



Automatic Segmentation of Breast Tumours on Ultrasound Images Using Genetic Algorithms and Morphological Operators

PACS: 43.80.Qf

Alvarenga, André V.^{1,2}; Pereira, Wagner C.A.²; Infantosi, Antonio F.C.²; Azevedo, Carolina M.³

¹Laboratory of Ultrasound, Division of Acoustics and Vibration Metrology, Directory of Scientific and Industrial Metrology, National Institute of Metrology, Standardization, and Industrial Quality, Duque de Caxias, Rio de Janeiro, Brazil; avalvarenga@inmetro.gov.br; victor@peb.ufrj.br

²Biomedical Eng. Program / COPPE, Federal University of Rio de Janeiro, Rio de Janeiro, Brazil; wagner@peb.ufrj.br; afci@peb.ufrj.br

³Radiology Department / Brazilian National Cancer Institute (INCa) Rio de Janeiro, Brazil; azevedocma@bol.com.br

ABSTRACT

Segmentation is known as a complex problem in ultrasound (US) images due to their textural nature, which results from factors like tissue properties, speckle, and artefacts. Trying to overcome this difficulty, different segmentation procedures have been described to support radiologists in recognizing abnormal areas on US images. This paper proposes a procedure, based on Genetic Algorithms (GA) and Morphological Operators (MO), to automatically segment breast tumours on ultrasound images. We assume each chromosome represents a complete image processing sequence, composed of MOs and structuring elements (SE). The GA procedure generates 500 chromosomes of 93 genes, where the three first genes summed indicate the number of operations to be performed, the following 45 genes represent the sequence of operations, and the last 45 the diameter, in pixel, of the disk used as SE. The best image processing sequence (best chromosome) is selected using a proposed objective function based on the difference between the areas of the tumour depicted by an experienced radiologist and the one defined by the segmentation procedure. A set of 20 images was used to train the GA procedure, and the best chromosome was tested with 40 different images. The performance of the segmentation procedure was assessed by comparing the computer-delineated tumour contour against the one drawn by a radiologist, using an overlap ratio. All tested images present an area overlap ratio superior to 70%. This initial result encourages us to go further by increasing the number of images and improving the objective function.

INTRODUCTION

Early diagnostics is essential to increase breast cancer therapy efficacy and with this aim mammography is routinely used as the screening modality [1]. However, a significant number of benign solid masses are considered as suspicious and recommended for biopsy [2]. Horsch et al. [3] point that 10 % to 30 % of masses submitted to surgical biopsy are malignant. Trying to improve the diagnostics and reduce the number of unnecessary biopsies, ultrasound (US) breast image has emerged as the most important adjunct to mammography for patients with palpable mass and inconclusive mammograms [4].

Malignant tumours generally infiltrate the surrounding tissue, resulting in irregular or undefined contour on the US image, thus, from the tumour contour it is possible to establish a diagnostic hypothesis [5,6]. Besides, many of the US contour characteristics, such as spiculation, angular margin, and microlobulation have been described and are well recognized as related to malignant tumours [4-8].

Several researchers [5,9,10] have used the contour manually delineated by an experienced radiologist, even when computer-aided methods are proposed. However, operator-dependent procedures are highly susceptible to variations [4,7,8,11]. Hence, despite of the speckled nature

of sonograms, which makes image segmentation a formidable task, various researchers have developed segmentation procedures to support radiologists in recognising abnormal areas on US breast images. With this aim, different techniques as average radial derivative [3], watershed [12], deformable shape-based [13-15], and morphological operators (MOs) [16] have been applied.

The key to successful morphological image processing is to select to right sequence of MOs and respective structuring elements (SE). There is a great number of possible MO sequences, depending on the imaging application. The classical approach to morphological processing is based on the researcher experience, intuition and perception of the goals to design the algorithms based on MOs. However, the human expertise can only try a very limited number of combinations. On the other hand, a genetic algorithm (GA) procedure may test many different ways of combining MOs. The search performed by GA is not random; it is guided by the fitness of MOs sequence in the population. As the search progresses, GA gradually shifts the population to the portion of the space containing the best MO sequences.

In this work, we use GA in selecting the sequence of MOs, and their respective SEs, to automatically segment breast tumours on ultrasound images. Contours depicted by an experienced radiologist were used as gold standard, and the parameter overlap ratio (*RS*) was used as figure of merit to assess the performance of the proposed segmentation method.

