SEGMENTATION OF MEDICAL ULTRASOUND IMAGES USING ACTIVE CONTOURS

Oleg Michailovich

Department of ECE
University of Waterloo
Waterloo, ON, Canada N2L 3G1

Allen Tannenbaum

School of ECE
Georgia Institute of Technology
Atlanta, GA, USA 30332

ABSTRACT

Segmentation of medical ultrasound images (e.g., for the purpose of surgical or radiotherapy planning) is known to be a difficult task due to the relatively low resolution and reduced contrast of the images, as well as due to the discontinuity and uncertainty of segmentation boundaries caused by speckle noise. Under such conditions, useful segmentation results seem to be only achievable by means of relatively complex algorithms, which are usually computationally involved and/or require *a prior* learning. In this paper, a different approach to the problem of segmentation of medical ultrasound images is proposed. In particular, we propose to *preprocess* the images before they are subjected to a segmentation procedure. The proposed preprocessing modifies the images (without affecting their anatomic contents) so that the resulting images can be effectively segmented by relatively simple and computationally efficient means. The performance of the proposed method is tested in a series of both *in silico* and *in vivo* experiments.

Index Terms— Image segmentation, morphological signal processing, decorrelation, active contours, and medical ultrasound.

1. INTRODUCTION

Medical ultrasound imaging is currently considered to be on the leading edge in non-invasive medical imaging. Its cost-benefit ratio in terms of accessibility, portability, and safety far exceeds that of alternative technologies such as X-ray CT and MRI. The clinical use of medical ultrasound imaging has long expanded beyond the simple visualization and inspection of anatomic structures. Surgical and radiotherapy planning, intra-operative navigation, and tracking the progress of disease are nowadays among the standard applications of this imaging modality.

All the aforementioned applications of medical ultrasound involve the procedure of segmentation of ultrasound images. Although modern ultrasound scanners are capable of providing fairly good views of internal anatomy, the use of *automatic* segmentation techniques for accurately quantifying and analyzing the embedded structures seems to be limited. The relatively low resolution and reduced contrast of ultrasound images along with the discontinuity and uncertainty of segmentation boundaries caused by speckle noise are among the main reasons that make the task of segmenting ultrasound images difficult to carry out.

A multitude of methods for segmentation of ultrasound images have been proposed hitherto, most of which take advantage of *deformable models* [1]. Thus, for example, in [2] the shape of human prostate is described by a discrete model represented by a finite set of vertices and corresponding facets. Subsequently, the model is evolved under internal (elastic) and external (image-related) forces, and the evolution is terminated when the model "locks on" the local maxima of image gradient, which are presumed to be related to the prostate boundary. In [3], the performance of a similar segmentation approach is improved via incorporation of *a priori* information on possible shapes of the prostate. In this case, the model is described by a linear combination of the principal components of prostatic shapes which have been depicted by a radiologist beforehand. Apart from the necessity of prior learning, this method (as well as the previous one) seems to have the disadvantage of being dependent on the information contained in image edges, which are known to be unreliable features to use in ultrasound imagery. This fact was addressed in [4], where the segmentation method discriminates the prostatic tissue based on texture-related features via a classification procedure.

The watershed transformation was applied to the problem of segmenting ovarian ultrasound images in [5]. Although demonstrating reliable performance in detecting the shape of follicles, it seems to be impossible for the method to be extended to other organs, as it is based on the assumption that the follicle interior regions should be noticeably darker than the rest of the ultrasound image area.

In [6], the problem of segmentation of intravascular ultrasound images is addressed via introducing a knowledge-based approach, which incorporates a number of explicit constrains stemming from the known anatomy of blood vessels and their associated geometry. On contrary, no *a priori* geometric assumptions are used in the method of [7], where the three-dimensional shapes of blood vessels are represented by active surfaces. However, this method requires the boundaries of internal and external laminas to be well observable.

Deformable models have also proven useful in the field of cardiac elastography. Thus, for example, in [8], echocardiograms are segmented based on active appearance motion models, which are capable of accounting for both shape and intensity variations of myocardium during the cardiac cycle. The technique of active contours was used in [9] for left ventricular segmentation. As opposed to [8], this methods does not involve a training stage. However, it also seems to be useful for segmenting the object of interest with well observable boundaries.

