AN EFFICIENT SEGMENTATION METHOD FOR ULTRASOUND IMAGES BASED ON A SEMI-SUPERVISED APPROACH AND PATCH-BASED FEATURES

A. Ciurte[1,2], N. Houhou[2], S. Nedevschi[1], A. Pica[3], F.L. Munier[4], J.-Ph. Thiran[2], X. Bresson[5], M. Bach Cuadra[2]*

[1]Department of Computer Science, Technical University of Cluj-Napoca (UTCN), Roumania
[2]Signal Processing Laboratory (LTS5), Ecole Polytechnique Fédérale de Lausanne, Switzerland
[3]Radiation Oncology Department, Lausanne University Hospital (CHUV), Switzerland
[4]Ophthalmic Hospital Jules Gonin, Lausanne, Switzerland
[5]Computer Science Department, City University of Hong Kong, Hong Kong

ABSTRACT

Segmenting ultrasound images is a challenging problem where standard unsupervised segmentation methods such as the well-known Chan-Vese method fail. We propose in this paper an efficient segmentation method for this class of images. Our proposed algorithm is based on a semi-supervised approach (user labels) and the use of image patches as data features. We also consider the Pearson distance between patches, which has been shown to be robust w.r.t speckle noise present in ultrasound images. Our results on phantom and clinical data show a very high similarity agreement with the ground truth provided by a medical expert.

Index Terms— Ultrasonography, retinopathy, image segmentation, active shape model, bipartite graph.

1. INTRODUCTION

Despite the potential high resolution that ultrasound (US) images can provide in comparison with other medical imaging modalities such as Magnetic Resonance imaging (MRI) or Computer Tomography (CT), US image quality is actually much lower because of speckle noise, shadows, attenuation and signal dropout. These artifacts often lead to weak (or missing) edges and also to the presence of fake edges, making general segmentation methods fail. Even so, image-guided therapy planning and navigation, and computer aided diagnosis make increasing use of US as main image modality. Indeed, US imaging is of particular interest in areas such echocardiography, obstetrics and gynecology, breast cancer and intravascular diseases [1]. Recently, the use of ultrasound imaging is increasing in ophthalmology [2]. The ultimate goal of this work is to delineate precisely eye anatomical structures in 3D B-scan ultrasound images. This segmentation problem is part of a larger project that aims at improving the radiotherapy planning and treatment of retinoblastoma in childhood by fusion of CT, ultrasound and fundus of photograph [3].

A wide range of image features are used for US in the literature (please refer to [1] for an extended review). Intensity features are often used by means of analytical models of the gray-level distribution (Rayleigh distribution being the most used model [4] but also Fisher-Tippet, Exponential, Gamma, Gaussian and other distributions are used in the literature). Note that all these models are well-suited for the received echo signal while, in practice, clinical ultrasound devices log-compress the signal before visualization [5]. The main drawbacks of these gray-level distribution models are related to distribution parameter estimation and to its dependency on the imaging system settings such as dynamic range and gain. Intensity gradient features can also be used [6]. However, they strongly assume that there are homogeneous regions and thus, a very low level of speckle noise. Texture features have also proved to be successful in US segmentation, particularly statistical patterns of the intensity due to their advantage of being independent of the imaging system physics [7]. Added to image features, prior information is still needed for an accurate segmentation of US images. Thus, shape priors are often included, as well as the regularization term inside an energy functional in variational methods [8] or statistical parametric shape priors [9].

In this work we present a semi-supervised segmentation method for US imaging using image patches as features. The idea of image patches as feature was first introduced for texture synthesis [10] and image denoising [11] in natural images. In [12], Coupé et al. extend the non-local means filter [11] to reduce speckle noise in US images. In [13, 14], the patch-based approach was also successfully used for spatio-temporal registration and for motion or/and elasticity estimation in US image sequences of the heart.

