Pattern Recognition 48 (2015) 485–497

Contents lists available at ScienceDirect

Pattern Recognition

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

Fully automatic segmentation of breast ultrasound images based on breast characteristics in space and frequency domains

Min Xian [a,b], Yingtao Zhang [b], H.D. Cheng [a,*]

[a] Department of Computer Science, Utah State University, Logan, UT 84322, USA
[b] School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China

ARTICLE INFO

Article history:
Received 18 February 2014
Received in revised form
25 May 2014
Accepted 12 July 2014
Available online 12 August 2014

Keywords:
Breast ultrasound (BUS) images
Computer-aided diagnosis (CAD)
Region of interest (ROI) generation
Automatic segmentation

ABSTRACT

Due to the complicated structure of breast and poor quality of ultrasound images, accurately and automatically locating regions of interest (ROIs) and segmenting tumors are challenging problems for breast ultrasound (BUS) computer-aided diagnosis systems. In this paper, we propose a fully automatic BUS image segmentation approach for performing accurate and robust ROI generation, and tumor segmentation. In the ROI generation step, the proposed adaptive reference point (RP) generation algorithm can produce the RPs automatically based on the breast anatomy; and the multipath search algorithm generates the seeds accurately and fast. In the tumor segmentation step, we propose a segmentation framework in which the cost function is defined in terms of tumor's boundary and region information in both frequency and space domains. First, the frequency constraint is built based on the newly proposed edge detector which is invariant to contrast and brightness; and then the tumor pose, position and intensity distribution are modeled to constrain the segmentation in the spatial domain. The well-designed cost function is graph-representable and its global optimum can be found. The proposed fully automatic segmentation method is applied to a BUS database with 184 cases (93 benign and 91 malignant), and the performance is evaluated by the area and boundary error metrics. Compared with the newly published fully automatic method, the proposed method is more accurate and robust in segmenting BUS images.

1. Introduction

Occurring in high frequencies, breast cancer is one of the leading causes of cancer death among females worldwide [1,2]. Finding early signs and symptoms of breast cancer by clinical examination is the key to reduce the mortality [3,4], and breast ultrasound (BUS) is a major imaging modality for early detection [5]. However, clinical experience and expert knowledge are necessary to achieve correct diagnosis [1,6]. These make the human perception-based diagnosis suffer from considerable intra- and inter-observer variabilities. In order to overcome such drawback, computer-aided diagnosis (CAD) systems have been studied [7–10].

Segmentation is a critical step in a BUS CAD system. Manual segmentation methods are time-consuming and tedious, and suffer from great individual variability [11]. Semi-automatic segmentation methods solved the problem partially. Nevertheless, some interactions were still required which prevented thewidespread applications of BUS CAD systems. Therefore, driven by clinical needs and related applications, it is necessary and essential to develop automatic segmentation methods having the ability to reduce dependencies on operators and ultimately lead to a fully automated CAD system.

Automatic segmentation method usually refers to the segmentation process requiring little user intervention or no intervention at all; the former is called semi-automatic segmentation, while the latter is called fully automatic segmentation. The fully automatic methods are characterized by the prior constraints, such as shape, appearance and spatial location of the lesions. Several works discussed fully automatic segmentation for BUS images [13–17].

In [14], it employed gray-level gradient as segmentation constraint to detect tumors on BUS images automatically. The method located ROIs by the radial gradient index (RGI) filtering technique. The points of interest (POIs) at the centers of the ROIs were selected as the seeds; then a region growing algorithm based on the maximum average radial gradient (ARD) was applied to obtain the tumor contours. The method was tested on a database of 400 patients (757 images) and 75% tumors were correctly detected at an overlap of 0.4 with radiologist tumor outlines.

In [15], the authors proposed a hybrid completely automatic segmentation method which combined the region-based and boundary-based techniques. They developed a mathematical formulation of the empirical rules (intensity, texture and location) to generate tumor seeds automatically. Region growing and

* Corresponding author. Tel.: +1 435 7972054; fax: +1 435 7973265.
E-mail address: hengda.cheng@usu.edu (H.D. Cheng).

http://dx.doi.org/10.1016/j.patcog.2014.07.026
0031-3203/

directional gradient of the image were used to find the boundary points after seed generation. Then, a rule based on Euclidean distances between candidate boundary points and the seed were employed to refine the boundary points by eliminating the outliers. Finally, the refined boundary points serve as the initial boundary of a deformable model. Area error metrics were used to measure the performance of the algorithm on 42 BUS images, the true positive ratio (TPR) was 75.04%, false positive ratio (FPR) was 20.86%, and false negative ratio (FNR) was 24.96%, respectively.

Liu et al. [16] described a fully automatic method based on an active contour mode. The method was divided into two steps: ROI generation and ROI segmentation steps. In the first step, a supervised local texture classifier (a support vector machine) was trained to find the candidate ROIs. Vertical position and distance to the image center were employed to determine the final ROI. In the second step, it used an active contour model which combined the global statistical information and local edge information to locate the final tumor contour. The performance was evaluated by using area error metrics on a database of 103 BUS images (48 benign and 55 malignant) and the TPR was 91.31%, FPR was 8.69% and FNR was 7.26%, respectively.

