A Novel Approach to Breast Ultrasound Image Segmentation Based on the Characteristics of Breast Tissue and Particle Swarm Optimization

Yanhui Guo [1, 2], H.D. Cheng[1,2], Jiawei Tian[3], Yingtao Zhang[1]

[1]School of Computer Science and technology, Harbin Institute of Technology, Harbin, China, 150001
[2]Department of Computer Science, Utah State University, Logan, UT 84322 U.S.A.
[3]Second Affiliated Hospital of Harbin Medical University, Harbin, China.

Abstract

Breast cancer occurs to over 8% women during their lifetime, and is a leading cause of death among women. Sonography is superior to mammography in its ability to detect focal abnormalities in the dense breasts and has no side-effect. In this paper, we proposed a novel automatic segmentation algorithm based on the characteristics of breast tissue and the eliminating particle swarm optimization (EPSO) clustering analysis. The characteristics of mammary gland in breast ultrasound (BUS) images are analyzed and utilized, and a method based on step-down threshold technique is employed to locate the mammary gland area. The EPSO clustering algorithm employes the idea of "survival of the superior and weeding out the inferior". The experimental results demonstrate that the proposed approach can segment BUS image with high accuracy and low computational time. The EPSO clustering method reduces the computational time by 32.75% compared with the standard PSO clustering algorithm. The proposed approach would find wide applications in automatic lesion classification and computer aided diagnosis (CAD) systems of breast cancer.

Keywords: Image segmentation, ultrasound image, partical swarm optimization, clustering.

1. Introduction

Breast cancer is still one of the most common cancers and a leading cause of death among women [1, 2]. Since the underlying molecular mechanism of this disease still remains unknown, early detection and diagnosis are very essential in reducing the mortality. More and more emphases are on early detection and diagnosis of breast cancer [3-5].

Currently, breast ultrasound (BUS) imaging is a valuable method in early detection and classification of breast lesions [6]. Sonography was more effective for women younger than 35 years old of age [7]. The results [8] show that the denser the breast parenchyma, the higher the detection accuracy of malignant tumors using US. The accuracy rate of breast ultrasound imaging has been reported to be 96-100% in the diagnosis of simple benign cysts [9]. It has been shown that breast sonography is superior to the mammography in detecting focal abnormalities in the dense breasts of adolescent women [10]. Furthermore, sonography can display mass obscured mammgraphi-

Proceedings of the 11th Joint Conference on Information Sciences (2008)
Published by Atlantis Press

cally by dense tissue, and it is low cost, portable and has no ionizing radiation [6].

In this paper, aiming to clinical application, we not only employ the entire images for segmentation, but also utilize the breast US image characteristics, and propose a novel automatic segmentation algorithm. The characteristics of mammary gland in breast ultrasound images are utilized and a method based on step-down threshold technique is employed to extract the mammary gland area. Furthermore, a new eliminating PSO clustering algorithm is proposed based on the idea of "survival of the superior and weeding out the inferior" to process the images quickly and accurately.

2. Proposed approach

2.1. Speckle reduction

Speckle noise is inherent in ultrasound imaging, and tends to reduce the resolution and contrast, thereby, diminishing the diagnostic accuracy. In order to remove speckle noise, we implement an algorithm for speckle reduction based on two-dimension (2D) homogeneity histogram (homogram) and directional average filters [11].

2.2. Mammary gland region extraction

There are four layers of a breast ultrasound image: skin, subcutaneous tissue, mammary gland and muscle. The boundaries between these layers are quite blur as shown in Fig. 1. Due to the fact that the mass resides in the mammary gland region, it will improve the performance of breast cancer detection if the mammary gland region can be located correctly and accurately.

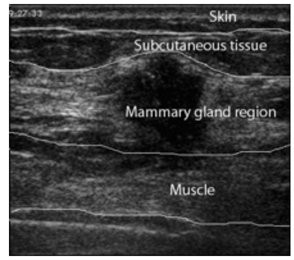


Fig. 1. Breast structure in ultrasound image

The breast ultrasound image after speckle reduction is transformed into binary image using the step-down threshold value $Th(m)$.

$$bw^m(i,j)=\begin{cases}0 & \tilde{g}(i,j)\le Th(m)\\ 255 & otherwise\end{cases} \quad (1)$$

where $\tilde{g}(i,j)$ is the gray level of $P(i,j)$ after speckle reduction, $Th(m)$ is the mth step-down threshold value, and $bw^m(i,j)$ is the binary image after the mth thresholding processing.

Searching binary image, some white regions $\mathrm{Re}W(n)$ are found, and the regions connected with the top and bottom margins will be eliminated.

Because the subcutaneous tissues margin and muscle lines have high gray levels, the subcutaneous tissues margin and muscle line region can be located among the white regions obtained. The ratio Ra, which is defined as a ratio between the width and height, is used to eliminate the false subcutaneous tissues and muscle region.

$$Ra(n)=max(RaWH(n)) \quad (2)$$

$$RaWH(n)=\frac{W(n)}{H(n)} \quad (3)$$

where $RaWH(n)$ is the ratio between the width and height of the nth candidate white region $\mathrm{Re}W(n)$.

