Automatic Seed Placement for Breast Lesion Segmentation on US Images

Joan Massich[1,⋆], Fabrice Meriaudeau[2], Melcior Sentís[3], Sergi Ganau[3], Elsa Pérez[4], Robert Martí[1], Arnau Oliver[1], and Joan Martí[1]

[1] Computer Vision and Robotics Group, University of Girona, Spain
jmassich@atc.udg.edu
[2] Laboratoire Le2i-UMR CNRS, University of Burgundy, Le Creusot, France
[3] Department of Breast and Gynecological Radiology, UDIAT-Diagnostic Center, Parc Taulí Corporation, Sabadell, Spain
[4] Department of Radiology, Hospital Josep Trueta of Girona, Spain

Abstract. Breast lesion boundaries have been mostly extracted by using conventional approaches as a previous step in the development of computer-aided diagnosis systems. Among these, region growing is a frequently used segmentation method. To make the segmentation completely automatic, most of the region growing methods incorporate automatic selection of the seed points. This paper proposes a new automatic seed placement algorithm for breast lesion segmentation on ultrasound images by means of assigning the probability of belonging to a lesion for every pixel depending on intensity, texture and geometrical constraints. The proposal has been evaluated using a set of sonographic breast images with accompanying expert-provided ground truth, and successfully compared to other existing algorithms.

Keywords: seed placement, ultrasound, segmentation, breast cancer.

1 Introduction

Breast cancer constitutes a leading cause of death for women in developed countries, and is most effectively treated when diagnosed at an early stage [8]. Digital Mammography is currently the most powerful screening tool for breast cancer [5], although ultrasound images can provide useful complementary information in cases where a tumor presence can be shielded due to dense glandular breast tissue [9]. Despite ultrasound imaging is a non-expensive and non-invasive technique with no side effects, its use in CAD systems is still under development. A feasible explanation is that performing automatic segmentation in US images is currently a challenge because they often suffer from poor quality and tend to generate artifacts: weak edges due to acoustic similarity between adjacent tissues, shadows as a consequence of the signal attenuation preventing to screen

---

any further, low contrast when the ultrasound wave is attenuated by the tissue media, or, speckle which is an unwanted collateral artifact produced by coherent interface of scatterers that appear as a granular structure superimposed on the image.

Among the reported techniques proposals for both guided and automatic segmentation of lesions in ultrasound images, region growing procedures that expand a seed accordingly to some criteria are widely used [6]. However, a proper selection of the seeds highly determines the final segmentation results.

The goal of this work is to compare three well known automatic procedures for selecting seed points [3,7,1] with a novel seed region selection methodology that makes use of texture and intensity features with geometric constraints. The experimental results have been obtained using a set of sonographic images with expert-provided ground truth, which have been tested using an already existing framework for segmenting breast lesions in ultrasound images [4].

2   Background

Given the noisy nature of the ultrasound images and the presence of other structures rather than lesions with similar acoustic properties, placing seed points on an ultrasound image with the aim to segment breast lesions is not a trivial task at all. Thus, an automatic seed placement procedure is usually required when dealing with fully automatic segmentation procedures. Three existing automatic seed placement procedures have been analyzed and tested according to their ability to later produce reliable segmented regions that match lesions:

– Pixel Rewarding (PR)[3]. To avoid manual delineation of the tumor boundaries, this proposal combines texture, intensity, gradient and a deformable model along with empirically determined domain specific knowledge to automatically find lesion margins in ultrasound images. Each pixel of the image is rewarded according to an assessment function using its position, intensity and texture. A recursive refinement stage removes outliers and provides a close estimate of the true boundary to a deformable model which produces the final segmentation. The deformable model operates on the directional gradient, making it more robust to noise. Its main advantage is its spatially constrained seed rewarding along with the fact that the lesion's appearance is obtained by means of a learning step. On the other hand, its major disadvantage remains in choosing an appropriated neighborhood for the term representing the probability mean of the surrounding pixels when calculating the pixel reward. If the neighborhood used is too small, it might incorrectly reward a noisy region; otherwise, if the used neighborhood is too large, a proper seed can be hidden due to its neighbors' low recall.
– Intensity Binarized Ranked Regions (IBRR) [7]. A score function to rank the regions not connected with the boundary or having intersection with the image center window is used, with no need of prior information of training

process. The function takes into account both the homogeneous texture features and the spatial features of the breast lesions.

