, or Prewitt scheme etc. for the external force, it shall obtain a worse final breast border. For obtaining a more precise breast border with GVF snake, a gradient adjusting stage (GA stage) is proposed to find a suitable gradient field of a mammogram for the external force for GVF snake. There are two stages to find the

Fig. 4. (a) Original mammogram and (b) initial breast border extraction.

adjusted gradient field (will be served as the adjusted edge map) of a mammogram, and are illustrated in the follows.

2.2.1. Stage 1: Gradient calculation

In Stage 1, the proposed algorithm utilizes Sobel filter to obtain the gradient field G of the original mammogram I. Sobel filter uses the following equation to find each pixel's gradient

$$
G(x, y) = |G_x(x, y)| + |G_y(x, y)|,
$$
\n(4)

where $G_x(x, y)$ and $G_y(x, y)$ indicate the vertical component and the horizontal component of the gradient of the pixel located at (x, y) , respectively.

2.2.2. Stage 2: Adjustment of gradient

Then in Stage 2, the following equation is proposed to adjust the gradient field G of the original mammogram I

$$
G_{\text{adjust}}(x, y) = w_0((G(x, y) - G_{\min})/(G_{\max} - G_{\min})) + w_1((T - I(x, y))/255),
$$
\n⁽⁵⁾

where G_{adjust} is the gradient image after adjusting. w_0 and w_1 are weights assigned by users, $0 \leq w_0 \leq 1$, and $w_1=1 - w_0$. Here w_0 is set to 0.63 and it is an experimental datum. T is the threshold men-tioned at Section [2.1.1.](#page-1-0) G_{min} and G_{max} represent the minimum value and the maximum value of the magnitude of the gradient field G, respectively.

The adjusted gradient of each pixel is between 0 and 1. This adjustment makes the breast border closer to the real boundary. Because the gray level of real breast boundary usually is similar to background (pixel value is low) and smooth. The proposed gradient field adjustment reduces the relative gradient of bright region and increases the relative gradient of dark region, such that the difference between the extracted breast border and the real breast border is reduced. In Ojala et al. scheme [\(Ojala et al.,](#page-5-0) [2001\)](#page-5-0), it is going to find a suitable global threshold to obtain breast boundary, then uses snake to smooth boundary. In our scheme, the threshold is used first to get the binary image IB and second to adjust the gradient field. GVF snake will be employed to process the initial contour IB_{edge} obtained in Section [2.1.4](#page-1-0) based on the image G_{adjust} to obtain the accuracy boundary.

2.3. GVF snake processing

In the GVF snake processing stage, the breast border image IBedge is deformed again and again in each iterative operation. The GVF snake iterative operations will stop while the two breast borders' difference of two adjacent iterative operations is no larger than a preassigned tolerance or the iteration reaches a given number of iterations.

In traditional snake model ([Kass et al., 1987](#page-5-0)), a curve $\mathbf{x}(s) = [x(s), y(s)], s \in [0,1],$ is found to minimize the energy functional

$$
E_{\text{snake}} = \int_0^1 E_{\text{int}} + E_{\text{ext}} ds
$$

=
$$
\int_0^1 \frac{1}{2} (\alpha |\mathbf{x}'(s)|^2 + \beta |\mathbf{x}''(s)|^2) + E_{\text{ext}}(\mathbf{x}(s)) ds,
$$
 (6)

where α and β are weighting parameters that control the snake's tension and rigidity, respectively. $\mathbf{x}'(s)$ and $\mathbf{x}''(s)$ represent the first and second derivatives of $\mathbf{x}(s)$ with respect to s. The external energy E_{ext} could be assigned as

$$
E_{\text{ext}}^1(x, y) = -|\nabla I(x, y)|^2, \qquad (7)
$$

$$
E_{\text{ext}}^2(x, y) = -|\nabla (G_{\sigma}(x, y) * I(x, y))|^2, \tag{8}
$$

where $G_{\sigma}(x,y)$ is a Gaussian function with a standard deviation σ and ∇ is the gradient operator. It has been shown that finding the curve $\mathbf{x}(s)$ which minimizes (6) is equivalent to solving the following equation

$$
\alpha \mathbf{x}''(s) - \beta \mathbf{x}'''(s) - \nabla E_{\text{ext}} = 0. \tag{9}
$$

Consider the partial derivative of x with respect to time t and set

$$
\mathbf{x}_t(s,t) = \alpha \mathbf{x}''(s,t) - \beta \mathbf{x}^{''''}(s,t) - \nabla E_{\text{ext}}.
$$
\n(10)

Then when the solution $x(s,t)$ stabilizes, one achieves a solution of (9).