MATERIAL AND METHODS

Database

Using a 7.5-MHz linear array B-mode 40-mm ultrasound probe (Sonoline – Sienna ©Siemens) with axial and lateral resolutions of 0.45 mm and 0.49 mm, respectively, 60 breast tumour US images were acquired on the Cancer National Institute (Brazil - Rio de Janeiro). For each image, a rectangular region of interest (ROI) was determined by a radiologist (Fig. 1a), and each tumour contour was manually depicted (Fig. 1b). The database was divided into a training set containing 20 samples, and a test set containing 40 ones. The 20 training images were chosen to be representative of the diversity found in the dataset.



Figure 1.- (a) Example of a ROI and the (b) tumour contour manually depicted by a radiologist.

Segmentation procedure

The proposed segmentation procedure assumes that each chromosome, which represents a image processing sequence, is composed by 93 genes, where the three first genes summed ($g_1+g_2+g_3$) indicate the number of operations to be performed, the following 45 genes ($g_4 - g_{48}$) represent the sequence of operations, and the other 45 ($g_{49} - g_{93}$) the diameter, in pixel, of the disk used as SE. The values associated to genes that represent image operations and SE diameter range from 0 to 15, defining a specific operation (Table I), and from 1 to 3, respectively. An example of an initial population with three chromosomes is presented in Fig. 2. Note that for the second chromosome $g_1+g_2+g_3$ indicates a total of 20 operations, hence, the numbers between g_4 and g_{23} indicate the operations to be performed and the genes g_{49} to g_{68} represent the respective diameter of the disk used as SE. In this work, the number of chromosomes and the number of generations are set as 500 and 200 correspondingly.

The proposed objective function (f_o) is defined as:

$$f_o = \sum_k \frac{Area |S_r(k) - S_o(k)|}{S_r(k)} * 100 + NFP(k)_{pixel} + NFN(k)_{pixel}, \quad (\text{Eq. 1})$$

where k is the number of images in the training group, NFP_{pixel} is defined as the number of pixels outside the reference region (S_r), depicted by a radiologist, that is included in the segmented region (S_o), and NFN_{pixel} pixel is the number pixel in S_r that is not present within S_o .

Table I.- List of the sixteen operations used in the segmentation procedure.

No.	Operation
0	Erosion of A
1	Dilatation of A
2	Reconstruction by erosion of A
3	Reconstruction by dilatation of A
4	Negative of A
5	Do nothing with A
6	Add A and B
7	Subtract A from B
9	Intersection between A and B
9	Union between A and B
10	Binarization of A
11	Minima imposition of A
12	Operator to remove structures of A that touch image boundaries
13	Opening by reconstruction of A
14	Closing by reconstruction of A
15	Remove the smallest structures from image A

g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9	...	g_{45}	g_{46}	g_{47}	g_{48}	g_{49}	g_{50}	g_{51}	g_{52}	g_{53}	g_{54}	...	g_{89}	g_{90}	g_{91}	g_{92}	g_{93}
10	3	5	10	8	5	12	9	8	...	4	10	7	9	1	2	3	2	1	2	...	3	1	2	2	1
5	15	0	0	3	7	12	15	2	...	2	4	12	1	2	1	2	3	2	1	...	2	1	2	3	2
12	11	7	12	10	13	10	9	1	...	3	12	8	3	3	2	1	1	2	3	...	2	1	2	3	3

Figure 2.- Example of three chromosomes representing an initial population. The three first genes summed ($g_1+g_2+g_3$) indicate the number of operations to be performed, the following 45 genes ($g_4 - g_{48}$) represent the sequence of operations, and other 45 ($g_{49} - g_{93}$) the diameter of the disk used as SE.

The GA algorithm was implemented in MATLAB® (Mathworks Inc., Natick, MA) and based on the following steps:

1. Randomly create an initial population of chromosomes, where the genes between g_1 and g_{48} (# of operations + respective operations) and g_{49} to g_{93} (SE diameters) are integer numbers ranging from 0 to 15 and 1 to 3, respectively;
2. Each chromosome, which represents a MO sequence, is applied to the images in training set.
3. The objective function (f_o) is calculated to each chromosome;
4. Ninety-five percent of chromosomes with the lowest values of f_o are selected to generate the offspring by (a) reproduction, (b) crossover and (c) mutation operations;
 - (a) Reproduce an existing chromosome by copying it into the new population.
 - (b) Create two new chromosomes from two existing ones by using the crossover operation at a randomly chosen starting point.
 - (c) Create a new chromosome from an existing one by randomly mutating a gene.
5. Calculates f_o from offspring.
6. Reinsert 25% of the best chromosomes from offspring to the population;
7. If the convergence criterion has not been reached, return to step 2;
8. The chromosome that has given the optimum value is designated as the one containing the best MOs sequence and respective diameters to SEs disk.