The proposed segmentation strategy is built on our previous work for interactive segmentation of textural images [15]. However, we consider here a different distance between patches based on [12], which is robust against speckle noise

*This work is supported by PRODOC Project of UTCN and by the Center for Biomedical Imaging (CIBM) of the Geneva-Lausanne Universities and the EPFL, as well as the foundations Leenaards and Louis-Jeantet.

in ultrasound images.

2. METHOD

The proposed method assumes solving an energy minimization problem:

$$u^* = \underset{u}{\operatorname{argmin}}\, E(u), \tag{1}$$

whose equilibrium state u^* corresponds to the segmentation solution. The proposed energy function $E(u)$ for this problem is the one introduced in [15]:

$$E(u) = \int_{\Omega\times\Omega} |\nabla_w u|^2 dxdy + \int_{\Omega} \lambda(x)(u-f)^2 dx + \beta TV(u), \tag{2}$$

where the first two energy terms correspond to the continuous min-cut formulation [16] with label constraints (see [15] for more details). The last energy term is the classic TV regularization term [17], that smoothes contour irregularities.

2.1. Continuous Min-cut

The proposed method uses non-local information, represented by a graph. The pixels in the image I are considered to be the nodes of the graph and a pair of pixels x and y are considered to be connected to each other, according to the value of the weight $w_{x,y}$ defined as:

$$w(x,y) = \begin{cases} \frac{|x-y|_2}{\sigma_1^2} + \frac{d(F(x),F(y))}{\sigma_2^2}, & \text{for } x,y \in N_{a\times a}(x) \\ 0, & \text{otherwise} \end{cases}, \tag{3}$$

where $N_{a\times a}(x)$ is a square window of size $a \times a$ around x, $F(x)$ is a square patch of size $b \times b$ centered on x, $d(\cdot,\cdot)$ is the distance between two different patches and σ_1, σ_2 are the scaling parameters. The parameters a, b, σ_1 and σ_2 are set in Section 2.5. In [15], the authors use the mean square difference as a patch distance measure in the context of natural images segmentation. This approach was originally developed for additive white Gaussian noise. For ultrasound segmentation however, we need to take into account the presence of the speckle noise. In this aim, we introduce the Pearson distance to measure the similarity between patches, previously used in [12] for speckle removal in ultrasound denoising problems. The formulation of the Pearson distance is made under the assumption of the following general speckle model: $I(x) = \tilde{I}(x) + \sqrt{\tilde{I}(x)}\eta(x)$, where $\tilde{I}(x)$ is the original image, $I(x)$ is the observed image and $\eta(x) \sim N(0,\sigma^2)$ is Gaussian noise. Thus, for each pixel x can be made the assumption that $I(x)|\tilde{I}(x) \sim N(\tilde{I}(x), \tilde{I}(x)\sigma^2)$. Using a Bayesian formulation over the proposed model (defined in [12]), the Pearson distance can be defined as follows:

$$d(F(x),F(y)) = \frac{1}{2}\sum_{k=1}^{B} \frac{(F^{(k)}(x) - F^{(k)}(y))^2}{F^{(k)}(y)}, \tag{4}$$

where $F(x) = (F^{(1)}(x), \ldots, F^{(B)}(x))$, $B = b \times b$ elements, and $F^{(k)}(x)$ the intensity of k^{th} element of the patch centered in x.

With the graph formulation previously described, we further consider the segmentation problem as a bi-partitioning of the graph into two disjoint subsets A_1 and A_2, while minimizing the subsets inter-similarity, known as min-cut or graph flow. The graph partition energy is known to be equivalent to the graph Laplacian operator (see e.g. [18]) defined as:

$$cut(A_1,A_2) = \frac{1}{8}\sum w_{xy}(\tilde{u}(x) - \tilde{u}(y))^2 \tag{5}$$

where $\tilde{u}(x)$ equals to -1 if $x \in A_1$ and 1 if $x \in A_2$. Let $u(x) = \frac{\tilde{u}(x)+1}{2}$, so that $u(x) \in [0,1]$. Thus, Eq. 5 can be written in the continuous form [19] as follows:

$$\frac{1}{2}\int\int_{\Omega\times\Omega} w(x,y)(u(x)-u(y))^2 dxdy = \frac{1}{2}\int\int_{\Omega\times\Omega} |\nabla_w u|^2 dxdy, \tag{6}$$

where $\nabla_w u$ is the graph-based gradient of u.

