

Combinative Multi-scale Level Set Framework for Echocardiographic Image Segmentation

Ning Lin¹, Weichuan Yu², and James S. Duncan^{1,2}

¹ Department of Electrical Engineering, Yale University

² Department of Diagnostic Radiology, Yale University

PO Box 208042, New Haven, CT 06520-8042, USA

{James.Duncan, ning.lin}@yale.edu, weichuan@noodle.med.yale.edu

Abstract. In the automatic segmentation of echocardiographic images, *a priori* shape knowledge is used to compensate poor features in ultrasound images. The shape knowledge is often learned via off-line training process, which requires tedious human effort and is unavoidably expertise-dependent. More importantly, a learned shape template can only be used to segment a specific class of images with similar boundary shapes.

In this paper, we present a multi-scale level set framework for echo image segmentation. We extract echo image boundaries automatically at a very coarse scale. These boundaries are then not only used as boundary initials at finer scales, but also as an external constraint to guide contour evolutions. This constraint functions similar to a traditional shape prior. Experimental results validate this combinative framework.

1 Introduction

Echocardiography is a widely used imaging technique in clinical diagnosis of heart disease. In order to improve the diagnosis performance and to reduce the dependency of human expertise, it is desired to automatically estimate important indices such as left ventricle (LV) deformation from echo images directly. This requires a reliable, automatic segmentation of LV boundary. However, the ultrasound images are always accompanied with degradations including intensity inhomogeneity [1], distortion, and speckle noise [2] which cause the failure of simple image feature-based thresholding methods. Currently, reasonable segmentation results are obtained mainly using tedious, interactive methodology. Automatic segmentation of echo images still remains a challenging topic.

Some groups have attacked this topic by modeling the physical principle of ultrasound imaging degradations. Different filtering methods and statistic models are presented to correct the intensity inhomogeneity [1] and reduce speckle noise [3–5]. It has been shown that the performance of image feature-based thresholding methods will be improved after the correction and de-noising. But the evaluation of these physical and statistic models still needs to be addressed.

Some other groups proposed to segment original echo images directly. They used sophisticated algorithms to combine *a priori* shape knowledge [6], texture information [7], spatio-temporal continuity of neighboring images [8, 9], and motion information

[10]. Among these features, *a priori* shape knowledge has been proven to be a powerful constraint in the segmentation of noisy images [11]. Usually, the shape knowledge is learned via interactive training process. Though training process can take place off-line, a lot of expertise effort is still needed. More importantly, a learned shape template can only be used to segment a specific class of images with similar boundary shapes.

The main contributions of this paper are: 1. we demonstrate that a region- and edge-based level set method can be used to segment echo images at a coarse scale. 2. The boundary shape is used in a multi-scale analysis framework **not only** as boundary initials at finer scales, **but also** as an additional constraint to guide contour evolutions. Note that the term *boundary shape* here does not mean a probabilistic template prior. Rather, it is the form of a deterministic boundary.

The rest of the paper is organized as follows: Section 2 explains our combinative multi-scale level set framework in detail. Section 3 presents validation and comparison experiments. Section 4 concludes the paper.

2 Combinative Multi-scale Level Set Framework

Before we describe our method, it is worth reviewing the state-of-the-practice of echo imaging technique for a better understanding of our motivation. 3D echo (3DE) imaging was introduced to provide 3D volume and surface information of the heart with greater accuracy than 2D echo (2DE) imaging [5]. Though the newest 3DE system uses a linear-array transducer to get a 3D volume which consists of almost parallel image planes, the popular imaging system nowadays uses a transducer rotating around a fixed axis to get a set of 2DE planes at different angles. When we use a rotating transducer, we have to face the problem of reconstructing 3D volume or surface from a set of 2D image planes. As simple interpolation methods either introduce artifacts or leave the gap between two neighboring planes unfilled (specially at some distance from the rotating axis), it would be logical to segment 2D image planes at first and then construct a 3D boundary surface or a 3D volume using 2D boundaries [12]. At the same time, it is worth mentioning that automatic segmentation of 2DE images is still an unsolved problem. Based on this thought and the available data, we focus on 2DE segmentation in this paper, though it would be straightforward to extend our method to 3DE segmentation once the linear-array transducer is available.

In the following, we list our algorithm and then explain each step in detail.