The segmentation procedure performance was assessed using the overlap ratio (RS), defined as:

$$RS = \frac{Area(S_r \cap S_o)}{Area(S_r \cup S_o)} \quad (\text{Eq. 2})$$

where the symbols \cap and \cup indicate the areas intersection and union, respectively. So if the areas have the same shape and size and are in the same position, the overlap ratio is the unity.

RESULTS

The best sequence of operations and respective SEs, defined by the GA procedure using 20 images, is presented on Table II. This operations sequence was applied to another set of 40 images, and the result of the proposed method (red contour) applied to different images are presented in Figure 3. For the sake of comparison the tumour contours delineated by the radiologist are also depicted (yellow contour).

The tumour shown in Figure 3a presents well-defined boundaries, and its automatic (Figure 3b - red) and manually-defined (Figure 3b - yellow) contours are very similar. However, when a tumour presents shadowed edges (Figure 3c) the radiologist tends to overestimate tumour margins (Figure 3d – yellow). In that case, the automatic-depicted contour tends to be smaller than the manually-depicted one.

In the histogram of Figure 4, it can be clearly noted that all the tested images have $RS > 0.7$ and 65% of them with $RS > 0.8$.

Table II.- Best sequence of operations selected by the GA procedure.

Best operation sequence	14	5	0	6	3	3	1	4	3	10	12	1	0	8	14	15
Respective SE diameters	3	2	3	3	1	3	3	2	2	1	3	3	2	3	3	3

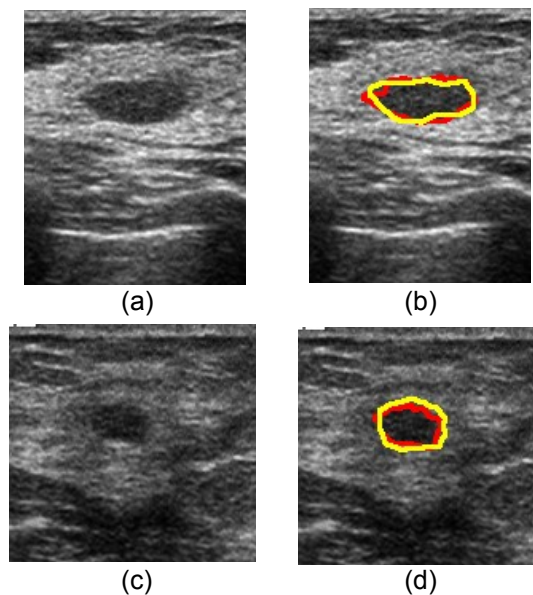


Figure 3.- Original images (a & c) and respective contours depicted using the proposed segmentation procedure (red), and by radiologist (yellow).

DISCUSSION & CONCLUSION

The proposed segmentation procedure, which used genetic algorithm in selecting the sequence of Morphological Operators, has shown to be adequate to segment breast tumour on ultrasound images. All 40 test image presented RS values superior to 0.70, and an optimal performance ($RS > 0.8$) was obtained for 65% of them. The initial results points out that this automatic method constitutes an improvement compared to the semi-automatic approach described in [16]. However, it will be necessary to test this procedure using a larger number of images, and to compare results quantitatively.

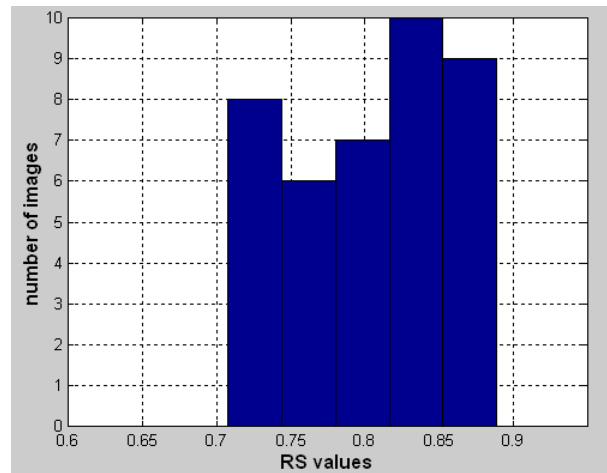


Figure 4.- RS values for the 40 breast tumours segmented.