2.2. Semi-supervision with labels

The 2^{nd} term of Eq. 2 developed in [20] introduces hard label constraints given by the function $f \in [0,1]$. This information is given by the user, assuming that for some pixels in the image are known to belong to the object ($f = 1$) or to the background ($f = 0$) with a degree of confidence $\lambda(x)$. For example, a user can select manually this region/s, with some clicks, which will correspond to the initialization of the segmentation process.

2.3. Minimization process

In order to use fast numerical schemes, energy (2) can be regularized as follows [15]:

$$E(u,v,s) = \int_{\Omega\times\Omega} |\nabla_w u|^2 dxdy + \int_{\Omega} \lambda(x)(v-f)^2 dx + \\ +\beta TV(s) + \int_{\Omega} \frac{1}{2\theta_v}(u-v)^2 dx + \int_{\Omega} \frac{1}{2\theta_s}(s-v)^2 dx, \tag{7}$$

where $v, s \in [0,1]$ are the coupling terms, that are forced to be almost equal to u when $\theta_v \to 0$ and $\theta_s \to 0$.
The solution of (7) is reached by an iterative process performed successively over u, v and s as follows:

$$u^{n+1} = \frac{\theta_v \int_\Omega w(x,y)u^n(y)dy + v^n(x)}{\theta_v \int_\Omega w(x,y)dy + 1},$$

$$v^{n+1} = \begin{cases} \frac{\theta_s u^{n+1} + \theta_v s^n}{\theta_s + \theta_v}, & \text{if } \lambda = 0 \\ f, & \text{if } \lambda = \infty \end{cases},$$

$$s^{n+1} = v^{n+1} - \theta div p^n,$$

where p is given by [21]:

$$p^{n+1} = \frac{\beta(p^n + 1/8\nabla(\text{div}p^n - v^{n+1}/\theta_s))}{\beta + 1/8|\nabla(\text{div}p^n - v^{n+1}/\theta_s)|}.$$

2.4. Initialization

Our algorithm is designed in a flexible framework which allows both, manual and automatic initialization. For the manual initialization, the user is required to make some clicks to define the object (see the red contour in second row of Fig. 2, corresponding also to the initial segmentation contour), and, optionally, the background labels (see the blue contour in second row of Fig. 2). We propose also an automatic initialization for the specific segmentation of the vitreous body in order to segment the tumor plus retina structures in US eye images. Such automation is performed by introducing priors regarding anatomical characteristics of the eye such as its ellipsoidal shape. We first define a region of interest as the inside of the eyeball, everything else being consider as background. The eyeball segmentation is done by model-based registration, as described in [3], by combining 3D CT data with the anatomical information stored in an anatomical eye model. Then, since retinoblastoma is always located on the back of the eye, the half front part of the eye is considered the *object label* (see first row of Fig. 3) for the segmentation of vitreous body. The retina region (healthy plus pathologic) is thus defined as the complementary area of the vitreous body inside the eyeball (see second row of Fig. 3).