The present study demonstrates that the performance of segmentation algorithms can be substantially improved via applying the latter to *preprocessed* rather than original images. In particular, the proposed method segments ultrasound images in the logarithmic domain, after the images have been subjected to the processes of decorrelation and outlier shrinkage. It is shown that this preprocessing is capable of considerably improving the separability of segmentation classes, thereby increasing the accuracy of resulting segmentation. The method by means of which the segmentation is performed is based on the technique of *active contours*. In particular, we adopt the

region-based formulation of active contours as introduced in [10], which seems to be advantageous in the situations, when objects of interest have poorly observable boundaries due to, e.g., dropouts and/or shadowing artifacts.

2. IMAGE FORMATION MODEL

Under the assumption of linear wave propagation and weak scattering, the backscattered signal and the tissue reflectivity function are known to obey a simple Fourier transform relationship with respect to each other. In this case, a radio-frequency (RF) image is considered to be result of the convolution of the point-spread function (PSF) of the imaging system with the tissue reflectivity function. Denoting by $g(n,m)$, $f(n,m)$, and $h(n,m)$ the RF-image, the tissue reflectivity function, and the PSF, respectively, the convolution model is formally given by:

$$g(n,m) = f(n,m) * h(n,m) + u(n,m), \tag{1}$$

where n and m denote the axial and lateral (or radial and angular, for B-scan sector images) indices of the image samples. The term $u(n,m)$ is added to describe measurement noises as well as all the physical phenomena, which are not accounted for by the convolution model.

Due to the natural intricacy of most biological tissues and the fact that tissue heterogeneity is generally formed by numerous small "independent" structures, the samples of the reflectivity function can be assumed to be uncorrelated. Moreover, these samples are likely to be independent of noise, which can be reasonably assumed to be Gaussian and white. In this case, the power density $P_g(\omega_1,\omega_2)$ of the RF-image can be shown to be given by:

$$P_g(\omega_1,\omega_2) = \sigma_f^2 \left|H(\omega_1,\omega_2)\right|^2 + \sigma_u^2, \tag{2}$$

where $H(\omega_1,\omega_2)$ is the Fourier transform of the PSF $h(n,m)$, while σ_f^2 and σ_u^2 denote the variances of the reflectivity function and noise, respectively.

From (2), it is clear that the correlation properties of ultrasound images are defined by the form of the PSF, which is typically a band-limited function. It goes without saying that these correlation properties have to be taken into account, while designing a segmentation algorithm. Alternatively, one can try to find an operator that can transform a received RF-image into another RF-image, whose samples correlate less than those of the original. Such an operator can be defined as a linear filter $l(n,m)$, whose transfer function is given by [11]:

$$L(\omega_1,\omega_2) = \frac{1}{\sqrt{|H(\omega_1,\omega_2)|^2 + \sigma_u^2/\sigma_f^2}}. \tag{3}$$

Applying (3) to an RF-image results in "flattening" its power spectrum and, therefore, reduces the correlation between the RF-image samples. The constant $\lambda = \sigma_u^2/\sigma_f^2$ can be thought of as a tunable parameter which can be set empirically, so that its optimal value would result in maximal decorrelation, while avoiding undesirable artifacts caused by over-amplification of high frequencies. In order to implement the spectrum equalization using the filter (3), the power spectrum of the PSF needs to be estimated. In the current study, the estimation is performed using the method described in [11, Section C].

The decorrelation of RF-images constitutes the first stage of the proposed preprocessing. After this stage is completed, the (decorrelated) RF-images are subjected to the standard process of frequency demodulation and envelope detection. Subsequently, the second preprocessing stage is applied to the resulting envelope images.