In GVF snake model ([Jain, 1989](#page-5-0)), an edge map $f(x,y)$ is assigned first. Usually, $f(x,y) = -E_{ext}(x,y)$. In the proposed EMA GVF snake model, the edge map f is assigned to the adjusted gradient image G_{adjust} . Then a gradient vector flow (GVF) field is defined as the vector field $V(x,y) = (u(x,y), v(x,y))$ which minimizes the energy functional

$$
\varepsilon = \iint \mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla f|^2 |V - \nabla f|^2 dxdy, \qquad (11)
$$

where μ is a regularization parameter.

The GVF field can be acquired via solving the following Euler– Lagrange equations

$$
\mu \nabla^2 u - (u - f_x)(f_x^2 + f_y^2) = 0, \tag{12}
$$

$$
\mu \nabla^2 v - (v - f_y)(f_x^2 + f_y^2) = 0, \tag{13}
$$

where ∇^2 is the Laplacian operator. Furthermore, Eqs. (8) and (9) can be solved via treating u and v as functions of time and solving

$$
u_t(x, y, t) = \mu \nabla^2 u(x, y, t) - (u(x, y, t) - f_x(x, y)) \cdot (f_x(x, y)^2 + f_y(x, y)^2),
$$
\n(14)

$$
v_t(x, y, t) = \mu \nabla^2 v(x, y, t) - (v(x, y, t) - f_y(x, y)) \cdot (f_x(x, y)^2 + f_y(x, y)^2).
$$
\n(15)

In the proposed EMA GVF snake model, the GVF field $V = (u, v)$ is initially given as (f_x, f_y) .

Once the final GVF field $V = (u, v)$ is obtained, the following dynamic GVF snake equation

$$
\mathbf{x}_t(s,t) = \alpha \mathbf{x}''(s,t) - \beta \mathbf{x}^{'''}(s,t) + V \tag{16}
$$

is applied to find the curve $x(s)$, which is the desired final border. In the proposed EMA GVF snake, the curve $x(s)$ is initially assigned to the breast border given in IB_{edge} . This equation can be solved in similar fashion to the traditional snake. For more details of snake and GVF snake, one is referred to [Kass et al. \(1987\) and Xu and Prince](#page-5-0) [\(1998\)](#page-5-0), respectively. Fig. 5 shows the output result of the presented algorithm.

3. Experiment results

Three hundred and twenty-two mammograms of the Mammogram Image Analysis Society (MIAS) database are used to show the performance of the proposed scheme. Each mammogram is 50 μ m/ pixel. And, the size of each mammogram is $1024 * 1024$ and bitdepth of 8 bits. Each of these mammograms is labeled by an expert radiologist with hand carefully to construct the ground truth extrapolation. The tool used for hand labeling is taken from the

Fig. 5. (a) Original mammogram and (b) final extracted breast border.

Table 1

The ME1, ME2, and RFAE between the manually labeled breasts and the breasts extracted by initial segmentation, GVF snake, EMA GVF snake, and [Ojala et al. \(2001\)](#page-5-0) for the 322 mammographic images of the MIAS database.

		Max	Min	Mean	Std
Initial segmentation (obtained via Section 2.1)	ME ₁ ME ₂ RFAE	13.9607 0.0959 0.2177	0.1722 0.0017 0.0001	3.4128 0.0286 0.0635	2.4797 0.0173 0.0356
Using GVF snake	ME ₁	18.4955	1.5657	9.4935	2.9883
	ME ₂	0.9078	0.0165	0.0978	0.0630
	RFAE	0.8737	0.0380	0.2339	0.0828
Using EMA GVF snake	ME ₁	12.4108	0.2105	2.2818	1.8480
	ME ₂	0.0747	0.0018	0.0144	0.0077
	RFAE	0.1688	0.0000	0.0166	0.0161
Ojala et al. (2001)	ME ₁	24.8596	0.6281	3.7469	12.8618
	ME ₂	45.1840	0.8103	4.1454	19.0790
	RFAE	0.5406	0.0138	0.0782	0.0050

tool presented in [Hoover et al. \(1994, 1996\),](#page-5-0) which was used to create hand-labeled images for estimating image segmentation schemes.