2.5. Parameter setting

The searching window, a, influences the speed of convergence: the higher a, the faster the convergence. There is a though compromise between the computational time to build the weighted matrix w and the time of convergence itself. The patch size, b, depends on the texture to be segmented, so that the patch contains a complete texture pattern. We notice a coarseness of the segmentation result and a loss of details for high b values. The scaling parameter, σ_1^2, is empirically fixed to be equal to $3b^2$ [15]. Finally, σ_2 is chosen such that enough points are selected for diffusion and that only pertinent points are selected (low value of σ_2 implies that only the best pixels are kept). In the further experiments we use the following settings for the studied parameters: $a = 7$, $b = 3$ and $\sigma_2 = 0.004$. *Our experiments show that good results are achieved for the same parameters in different subjects, structures, and image sequences*.

3. VALIDATION

Our clinical data set consists of 3 children with retinoblastoma (calcified and non calcified tumors). The ophthalmic 2D ultrasound images used here are slices from 3D volumes acquired with OTI Ophthalmic Technologies Inc. (OTI), as in first row of Fig. 1, with isotropic resolution of 0.1 mm and image size of 206×217 pixels. Second row of Fig. 1 shows the manual labeling of structures of interest: vitreous and lens (in yellow), the tumor (in red), the retina (in green), the optic nerve (in cyan). We apply the classical Active contours without edges (ACWE) [22] (first row of Fig. 2) as a reference for comparison. Third row of Fig. 2 and second row of Fig. 3 show the results for our method with manual and automated initialization respectively. Visual inspection shows that ACWE is not adapted for the challenging US images. On the contrary, our proposed segmentation method provides a very high similarity agreement with the ground truth, for both manual and automatic initializations. Quantitative evaluation on the 3 tumor segmentation is presented in Table 1 in comparison with manually delineated structures. Dice similarity measure (DSM) is shown for tumor only with manual initialization and for tumor plus retina and vitreous for the automated initialization. Our proposed segmentation method performs quite well because a DSM > 0.7 is considered as a good agreement with the ground truth.

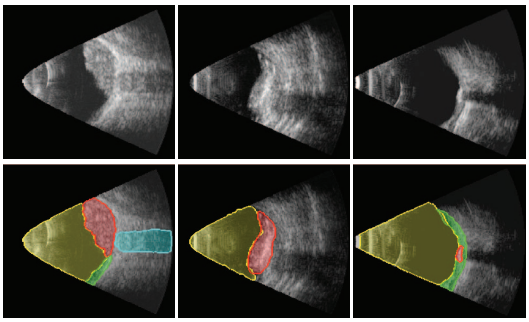

Fig. 1: Clinical data: original images (top) and ground truth (GT) by manual labeling (bottom).

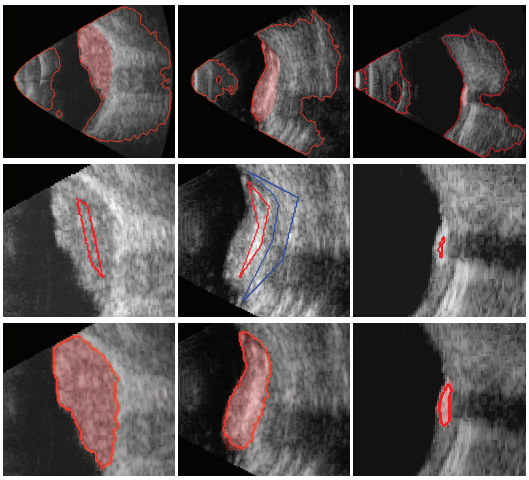

Fig. 2: Tumor segmentation (tumor GT is filled with red labels). 1^{st}, row ACWE results [22]; 2^{nd} row, manual initialization; 3^{rd} row, results of our algorithm.

Comparison of the Gaussian model [15] against the speckle model proposed here is shown by means of the optic nerve segmentation in Fig. 4. Visual inspection shows very similar results. DSM for speckle model being slightly better than for Gaussian model (0.9352 and 0.9288, respectively).

Results on breast tumor phantom data are shown in Fig. 5.