3. SPECKLE NOISE AND OUTLIER SHRINKAGE

A coherent nature of the formation of ultrasound images causes the latter to be severely contaminated by speckle noise. Thus, before deriving any segmentation procedure, the properties of the noise should be carefully analyzed. In ultrasound imaging, however, a universally agreed upon definition of a model for speckle noise still seems to be lacking. One of the most frequently used models (whose feasibility has been verified via its extensive practical use) represents the noise as a multiplicative process. Specifically, the observed envelope image $p(n,m)$ is modeled as:

$$p(n,m) \simeq d(n,m)\,\xi(n,m), \tag{4}$$

where $d(n,m)$ is the *true* image and $\xi(n,m)$ is the speckle noise.

Working with multiplicative noises is generally considered to be more difficult than the additive noise case. Thus, in order to convert $\xi(n,m)$ into an additive noise, *homomorphic processing* can be used that takes advantage of the logarithmic transformation applied to both parts of (4). Denoting the logarithms of $p(n,m)$, $d(n,m)$, and $\xi(n,m)$ by $p_l(n,m)$, $d_l(n,m)$, and $\xi_l(n,m)$, respectively, the measurement model in the log-transform domain becomes:

$$p_l(n,m) \simeq d_l(n,m) + \xi_l(n,m). \tag{5}$$

Let us assume, for the moment, that the samples of the reflectivity function behave as realizations of a zero-mean, white Gaussian noise of variance σ_f^2. In this case, it is straightforward to show that the samples of $\xi(n,m)$ obey a Rayleigh distribution, and, consequently, the log-transformed speckle noise $\xi_l(n,m)$ is distributed according to the Fisher-Tippett probability density given by:

$$P(x) = 2\exp\left\{(2x - \log 2\sigma_f^2) - \exp\left\{2x - \log 2\sigma_f^2\right\}\right\}. \tag{6}$$

It is interesting to note that the Fisher-Tippett distribution possesses an approximant in the form of a Gaussian density. The latter is obtained by replacing the inner exponent in (6) by the first three terms of its Taylor series expansion, resulting in:

$$P(x) \simeq 2e^{-1}\exp\left\{-\frac{1}{2}\left(\frac{x - \log\sqrt{2}\sigma_f}{0.5}\right)^2\right\}. \tag{7}$$

An interesting fact about the approximation above is that it has a constant variance of 0.25, implying that the additive noise in (5) may be roughly viewed as white Gaussian with a *fixed* variance. Moreover, the variance σ_f^2 of the underlying reflectivity function defines the *mean value* of the distribution. This implies that the biological tissues represented by reflectivity functions of different variances can be discriminated based on their mean intensity in the log-trasnform domain of corresponding envelope images. Note that the reasonability of such a representation of biological tissues was recently advocated in [12].

Unfortunately, the approximation (7) is acceptable only in close proximity of the mean value of the original *pdf* (6). Moreover, as compared to the Gaussian, the *pdf* of the Fisher-Tippett distribution is asymmetric and leptokurtic. As a result, a noise "produced" by the Fisher-Tipett distribution may be viewed as a white Gaussian noise contaminated by occasional transients or outliers. Unfortunately, the latter can significantly bias estimating the mean value of the Gaussian approximation (7).

The above difficulty can be overcome via "gaussianization" of $\xi_l(n,m)$ that consists of estimating and subsequently subtracting its

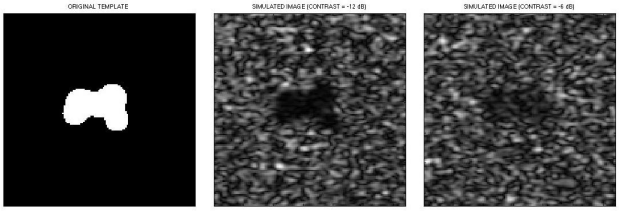

Fig. 1. (*Left*) Original template; (*Center*) Simulated envelope at -12 dB contrast; (*Right*) Simulated envelope at -6 dB contrast.

spiky component. The latter can be estimated as robust residuals of $p_l(n,m)$ computed according to:

$$r(n,m) = \text{sign}\left(\Delta p_l(n,m)\right)\ \left(|\Delta p_l(n,m)| - \lambda\right)_+ . \qquad (8)$$