– Gradient-Based (GB)[1]. After initial Radial Gradient Index [2] filtering, the lesion candidates are segmented from the background by maximizing an Average Radial Gradient (ARD) index for regions grown from the detected points. A round robin analysis to assess the quality of the classification of lesion candidates into actual lesions and false-positives by a Bayesian neural network is used, yielding to a good overall performance. The main drawback of this seed selection procedure is its associated computational cost, which has been partially solved by means of subsampling techniques. However, due to the comprehensive nature of the seed determination, the method remains unadvisable for online applications.

3   ITG: A Novel Seed Placement Methodology for Region Selection

Characterizing breast lesions by means of image analysis techniques usually combines intensity and texture as high specificity features [9]. Besides, it is a fact that radiologists tend to center the lesions when acquiring the images [3]. Thus, the proposed methodology makes use of Intensity, Texture and Geometric constraints (ITG) and takes advantage of the mentioned statements in order to select a seed region for further region growing expansion, as is shown in Figure 1. The proposal evaluates the probability of a pixel being part of a lesion depending on its intensity, texture and position to generate a joint probability or total probability plane.

Afterwards, the largest region composed by connected pixels with a posterior probability that satisfies the imposed confidence level of being a lesion is selected. In order to compute the posterior probability, a Bayesian framework is assumed accordingly to equation 1.

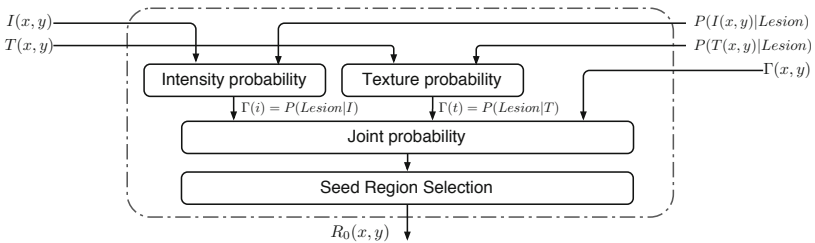$$P(Lesion|I,T) = \frac{P(I,T|Lesion) \cdot P(Lesion)}{P(I,T)} \tag{1}$$

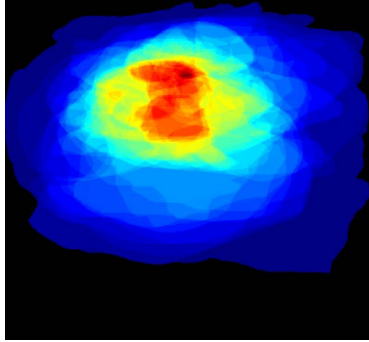**Fig. 1.** Block diagram describing the seed region selection proposal

Fig. 2. Lesion occurrence on a normalized grid, where the probability is represented according to a color code, from blue (lowest) to red (highest)

Where Intensity (I) and Texture (T) are two Independent and Identically Distributed (IID) features, and $P(Lesion)$ is assumed to be a centered multivariate Gaussian distribution proportional to the image. This is a reasonable assumption, since most of the lesions are centered as corroborates the probability map obtained from the dataset ground truth delineations (see fig. 2). Notice that the denominator $P(I, T)$ can be ignored since is common for the two classes $\{Lesion, \overline{Lesion}\}$ and cancels out. Thus, the final posterior probability can be calculated accordingly to equation 2 where $P(I|Lesion)$ and $P(T|Lesion)$ are the Intensity and Texture Probability Density Function (pdf) determined during the training step.