References: [1] <http://www.inca.gov.br/estimativa/2006/>

[2] M.A. Dennis, S.H. Parker, A.J. Klaus, A.T. Stavros, T.I. Kaske, S. B. Clark: Breast biopsy avoidance: the value of normal mammograms and normal sonograms in the setting of a palpable lump. *Radiology* **219** (2001) 168–191.

[3] K. Horsh, M. L. Giger, L. A. Venta, C. J. Vyborny: Automatic segmentation of breast lesions on ultrasound. *Medical Physics* **28** (2001) 1652–1659.

[4] P. Skaane: Ultrasonography as adjunct to mammography in the evaluation of breast tumors. *Acta Radiologica Supplementum* **40** (1999) 1–47.

[5] Y. Chou, C. Tiu, G. Hung, S. C. Wu, T.Y. Chang, H.K. Chiang: Stepwise Logistic Regression Analysis of Tumour Contour Features for Breast Ultrasound Diagnosis,” *Ultrasound in Medicine & Biology* **27** (2001) 1493–1498.

[6] S.Y. Chiou, Y.H Chou, H.J. Chiou, H.K. Wang, C.M. Tiu, L.M. Tseng, C.Y. Chang: Sonographic features of nonpalpable breast cancer: a study based on ultrasound–guided wire–localized surgical biopsies. *Ultrasound in Medicine & Biology* **32**, No.9, (2006) 1299–1306.

[7] S. Huber, J. Danes, I. Zuna, J. Teubner, M. Medl, S. Delorme: Relevance of sonographic B–mode criteria and computer–aided ultrasonic tissue characterization in differential diagnosis of solid breast masses. *Ultrasound in Medicine & Biology* **26** (2000) 1243–1252.

[8] G. Rahbar, A. C. Sie, G. C. Hansen, J. S. Prince, M. L. Melany, H. E. Reynolds, V. P. Jackson, J. W. Sayre, L. W. Basset: Benign versus malignant solid breast masses: US differentiation. *Radiology* **213** (1999) 889–894.

[9] C. Chen, Y. Chou, K. Han, G. Hung, C. Tiu, H. Chiou, S. Chiou: Breast lesions on sonograms: computer–aided diagnosis with nearly setting–independent features and artificial neural networks. *Radiology* **206** (2003) 504–514.

[10] F. Lefebvre, M. Meunier, F. Thibault, P. Laugier, G. Berger: Computerized ultrasound B–scan characterization of breast nodules. *Ultrasound in Medicine & Biology* **26** (2000) 1421–1428.

[11] R. G. Barr, “Breast ultrasound: a bright future”, *Medica Mundi* **45**, 8–13 (2001)

[12] Y. Huang, D. Chen: Watershed segmentation for breast tumor in 2–D sonography. *Ultrasound in Medicine & Biology* **30** (2004) 625–632.

[13] A. Madabhushi, D.N. Metaxas: Combining low–, high–level and empirical domain knowledge for automated segmentation of ultrasonic breast lesions. *IEEE Trans. Med. Imaging* **22** (2003) 155–169.

[14] Y. Huang, Y. Jiang, D. Chen, W. K. Moon: Level Set Contouring for Breast Tumor in Sonography. *Journal of Digital Imaging* **20** (2007)

<http://www.springerlink.com.w10002.dotlib.com.br/content/ct028tvj2j02u42k/fulltext.html>.

- [15] R. F. Chang, W. J. Wu, W. K. Moon, D. R. Chen: Automatic ultrasound segmentation and morphology based diagnosis of solid breast tumors. *Breast Research and Treatment* **89** (2005) 179–185.
- [16] A. V. Alvarenga, A. F. C. Infantosi, W. C. A. Pereira, C. M. Azevedo: Application of morphological operators on the segmentation and contour detection of ultrasound breast images. *Brazilian Journal of Biomedical Engineering* **19**, No.2 (2003) 91–101.