Here $\Delta p_l(n,m)$ denotes difference between $p_l(n,m)$ and its median-filtered version, λ is a predefined threshold, and the operator $(x)_+$ returns x if $x > 0$ and zero otherwise. It was observed in [11] that in most cases, the robust residuals $r(n,m)$ correspond to the outliers of the spiky noise, when the size of the median filter is set to be 3×3 (or 5×5) and the threshold λ is set to a level such that 93-95% of the differences $|\Delta p_l(n,m)|$ do not exceed the predefined threshold. In this case, subtracting $r(n,m)$ from $p_l(n,m)$ results in suppressing the spiky component of the Fisher-Tippett noise. Moreover, the noise contaminating the difference signal $p_l(n,m) - r(n,m)$ behaves very similarly to white Gaussian noise, while $d_l(n,m)$ remains practically unchanged.

Rejecting the outliers of the log-transfromed speckle noise constitutes the second and the last stage of the proposed preprocessing. Note that, though the consideration above relied on the assumption that the samples of $f(n,m)$ are Gaussian distributed, it can be shown that (8) can be effectively used for "gaussianization" of more general types of noises [11].

4. SEGMENTATION BY ACTIVE CONTOURS

In order to facilitate the discussion, we confine the derivations below to the case of two classes and use a continuous rather than discrete formulation. In this case, the segmentation problem is reduced to the problem of partitioning the domain of definition $\Omega \subset \mathbb{R}^2$ of an image $p(z)$ (with $z \in \Omega$) into two mutually exclusive and complementary subsets Ω_- and Ω_+. These subsets can be represented by their respective characteristic functions χ_- and χ_+, which can, in turn, be defined by means of a *level-set function* $\phi(z) : \Omega \to \mathbb{R}$ in the following manner. Let $\mathcal{H}$ be the Heaviside function defined in the standard way as:

$$\mathcal{H}(z) = \begin{cases} 1 & \text{if } z \geq 0; \\ 0 & \text{if } z < 0. \end{cases} \qquad (9)$$

Then, one can define $\chi_-(z) = \mathcal{H}(-\phi(z))$ and $\chi_+(z) = \mathcal{H}(\phi(z))$, with $z \in \Omega$.

Given a level-set function $\phi(z)$, its *zero level set* $\{z \mid \phi(z) \equiv 0, z \in \Omega\}$ is used to implicitly represent a curve – *active contour* – embedded into Ω. For the sake of concreteness, we associate the subset Ω_- with the support of the object of interest, while Ω_+ is associated with the support of corresponding background. In this case, the objective of active-contour-based image segmentation is, given an initialization $\phi_0(z)$, to construct a *convergent* sequence of level-set functions $\{\phi_t(z)\}_{t>0}$ (with $\phi_t(z)|_{t=0} = \phi_0(z)$) such that

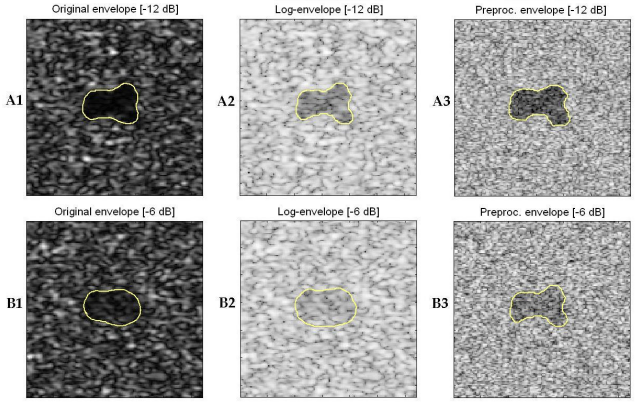

Fig. 2. (*A1-A3*) Segmentation of the original, log-, and preprocessed envelopes at -12 dB constrast; (*B1-B3*) Segmentation of the original, log-, and preprocessed envelopes at -6 dB constrast.

the zero level-set of $\phi_\infty(z)$ coincides with the boundary of the object of interest.