Shan et al. [12] proposed an automatic seed point selection algorithm. The method formulated the texture, spatial location and size of the candidate area. It ranked the regions obtained from the iterative thresholding by the empirical formula, and selected the center of the winning region as the seed. Then, a fully automatic segmentation method was developed based on the features from the spatial and frequency domains [13,17]. Two new features called phase in max-energy orientation (PMO) and radial distance (RD) were proposed. Combing with common texture and gray level features [15], the new feature set was utilized to classify tumor pixels from the background by a trained ANN. The algorithm was tested on a database of 60 BUS images. It achieved a TPR of 93.41%, FPR of 12.72% and FNR of 6.59% by using the area error metrics; and an average Hausdorff error (AHE) of 18.90 pixels and an average mean absolute error (AME) of 5.04 pixels by using the boundary error metrics.

Notice that the methods mentioned above are all followed a two-step strategy: first, locating the ROI, and then segmenting the tumor in the ROI. Although some of the fully automatic methods have achieved quite good performance on their own datasets, these methods still suffer from some drawbacks, which impede the widespread application of BUS CAD systems.

1) **Lack of biological foundations**. Some existing BUS segmentation methods [14,18–22] applied computer vision, image processing, pattern recognition and machine learning algorithms for processing BUS images. However, without incorporating with the biological background of the breast, the BUS CAD systems cannot work well.
2) **Non-robust constraints.** Although some fully automatic segmentation methods employed biological knowledge of breast explicitly or implicitly, most of the ROI generation methods [13–17] were dependent on certain hard or inflexible constraints, such as the fixed reference points (RPs) for selecting seeds or ROIs [14], fixed gray value for thresholding, fixed region size and spatial location for refining candidate ROIs [16], which make these methods difficult to achieve good generalization ability. For example, a fixed RP in the middle of the image may result wrong tumor detection, if the tumor is far away from the center of the image.
3) **Low ability to exclude structures similar to tumors**. The methods [10,16] relying only on the tumor features from the spatial domain cannot effectively distinguish the tumors from the hypoechoic neighboring regions, such as fat regions and shadowing artifacts.

To solve above problems, an accurate, robust, and fast fully automatic segmentation method is proposed in this paper. In the ROI generation step, the newly proposed adaptive RP generation algorithm and the multipath search algorithm can locate the tumor region quickly without utilizing any fixed and inflexible constraints. In the second step, we proposed a segmentation framework, which utilized the edge information in the frequency domain and information of the intensity distribution, position and pose of tumors in the space domain. It has high ability to exclude normal region from tumor region. The flowchart of the proposed approach is illustrated in Fig. 1.

The paper is organized as follows: in Section 2, the fully automatic ROI generation algorithm is presented; in Section 3, the segmentation method is illustrated; in Section 4, the experimental results are discussed; and the conclusion is given in Section 5.

2. Fully automatic ROI generation

ROI generation is a prerequisite step for automatic segmentation and classification in many BUS CAD systems. A ROI is usually a rectangular region [10,13,16,17] which provides the rough location of a lesion and excludes normal tissues as much as possible. In [10,16], a supervised local texture classifier was proposed to find the candidate ROIs, and the fixed constraints, such as region size and spatial location, were used to select the final ROI. In [14], the authors located ROIs by using RGI filtering method and gray level thresholding, and the constraining function and the threshold must be predefined. In [13,17], the authors proposed an iterative threshold selecting method to locate the ROIs, and the final ROI was selected based on the region size and the distance between the candidate region and the fixed reference point (image center). We can see that the inflexible constraints used in the above automatic ROI generation methods reduced their robustness. An ideal automatic ROI generation method should have the following characteristics:

1) **Fully automatic**. The operator-independent feature can avoid the subjective bias and tedious work caused by manual and semi-automatic methods.
2) **Complete coverage of tumor region**. It is the necessary condition for achieving high segmentation accuracy.

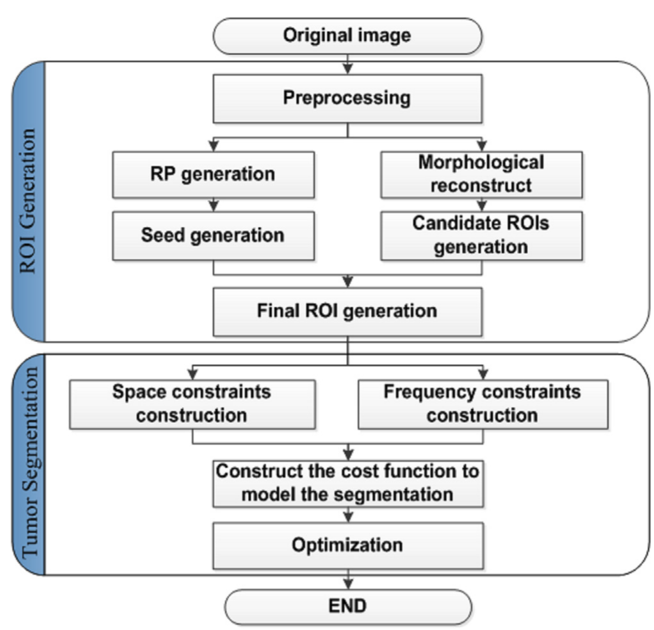

Fig. 1. Flowchart of the proposed segmentation method.