The subcutaneous tissue margin and chest muscle line detected above are the top and bottom margins of the mammary gland region which is extracted successfully from the subcutaneous tissue and the muscle layer.

2.3. Eliminating particle swarm optimization algorithm

Particle swarm optimization (PSO) algorithm is an evolutionary computation technique utilizing random search inspired by the mechanism of natural selection and genetics to emulate the evolutionary behaviors of biological systems. The PSO is introduced in [13], which simulates simplified swarm social models such as bird flocking and fish schooling.

Let P_i represent the ith particle, whose position and velocity in a d-dimensional space are defined as X_{id} and V_{id}, respectively. The position and velocity are updated according to the following formulas:

$$\begin{aligned} V_{id}(t) &= \omega V_{id}(t-1) \\ &+c_1 rand()(P_{id}(t-1)-X_i(t-1)) \quad (4) \\ &+c_2 rand()(P_{ig}(t-1)-X_i(t-1)) \end{aligned}$$

$$X_{id}(t) = X_{id}(t-1)+V_{id}(t-1) \quad (5)$$

where $X_{id}(t)$ is the position of the ith particle in a d-dimensional space at time step t, V_i is the velocity of $P_i(t)$. Parameters c_1 and c_2 are learning factors, ω is an inertia weight and $rand()$ is a random function.

2.4. Eliminating PSO Clustering

A modified PSO, eliminating PSO (EPSO), is based on the idea of "survival of the superior and weeding out the inferior". N particles are initialized whose velocities and positions are updated accordingly, and the positions' fitness values are calculated and sorted in a list with the descending order. Then L particles are eliminated whose fitness values are in the last L positions of the list. This will reduce the computational time, while the accuracy of the solution is not affected. The process is iterated until the maximum iteration number is reached or the minimum error condition is satisfied.

EPSO clustering is an algorithm based on k-means clustering and EPSO algorithm. The centers of the clusters are considered as the particles positions, and EPSO algorithm is employed to search the optimum solution by eliminating the "weakest" particles to speed up the computation. The k-means clustering is utilized to update the positions of the particles.

The procedure of clustering analysis based on EPSO and k-means clustering is described below:

(1) Select M particles (primary population number) and put them into the primary swarm $S(1)=\{P_1, P_1, \cdots P_M\}$, and initialize the positions X_{id} of swarm S using k-means clustering results;

(2) Randomly initialize the velocities V_{id};

(3) Evaluate the fitness of each particle $Fit(X_{id}(t))$;

(4) Compare the personal best of each particle in the new swarm $S(t+1)$ with its current fitness value, and set $P_{id}(t)$ to the better performance.

$$P_{id}(t+1)=\begin{cases} P_{id}(t) & Fit(P_{id}(t))>Fit(X_{id}(t)) \\ X_{id}(t) & Fit(P_{id}(t))\le Fit(X_{id}(t)) \end{cases}$$

(5) Set the global best $P_{gd}(t+1)$ to the position of the particle with the best fitness in the swarm;

(6) Sort the particles according to the fitness values. A new swarm $S(t+1)$ is obtained by eliminating the L particles whose fitness values are in the last L positions of the list;

(7) Optimize the position of each new particle in the new swarm $S(t+1)$ according to k-means clustering principle;

(8) Change the velocity vector $V_{id}(t+1)$ for each particle according to Eq. (27);

(9) Update each particle position in $S(t+1)$;

(10) Go to step (3), and repeat the process until the maximum iteration number or the minimum error is reached.

2.5. Mammary gland image segmentation based on EPSO clustering

EPSO clustering algorithm is applied to mammary gland image segmentation. In the proposed method, the number of particles, N is initialized to be 65, and the number of particles to be eliminated in each iteration, L is 5. Because the intensities of the pixels belong to lesion are very low, the group of pixels with the lowest intensities could be regarded as the lesion-like pixels. Then the mammary gland region is located by the following formula:

$$BW_M(i,j) = \begin{cases} 0 & g_M(i,j) \in C_1 \\ 255 & otherwise \end{cases} \quad (6)$$

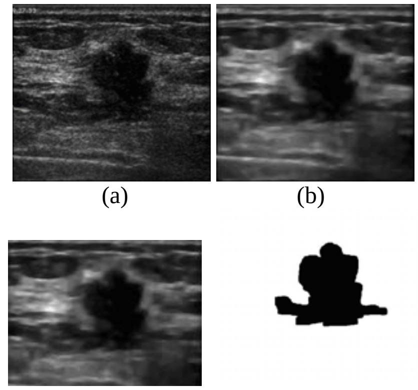
where $g_M(i,j)$ is the pixel in mammary gland region at the coordinates (i,j), and C_1 is the cluster with the lowest intensities.

3. Experimental results

The proposed algorithm is applied to a number of clinical ultrasound images. Figs. 2(a) and 3(a) are the original breast ultrasound images, where most of the bright areas are the breast and muscle tissues, and the suspicious tumor areas are corrupted by speckle noise.