Several performance metrics are often used to put into evidence which shows the different performance features of the segmentation algorithms. They are mean error function ME1 ([Karnan &](#page-5-0) [Thangavel, 2007](#page-5-0)), misclassification error function ME2 [\(Tsai, Chan,](#page-5-0)

[Lin, Yang-Mao, & Huang, 2008](#page-5-0)), and the relative foreground area error (RFAE) ([Sezgin & Sankur, 2004\)](#page-5-0). The ME1 function is calculated as follows:

$$
ME1 = \frac{1}{n} \sum_{i=1}^{n} \sqrt{(x.m_i - x_i)^2 + (y.m_i - y_i)^2},
$$
\n(17)

Fig. 6. (a) The original mammogram image, mdb136. (b) The extrapolation result of [Ojala et al. \(2001\).](#page-5-0) (c) The extrapolation result of the proposed scheme.

Fig. 7. (a) The original mammogram image, mdb054. (b) The ground truth extrapolation of (a). (c) The initial contour of (a). (d)-(i) The borders obtained by (d) EMA GVFsnake, (e) GVF-snake, (f) Sobel, (g) Canny edge, (h) 2-means and (i) watershed.

where *n* is total number of breast border points. (x m_i, γ, m_i) is the Cartesian coordinate of the ith point of the manual boundary of the breast and (x_i, y_i) is the Cartesian coordinate of the detected breast border point which is closest to the ith point of the manual breast boundary. If the manual boundary and the detected boundary are closer, the value of ME1 is closer to zero. The ME2 function is defined as follows:

$$
ME2 = 1 - \frac{TP + TN}{TP + FN + TN + FP} = \frac{FP + FN}{TP + FN + TN + FP}.
$$
\n(18)

The RFAE is defined as follows:

$$
\text{RFAE} = \begin{cases} \frac{\frac{TP+FN-(FP+TP)}{TP+FN}}{\frac{FP+TP}{FP+FN}} = \frac{FN-FP}{TP+FN} & \text{if } (FP+TP) < (TP+FN),\\ \frac{FP+TP- (TP+FN)}{FP+TP} = \frac{FP-FN}{FP+TP} & \text{if } (FP+TP) \geq (TP+FN), \end{cases} \tag{19}
$$

where TP indicates the areas of true positive, TN represents the areas of true negative, FP indicates the areas of false positive, and FN represents the areas of false negative respectively. $(FP + TP)$ indicates the extracted object and $(TP + FN)$ denotes the ground-truth object. Note that the object extrapolation is better while RFAE is smaller.

In this paper, ME1, ME2, and RFAE are taken to measure the efficiency of the presented breast region extrapolation algorithm. The performances of the extrapolation methods: initial segmentation given in Section [2.1](#page-1-0), GVF snake, EMA GVF snake and Ojala et al. (2001) on the 322 mammograms of the MIAS database are given in [Table 1.](#page-3-0)

[Table 1](#page-3-0) tells that all the means of ME1, ME2, and RFAE of GVF snake segmentation is larger than those of initial segmentation and EMA GVF snake segmentation. In fact, traditional GVF snake always pushes the breast border toward the brightest region (the interior of the breast). On the other hand, the EMA GVF snake always pulls the breast border toward the darker region. And the real breast border frequently exists in the darker region. [Table 1](#page-3-0) shows that the presented scheme can give a very precise extrapolation for mammograms.

Since Ojala et al. (2001) is effected by mammogram histogram, if histogram is not bimodal, such as, there are many peaks or many bin values are close in histogram, the thresholding segmentation may extract worse breast contour. Thus, we eliminated some worse results, such as the result shown in [Fig. 6,](#page-4-0) from the calculation of result errors of Ojala et al. (2001).

[Fig. 7](#page-4-0) is given to compare results obtained by using EMA GVF snake, GVF snake, Sobel, Canny edge, and watershed. In [Fig. 7,](#page-4-0) the breast borders obtained by Sobel and Canny edge are very rough and uncertain. So, these two methods cannot automatically detect the desired integrated components or closed connective edges as an initial segmentation. The result using watershed includes too many regions. Since the gray level of a real breast border is close to the dark background and various in each case, it is not easy for using a general region growing method to get an appropriate initial segmentation. Thus the method proposed in Section [2.1](#page-1-0) which can get the initial border and also provide a threshold for edge map adjusting is more suitable for initial breast border extraction.