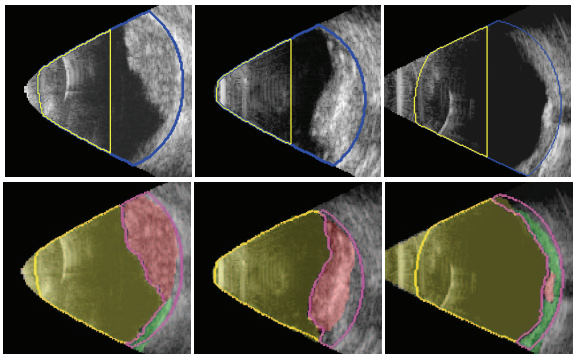

Fig. 3: Automated segmentation. Top: zoom on automatic initialization; bottom: zoom on vitreous body (yellow contour) and tumor+retina (pink contour).

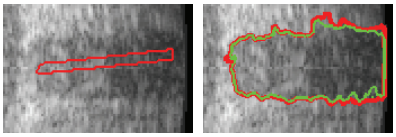

Fig. 4: Optic nerve segmentation: left, initialization, right, Gaussian noise model [15] in red and speckle model in green.

US phantom images were acquired with a HI VISION^{TM} 900 US system and present weak borders to be segmented. Both Gaussian and Speckle noise models show again high similarly and good results, visually and quantitatively, with the same DSM for both phantoms (0.95 for top and 0.94 for bottom phantom). Our algorithm is implemented in Matlab© and it takes in average 300 seconds with 5000 iterations in a INTEL Core 2 Duo, 2.66 GHz, 2 GB of RAM.

4. CONCLUSION AND FUTURE WORK

Our semi-supervised segmentation method based on the Pearson distance between patches produces similar results than the Euclidean distance between patches like in [15], both of providing a high similarity agreement with ground truth (in both phantom and clinical data). This may indicate that the most important elements of the proposed algorithm are the semi-supervised approach (with the user labels) and the use of

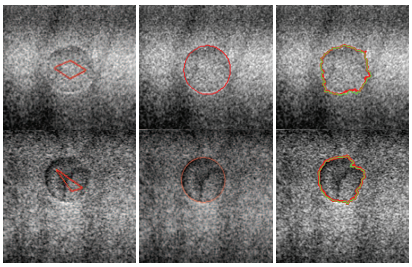

Fig. 5: Tests on 2 tumor breast phantom images (top and bottom); Gaussian model is in red, Speckle model is in green.

	Tumor	Tumor+Retina	Vitreous Body
Case 1	0.9625; 20.11	0.8941; 23.96	0.9726; 46.77
Case 2	0.9045; 16.92	0.7766; 16.92	0.9778; 49.30
Case 3	0.7487; 1.22	0.7442; 15.40	0.9312; 84.66

Table 1: Dice similarity measure (DSM) and surface in mm^2. DSM > 0.7 is considered a good agreement with the ground truth.

image patches as data features. Despite the fact that the proposed method specifically takes into account speckle noise, it seems that the choice of the distance between patches does not seem to influence too much the results, at least for this level of noise. Future work will investigate the robustness of the model w.r.t. the speckle noise. We will also further assess both Gaussian and Speckle noise models in comparison with some other methods used for US segmentation. Finally, an extension to 3D and a better optimization method will be developed.

5. REFERENCES

[1] J. A. Noble, D. Boukerroui, *Ultrasound Image Segmentation: A Survey*, IEEE Trans. on Medical Imaging, 25(8) (2006), 987-1010.

[2] A. Fenster, D. Downey, H. Cardinal, *Three-dimensional ultrasound imaging*, Phys. Med. Biol., vol 46(5), (2001) 67-99.

[3] M. Bach Cuadra et al. *Model-based Segmentation and Fusion of 3D Computed Tomography and 3D Ultrasound of the Eye for Radiotherapy Planning*, Computational Vision and Medical Image Processing, (2009) 53-58.