Fig. 2(a) has a lesion with some branches at the center, which are important sign for breast cancer diagnosis. In Fig. 2(d), the mass and its branches are segmented correctly.

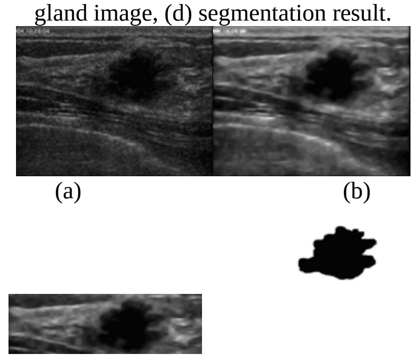
Fig. 3(a) displays a lesion with an ill-defined border. The segmentation result in Fig. 3(d) shows that the lesion border's integrity is persevered well. The mass is segmented correctly and the edge is distinct, which is more suitable for mass detection and classification.



(a) (b)

(c) (d)

Fig. 2 (a) Original image, (b) image after speckle reduction, (c) extracted mammary gland image, (d) segmentation result.



(a) (b)

(c) (d)

Fig. 3 (a) Original image, (b) image after speckle reduction, (c) extracted mammary gland image, (d) enhanced mammary gland image, (e) segmentation result.

4. Conclusions

Breast cancer is still one of the most common cancers and a leading cause of death among women. The automated segmentation of BUS images is an essential issue for CAD systems. However, most existing algorithms are only for segmenting the ROIs. In this paper, a breast ultrasound image segmentation algorithm based on EPSO clustering is proposed. The major advantage of the proposed algorithm is that it can handle the entire image instead of ROIs automatically and accurately, since it utilizes the characteristics of mammary gland of the breast ultrasound images. Also the algorithm has very low computational time complexity. The proposed approach may find wide applications in automatic lesion classification and computer aided diagnosis systems for breast cancer control.

5. Acknowledgement

The work was supported, in part, by Natural Scientific Research Innovation Foundation in Harbin Institute of Technology, Project HIT.NSRIF.2008.48, and Natural Science Foundation of China No.60873142 and No. 30670546.

6. References

[1] Y. Luo, J. Zhang, Y. Liu, A. C. Shaw, X. Wang, S. Wu, X. Zeng, J. Chen, Y. Gao, and D. Zheng, "Comparative proteome analysis of breast cancer and normal breast," *Molecular Biotechnology*, vol. 29, no. 3, pp. 233-244, 2005.

[2] A. Jemal, T. Murray, E. Ward, A. Samuels, R. C. Tiwari, A. Ghafoor, E. J. Feuer, and M. J. Thun, "Cancer Statistics, 2005," *CA Cancer J Clin*, vol. 55, no. 1, pp. 10-30, 2005.

[3] P. Hider and B. Nicholas, *The Early Detection and Diagnosis of Breast Cancer: A Literature Review: an Update*: Clearing House for Health Outcomes and Health Technology Assessment, Dept. of Public Health and General Practice, Christchurch School of Medicine, 1999.

[4] K. Kaul and F. M. Daguilh, "Early Detection of Breast Cancer:Is Mammography Enough?" *Hospital Physician*, vol. 9, pp. 49-55, 2002.

[5] E. Paci, "Mammography and beyond: developing technologies for the early detection of breast cancer," *Breast Cancer Res*, vol. 4, no. 3, pp. 123 - 125, 2002.

[6] K. Drukker, M. L. Giger, C. J. Vyborny, and E. B. Mendelson, "Computerized detection and classification of cancer on breast ultrasound," *Acad Radiol*, vol. 11, no. 5, pp. 526-35, 2004.

[7] L. W. Bassett, "Usefulness of mammography and sonography in women less than 35 years of age," *Radiology*, vol. 180, no. 3, pp. 831-835, 1991.

[8] H. Laine, J. Rainio, H. Arko, and T. Tukeva, "Comparison of breast structure and findings by X-ray mammography, ultrasound, cytology and histology: A retrospective study," *European Journal of Ultrasound*, vol. 2, no. 2, pp. 107-115, 1995.

[9] V. P. Jackson, "The role of ultrasound in breast imaging," *Radiology*, vol. 177, no. 2, pp. 305-311, 1990.

[10] C. Sohn, J. Blohmer, and U. M. Hamper, *Breast Ultrasound*: Thieme Publishing Group, 1998.

[11] Yanhui Guo, H. D. Cheng, J. Tian，, and Y. Zhang, "A Novel Approach to Speckle Reduction and its Application to Ultrasound Medical Image," *TR-USU-07-1, computer science department, Utah State University.*, 2007.

[12] Y. H. Guo, H. D. Cheng, J. H. Huang, J. W. Tian, W. Zhao, L. T. Sun, and Y. X. Su, "Breast ultrasound image enhancement using fuzzy logic," *Ultrasound in Medicine and Biology*, vol. 32, no. 2, pp. 237-247, 2006.

[13] J. Kennedy and R. Eberhart, "Particle swarm optimization," Neural Networks, 1995. Proceedings., IEEE International Conference on, pp. 1942-1948, 1995.