$$P(Lesion|I, T) = P(I|Lesion) \cdot P(T|Lesion) \cdot P(Lesion|x, y) \qquad (2)$$

The used texture measure is given by the Equation 3 and corresponds to the difference between the pixel intensity value $I(x, y)$ and the mean intensity value of its N nearest neighbors (here, 8-pixel neighborhood has been used).

$$T(x, y) = I(x, y) - \frac{1}{N} \sum_{\delta=0}^{N-1} I_\delta(x, y) \qquad (3)$$

Once determined the posterior probability, the probability plane is thresholded and the largest area from the foreground is selected as the seed region. The threshold has been empirically set at 0.8 as a good tradeoff between large foreground regions and low lesion belonging recall.

In summary, the proposed seed placement methodology makes use of five inputs to automatically determine a seed region: the intensity image, the texture image, the intensity and texture Probability Density Functions, and the seed location prior; along with a fixed parameter to split the probability plane into foreground and background.
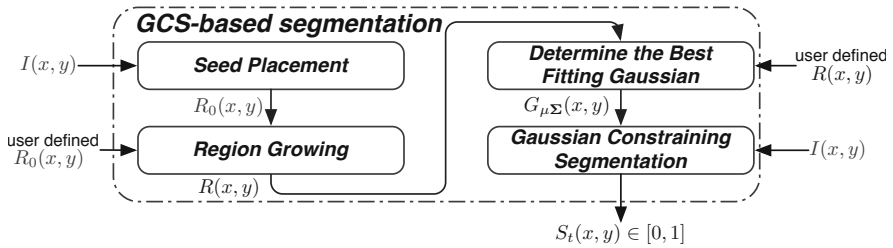
Fig. 3. Block diagram for the Gaussian Constraining Segmentation framework used to evaluate the proposal

4   Results

4.1   Experimental Setup

In order to evaluate the performance of the proposed methodology, a dataset of 25 sonographic images acquired at the *Hospital Dr. Josep Trueta* of Girona and the *UDIAT-Diagnostic Center* of Sabadell has been used. Since each image was annotated by seven radiology experts who provided the lesion delineations, the Simultaneous Truth and Performance Level Estimation (STAPLE) algorithm [10] to obtain the Hidden Ground Truth (HGT) has been used. The μ-coefficient proposed as a variance of the True-Positive Ratio (TPR) or Jaccard coefficient was then used in order to take into account the experts agreement by means of the HGT. The proposed ITG methodology along with the GB, PR, and IBRR procedures have been tested through the Gaussian Constraining Segmentation framework proposed by Massich et al. [4]. Figure 3 states the basic operations for such GCS-based segmentation framework: after an initial region $R_0(x, y)$ is determined, it is converted into a preliminary lesion delineation $R(x, y)$ by means of a region growing algorithm. Such lesion delineation is used to obtain a multivariate Gaussian function describing the shape, position and orientation of the lesion ($G_{\mu\Sigma}(x, y)$). Finally, the Gaussian Constraining Segmentation (GCS) procedure refines the segmentation by thresholding an intensity dependent function $\Psi(x, y)$ constrained by the multivariate Gaussian describing the lesion.

Figure 4 shows the segmentation results obtained for two clinical cases depending on the seed placement procedure. The blue delineation indicates the obtained seed region, while the red delineation indicates the Ground Truth, and the green delineation the obtained segmentation through the different region growing methods.