The above sequence of level-set functions can be constructed using the variational framework. In particular, in the current study, the optimal level set function is found as a minimizer of the following functional:

$$\mathcal{E}(\phi(z), c_-, c_+) = \int_\Omega |p(z) - c_-|\, \mathcal{H}(-\phi(z))\, dz + \qquad (10)$$
$$+ \int_\Omega |p(z) - c_+|\, \mathcal{H}(\phi(z))\, dz + \alpha \int_\Omega \| \bigtriangledown\, \mathcal{H}(\phi(z))\|\, dz.$$

In (10), the first two terms penalize the deviation of the image values from the mean intensities c_- and c_+ inside and outside of the active contour, respectively. At the same time, the last terms controls the smoothness of the active contour, and hence $\alpha > 0$ can be viewed as a regularization parameter.

The convergent sequence $\{\phi_t(z)\}_{t>0}$ is obtained as a gradient flow stemming from minimization of (10) as detailed in [10] *mutatis mutandis*. Note that the functional given by (10) can be thought of as a *robust* version of the functional minimized in [10].

5. EXPERIMENTAL RESULTS

The experimental study presented in this paper consists of two parts. In the first part, we tested the performance of the proposed segmentation method under controllable conditions using simulated data sets. The latter were generated using the "tissue" structure shown in the left subplot of Fig. 1, where the white region can be thought of as, e.g., a cyst, while the black region represents surrounding tissue. The ratio between the variances of the reflectivity functions corresponding to the "cyst" and the background were set to be equal to 0.25 and 0.5, thereby resulting in envelope images with the contrast ratios of -12 dB and -6 dB, respectively. The central and right subplots of Fig. 1 show two examples of the simulated envelopes computed using the Field II ® package for different contrasts. Note that segmentation is expected to reconstruct the boundary of the "cyst" region from corresponding RF-image.

To demonstrate the usefulness of the proposed preprocessing, the active-countor-based segmentation of Section 4 was applied to the original envelopes, their log-transformed versions, and to the preprocessed envelopes. Subplots A1-A2 of Fig. 2 show realizations of

Table 1. NMSE of segmentation of different types of envelope images

	Original	Logarithmic	Preprocessed
-12 dB	0.19	0.16	0.07
-6 dB	0.31	0.28	0.17

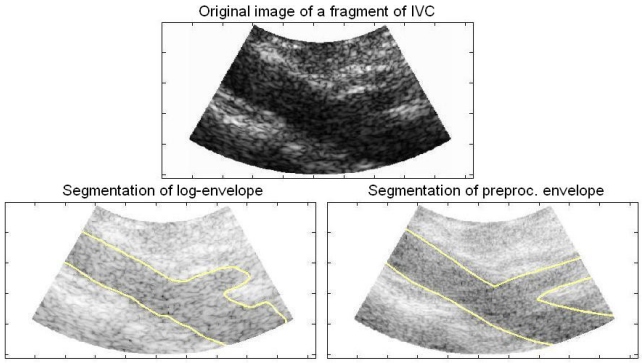

Fig. 3. (*Top*) Original image of IVC; (*Left*) Segmentation of the log-envelope; (*Right*) Segmentation of the preprocessed envelope.

the above envelopes along with the final active contours for the case of -12 dB contrast. The results obtained for the case of -6 dB contrast are exemplified in Subplots B1-B3 of the same figure. One can see that, in comparison with the original and log-transformed envelopes, the preprocessed images have noticeably better resolution and higher contrast. Consequently, segmenting the preprocessed images results in much more accurate estimation of the "cyst" geometry, even when the "cyst" is virtually indistinguishable (see Subplots B1-B3 of Fig. 2).

In order to quantitatively compare the segmentation results, the *normalized mean-squared error* (NMSE) was used, which is defined as $\mathrm{NMSE} = \mathcal{E}\left\{\|\mathrm{p_0} - \mathrm{p_e}\|_\mathrm{F}^2 / \|\mathrm{p_0}\|_\mathrm{F}^2\right\}$, where p_0 and p_e denote an original image and its estimate, and the subscript F stands for the Frobenius norm. The NMSE obtained in the simulation study are summarized in Table 1 for different types of the envelopes and contrast ratios. One can see that the preprocessing results in a substantial reduction of the NMSE values.

During the second stage of the experimental study, the proposed segmentation method was evaluated using *in vivo* data. The data were recorded from adult volunteers with a VIVID-3 scanner (GE Medical Systems, Inc.) using a 4 MHz, curved-array probe.