[4] J. Anquez, E.D. Angelini, I. Bloch, *Segmentation of fetal 3D ultrasound based on statistical prior and deformable model,* ISBI08, (2008) pp.17-20.

[5] Zhong Tao; H.D. Tagare, J.D. Beaty, *Evaluation of Four Probability Distribution Models for Speckle in Clinical Cardiac Ultrasound Images,* Medical Imaging, IEEE Trans. on, 11 (25), (2006) 1483-1491.

[6] C. Corsi, G. Saracino,A. Sarti, C. Lamberti, *Left ventricular volume estimation for real-time three-dimensional echocardiography,* Trans. on Medical Imaging, vol.21(9), (2002) 1202-1208.

[7] D. R. Chen, R. F. Chang, W. J. Kuo, M. C. Chen, and Y. L. Huang, *Diagnosis of breast tumors with sonographic texture analysis using wavelet transform and neural networks,* Ultrasound Med. Biol., vol. 28(10), (2002) 1301-1310.

[8] N. Paragios, M.-P. Jolly, M. Taron, R. Ramaraj, *Active Shape Models and Segmentation of the Left Ventricle in Echocardiography*, Lect. Notes in Computer Science, vol. 3459, (2005) 131-142.

[9] Y. Chen, F. Huang, H.D. Tagare, M. Rao, *A Coupled Minimization Problem for Medical Image Segmentation with Priors,* Int. J. of Computer Vision, vol. 71(3), (2007) 259-272.

[10] A.A. Efros, T.K. Leung, *Texture Synthesis by Non-parametric Sampling*, IEEE International Conference on Computer Vision, (1999).

[11] A. Buades, B. Coll, J. Morel, *A review of image denoising algorithms, with a new one*, Multiscale Modeling & Simulation, vol. 4(2), (2005) 490-530.

[12] P. Coupé, P. Hellier, C. Kervrann, C. Barillot, *Nonlocal means-based speckle filtering for ultrasound images*, IEEE Trans. Image Process, 18(10), (2009) 2221-29.

[13] M.J. Ledesma-Carbayo et al., *Spatio-temporal nonrigid registration for ultrasound cardiac motion estimation*, IEEE Trans. on Medical Imaging, vol.24(9), (2005) 1113-1126.

[14] J. Revell, M. Mirmehdi, D. McNally, *Computer vision elastography: speckle adaptive motion estimation for elastography using ultrasound sequences,* IEEE Trans. on Medical Imaging, vol.24(6), (2005) 755-766.

[15] N. Houhou, X. Bresson, A. Szlam, T.F. Chan, J.-Ph. Thiran, *Semi-Supervised Segmentation based on Non-local Continuous Min-Cut*, Scale Space and Variational Methods in Computer Vision, vol. 5567, (2009) 112-123.

[16] G. Strang, *Maximal flow through a domain* , Mathematical Programming, vol.26, (1983) 123-143.

[17] L.I. Rudin, S. Osher, E. Fatemi, *Nonlinear Total Variation Based Noise Removal Algorithms*, Physica D, vol. 60(1-4), (1992) 259-268.

[18] J. Shi, J. Malik, *Normalized cuts and image segmentation,* Pattern Analysis and Machine Intelligence, IEEE Trans. on , vol.22(8), (2000) 888-905.

[19] G. Gilboa, S. Osher, *Nonlocal operators with applications to image processing*, Multiscale Model. Simul, (2008).

[20] M. Unger, T. Pock, D. Cremers, H. Bischof, *Tvseg - interactive total variation based image segmentation*, British Machine Vision Conf., Leeds, UK (2008).

[21] A. Chambolle, *An algorithm for Total Variation minimization and applications*, J. of Mathematical Imaging and Vision., vol.20(1-2), (2004) 89-97.

[22] T.F. Chan, L.A. Vese, *Active Contours Without Edges*, IEEE Trans. on Image Processing, vol.10(2), (2001) 266-277.