4.2   Seed Region Location

The effect of the initial seed position cannot be neglected when evaluating the performance of the proposed methodology. Thus, Figure 5a illustrates the ten Areas-of-Interest to test the influence of the lesion center distance and orientation
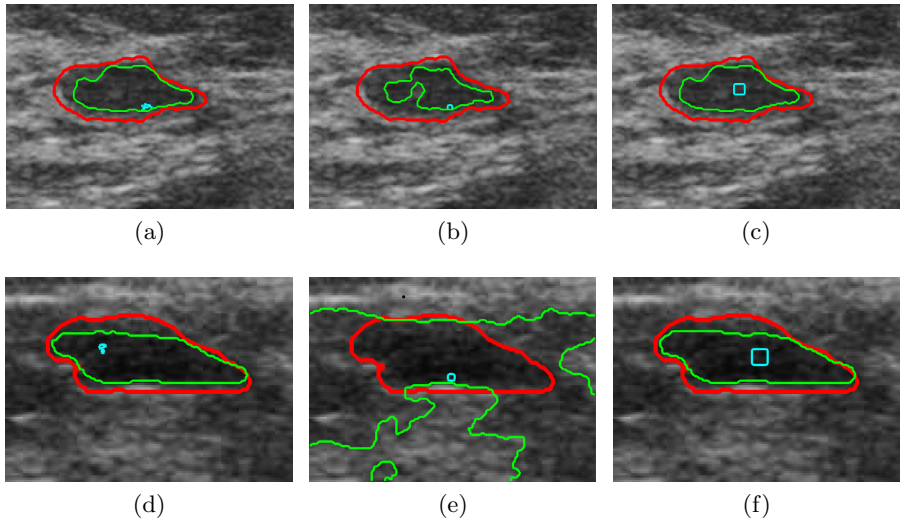
Fig. 4. Segmentation results: each row contains a clinical case (a-c,d-f), while each column corresponds to a different seed placement method: ITG (a,d), PR (b,e), and IBRR (c,f)

when seeding. The Areas-of-Interest have been selected as belonging to four different classes: out of the lesion (area 1), inside the lesion close to the boundaries (areas 2 to 5), inside the lesion but slightly shifted from the central part (areas 6 to 9), and central part of the lesion (area 10). For evaluation purposes, the region growing algorithm has been applied to each of the ten Areas-of-Interest using 15 randomly sampled seed regions for every area of interest. Figure 5b shows the segmentation results for each Area-of-Interest according to the μ value. It clearly shows that to achieve good segmentation results highly depends on the location of the seed regions within the lesion (the best segmentation results are achieved when placing the seed in the areas 6 to 10). The figure also indicates that three main classes a to c can be identified: (a) Areas-of-Interest 6 to 10 that correspond to the inner lesion area, (b) Areas-of-Interest 2 to 5 that correspond to the boundary area, and (c) Area-of-Interest 1 that corresponds to anywhere outside the lesion. The results indicates that the better segmentation results are achieved when the seed is placed in the (a) Areas-of-Interest (the inner lesion area, away from the boundaries), but not necessarily in the innest region.

4.3 Methodology Evaluation

Besides determining the role of the seed region location in terms of the achieved segmentation results, the performance of the proposed methodology has been also evaluated by comparing to the methods referred in section 2 (Pixel Rewarding, Intensity Binarized Ranked Regions, and Gradient-Based), as is shown in Figure 6: the first plot (Figure 6a) shows the ability of each methodology to place
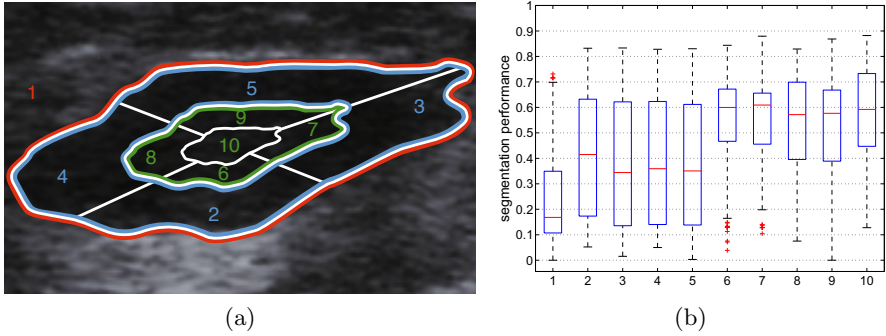
Fig. 5. (a) The 10 Areas-of-Interest to place seed regions, and (b) segmentation results for each Area-of-Interest in terms of the μ-coefficient
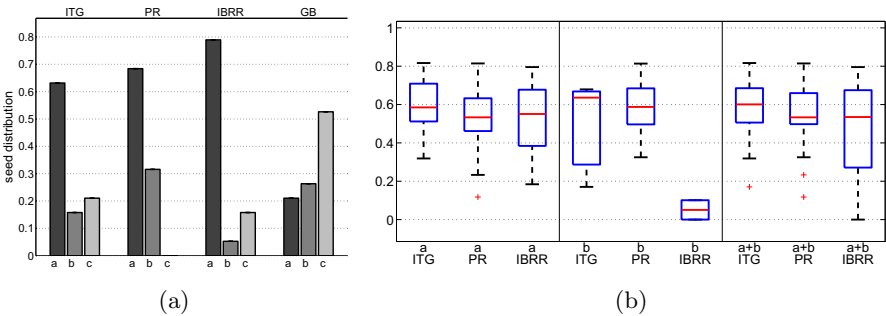