The upper subplot of Fig. 3 shows *in vivo* ultrasound image of a fragment of inferior vena cava (IVC). One can see that the vessel boundaries are discontinuous and poorly observable due to contrast variations and low resolution. For these reasons, segmenting the log-envelope image, as demonstrated by the leftmost subplot of Fig. 3, results in a far-from-perfect reconstruction of the geometry of the IVC. On the other hand, the vessel geometry is virtually perfectly reconstructed by segmenting the corresponding preprocessed image, as shown in the rightmost subplot of the same figure. In general, it was observed that the segmentation of preprocessed images agreed with the manual segmentation performed by a radiologist within 8% error bounds, while for the unprocessed images this error was about 21%.

6. CONCLUSIONS

In this paper, a novel approach to the problem of segmentation of medical ultrasound images has been proposed. It is proven conceptually and experimentally, that the performance of image segmentation can be substantially improved via properly preprocessing the images to be segmented. A distinctive feature of the proposed preprocessing lies in the fact that it allows substantially improving the separability of segmentation classes without affecting the contents of the images being segmented. This feature makes it possible for the preprocessing to be used in combination with a variety of segmentation tools.

7. REFERENCES

[1] T. McInerney and D. Terzopoulos, "Deformable models in medical image analysis: A survey," *Medical Image Analysis*, vol. 1, no. 2, pp. 91–108, 1996.

[2] A. Ghanei, H. Soltanian-Zadeh, A. Ratkewicz, and F. Yin, "A three-dimensional deformable model for segmentation of human prostate from ultrasound images," *Med. Phys.*, vol. 28, no. 10, pp. 2147–2153, 2001.

[3] D. Shen, Y. Zhan, and C. Davatzikos, "Segmentation of prostate boundaries from ultrasound images using statistical shape model," *IEEE Trans. Med. Imag.*, vol. 20, no. 4, pp. 539–551, 2003.

[4] W. D. Richard and C. G. Keen, "Automated texture-based segmentation of ultrasound images of the prostate," *Comput. Med. Imaging Graph.*, vol. 20, no. 3, pp. 131–140, 1996.

[5] A. Krivanek and M. Sonka, "Ovarian ultrasound image analysis: Follicle segmentation," *IEEE Trans. Med. Imag.*, vol. 17, no. 6, pp. 935–944, 1998.

[6] M. Sonka, X. Zhang, M. Siebes, M. S. Bissing, S. C. DeJong, S. M. Collins, and C. R. McKay, "Segmentation of intravascular ultrasound images: A knowledge-based approach," *IEEE Trans. Med. Imag.*, vol. 14, no. 4, pp. 719–732, 1995.

[7] J. D. Klingensmith, R. Shekhar, and D. G. Vince, "Evaluation of three-dimensional segmentation algorithms for the identification of luminal and medial-adventitial borders in intravascular ultrasound images," *IEEE Trans. Med. Imag.*, vol. 19, no. 10, pp. 996–1011, 2000.

[8] J. G. Bosch, S. C. Mitchell, B. P. F. Lelieveldt, F. Nijland, O. Kamp, M. Sonka, and J. H. C. Reiber, "Automatic segmentation of echocardiograms sequences by active appearance motion models," *IEEE Trans. Med. Imag.*, vol. 21, no. 11, pp. 1374–1383, 2002.

[9] C. Corsi, G. Saracino, A. Sarti, and C. Lamberti, "Left ventricular volume estimation for real-time three-dimensional echocardiography," *IEEE Trans. Med. Imag.*, vol. 21, no. 9, pp. 1202–1208, 2002.

[10] T. Chan and L. Vese, "Active contours without edges," *IEEE Trans. Image Processing*, vol. 10, no. 2, pp. 266–277, 2001.

[11] O. Michailovich and A. Tannenbaum, "De-speckling of ultrasound images," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 53, no. 1, pp. 64–78, 2006.

[12] O. Husby, T. Lie, T. Lango, J. Hokland, and H. Rue, "Bayesian 2-D deconvolution: A model for diffuse ultrasound scattering," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 48, no. 1, pp. 121–130, 2001.