4. Conclusions

An algorithm for extrapolation of the breast region from mammographic images is developed in this paper. This algorithm uses the binarization scheme, the morphological opening operation, and the morphological erosion operation to obtain a rough initial extrapolation of the breast region from a mammogram. Then GVF snake is employed to process the rough initial breast border based on an adjusted edge map to obtain the final extrapolation of breast region. The traditional GVF snake always pushes the breast border toward the brightest region (the interior of the breast). However, the real breast border is frequently exists in the darker region and the rough initial breast border is always in the interior of the breast. The EMA GVF snake segmentation pulls the breast border toward the darker region. Therefore, the presented algorithm provides a suitable edge map for GVF snake to determine the external force used in snake scheme such that the snake scheme pulls the breast border outward to the real breast border. The presented algorithm is applied over the 322 mammograms of the MIAS database. The experimental results show the breast border extrapolated by the presented EMA GVF snake algorithm approximately follows the breast border extrapolated by an expert radiologist. The experimental results also show the presented algorithm is more robust and more precise than the traditional GVF snake scheme and Ojala et al. (2001) for the breast extrapolation from the mammograms.

For future works, we are going to detect the nipple and muscle regions. Then we are going to discover the suspicious regions of tumor and microcalcification. Finally, an automatic breast cancer detection system could be constructed.

References

- Bick, U., Giger, M. L., Schmidt, R. A., Nishikawa, R. M., Wolverton, D. E., & Doi, K. (1995). Automated segmentation of digitized mammograms. Academic Radiology, 2, 1–9.
- Delogu, P., Fantacci, M. E., Kasae, P., & Reticoa, A. (2007). Characterization of mammographic masses using a gradient-based segmentation algorithm and a neural classifier. Computers in Biology and Medicine, 37, 1479–1491.
- Gonzalez, R. C., & Woods, R. E. (1992). Digital image processing. Reading, MA: Addison-Wesley.
- Hoover, A., Jean-Baptiste, G., Goldgof, D., & Bowyer, K. (1994). A methodology for evaluating range image segmentation techniques. In Second IEEE workshop applications computer vision, Sarasota, FL (pp. 264–271).
- Hoover, A., Jean-Baptiste, G., Jiang, X., Flynn, P. J., Bunke, H., Goldgof, D., et al. (1996). An experimental comparison of range image segmentation algorithms. IEEE Transactions on Pattern Analysis and Machine Intelligence, 673–689.
- Jain, A. K. (1989). Fundamentals of digital image processing. Prentice Hall Int. (pp. 200–202).
- Karnan, M., & Thangavel, K. (2007). Automatic detection of the breast border and nipple position on digital mammograms using genetic algorithm for asymmetry approach to detection of microcalcifications. Computer Methods and Programs in Biomedicine, 87, 12–20.
- Kass, M., Witkin, A., & Terzopoulos, D. (1987). Snakes: Active contour models. International Journal of Computer Vision, 1(4), 321–331.
- Kom, G., Tiedeu, A., & Kom, M. (2007). Automated detection of masses in mammograms by local adaptive thresholding. Computers in Biology and Medicine, 37, 37–48.
- Mendez, A. J., Tahoces, P. G., Lado, M. J., Souto, M., Correa, J. L., & Vidal, J. J. (1996). Automatic detection of breast border and nipple in digital mammograms. Computer Methods and Programs in Biomedicine, 49, 253–262.
- Ojala, T., Nappi, J., & Nevalainen, O. (2001). Accurate segmentation of the breast region from digitized mammograms. Computerized Medical Imaging and Graphics, 25, 47–59.
- Rojas Domiınguez, A., & Nandi, A. K. (2008). Detection of masses in mammograms via statistically based enhancement multilevel-thresholding, segmentation and region, selection. Computerized Medical Imaging and Graphics, 32, 304–315.
- Sezgin, M., & Sankur, B. (2004). Survey over image thresholding techniques and quantitative performance evaluation. Journal of Electronic Imaging, 13(1), 146–165.
- Subashini, T. S., Ramalingam, V., & Palanivel, S. (2010). Automated assessment of breast tissue density in digital mammograms. Computer Vision and Image Understanding, 114(1), 33–43.
- Tsai, M. H., Chan, Y. K., Lin, Z. Z., Yang-Mao, S. F., & Huang, P. C. (2008). Nucleus and cytoplast contour detector of cervical smear image. Pattern Recognition Letters, 29, 1441–1453.
- Xu, C., & Prince, J. L. (1998). Snakes, shapes, and gradient vector flow. IEEE Transactions on Image Processing, 7(3), 359–369.