Fig. 6. Comparison between the proposed method ITG and the PR, IBRR, and GB methods: (a) distributions of seed region location, and (b) final segmentation performance depending on the seed location

the selected seed regions along the three main classes *a* to *c* for the Areas-of-Interest, while the boxplot (Figure 6b) indicates the mean and standard variation of the final segmentation results for each methodology. Although the PR and IBRR methods place more seeds in the central area than the ITG method, this new proposed method has the highest performance in terms of final segmentation results, as can be observed in Figure 6b.

5 Conclusions

The importance of a good seed placement for a region growing-based segmentation procedure has been stated. A new automatic seed placement algorithm for breast lesion segmentation on ultrasound images has been proposed. The proposal makes use of the intensity, texture and geometrical constraints to evaluate the probability of a pixel being part of the lesion. Performance of the new proposal has been successfully evaluated in terms of segmentation results on

a dataset of 25 sonographic images, and compared to three existing automatic procedures. Future work includes to assess the robustness of the new proposed methodology using a larger database.

References

1. Drukker, K., Giger, M.L., Kupinski, K.M.A., Vyborny, C.J., Mendelson, E.B.: Computerized lesion detection on breast ultrasound. Medical Physics 29(7), 1438–1446 (2002)
2. Kupinski, M.A., Giger, M.L.: Automated seeded lesion segmentation on digital mammograms. IEEE Transactions on Medical Imaging 17(4), 510–517 (1998)
3. Madabhushi, A., Metaxas, D.: Automatic boundary extraction of ultrasonic breast lesions. In: Proceedings of IEEE International Symposium on Biomedical Imaging, pp. 601–604 (2002)
4. Massich, J., Meriaudeau, F., Pérez, E., Martí, R., Oliver, A., Martí, J.: Lesion Segmentation in Breast Sonography. In: Martí, J., Oliver, A., Freixenet, J., Martí, R. (eds.) IWDM 2010. LNCS, vol. 6136, pp. 39–45. Springer, Heidelberg (2010)
5. Moore, S.K.: Better breast cancer detection. IEEE Spectrum 38(5), 50–54 (2001)
6. Noble, J.A., Boukerroui, D.: Ultrasound image segmentation: A survey. IEEE Transactions on Medical Imaging 25(8), 987–1010 (2006)
7. Shan, J., Cheng, H.D., Wang, Y.: A novel automatic seed point selection algorithm for breast ultrasound images. In: 19th International Conference on Pattern Recognition (2008)
8. Sivaramakrishna, R., Powell, K.A., Lieber, M.L., Chilcote, W.A., Shekhar, R.: Texture analysis of lesions in breast ultrasound images. Computerized Medical Imaging and Graphics 26(5), 303–307 (2002)
9. Stavros, A.T., Rapp, C.L., Parker, S.H.: Breast ultrasound. Lippincott Williams & Wilkins (2004)
10. Warfield, S.K., Zou, K.H., Wells, W.M.: Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. IEEE Transactions on Medical Imaging 23(7), 903–921 (2004)