1. For each 2DE image, construct a Gaussian pyramid.
2. Initialize the level set inside the LV chamber at the coarsest scale level of the Gaussian pyramid.
3. Apply region homogeneity and edge-based level set method to find out the LV chamber boundary.
4. Interpolate the boundary to a finer scale level. Use edge-based level set constrained by boundary similarity to refine the contour.
5. If the finest scale level has been reached, stop.
Otherwise, go to step 4.

2.1 Intensity Distribution in Gaussian Pyramid

The intensity distribution of echo images is not clear yet, but it is definitely not Gaussian. This explains why additive Gaussian noise model performs poorly in the segmentation of echo images. This situation changes, however, at higher levels of Gaussian pyramid (the reader is referred to [13] for details about Gaussian pyramid). The high-frequency noise is smoothed out with the increase of pyramid level. Besides, the neighboring pixels at higher levels are more likely to be independent as subsampling reduces their correlation. According to Central Limit Theorem, the gray values of these pixels after Gaussian smoothing may be approximated as a Gaussian distribution. Therefore, we may use the additive Gaussian noise model at the highest pyramid level to extract boundaries. At lower pyramid levels, though the additive Gaussian noise model is not valid, the boundary shapes remain similar, as shown in figure 1. In the next subsection, we will use this shape similarity as an additional constraint in the level set methods.

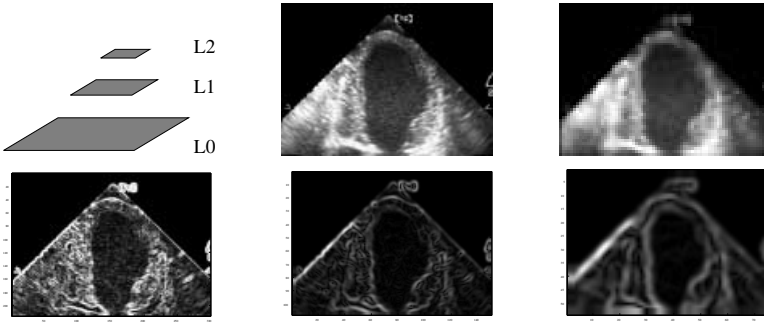


Fig. 1. Top: The principle of Gaussian pyramid algorithm (left). The original image (middle) is denoted as level L0. A 5×5 2D Gaussian kernel is used for smoothing before subsampling the image. The expanded image L2 (right) is smoother than the original image, while the boundary shapes are similar at both L0 and L2. **Bottom:** The derivative amplitudes of L0 (left), of L1 (middle), and of L2 (right). The disturbance of speckle noise decreases with the increase of pyramid level. At the same time, the boundary of the LV chamber is blurred as well.

2.2 Level Set Methods at Different Scale Levels

In this subsection, we explain the reason of using different level set methods at step 3 and step 4 in our algorithm. We also give the details of the level set implementation.

Finding a contour can be described as an energy minimization problem. For example, finding a 2D contour \mathcal{C} in an image $I(x, y)$ using edge information is equivalent to minimizing the following energy function

$$F_1 = \int_0^1 g(I(\mathcal{C}(q))) |\mathcal{C}'(q)| dq, \quad (1)$$

here $g(I) = \frac{1}{1+|\nabla I|^2}$ is a monotonously decreasing function of image gradients and $|C'(q)|dq$ denotes the unit length of the contour (see [14] for details).

For echo images at different pyramid levels, solely edge-based detection techniques may not work due to weak edges (cf northeast direction of L2 in figure 1). Thus, we need additional constraints to guide the segmentation. The *Active Contour without Edge* method [15] shows promising results in segmenting images with weak edges. This method is based on the Gaussian noise assumption and is therefore an appropriate constraint for echo images at very coarse scales, as we mentioned in section 2.1.

In an image $I(x, y)$ satisfying the Gaussian noise model, assume that we have a foreground object and a background with intensity mean values c_1 and c_2 , respectively. According to [15], the object contour \mathcal{C} can be found by minimizing

$$F_2 = \frac{1}{2\sigma_1^2} \iint_{\text{inside}(\mathcal{C})} |I(x, y) - c_1|^2 dx dy + \frac{1}{2\sigma_2^2} \iint_{\text{outside}(\mathcal{C})} |I(x, y) - c_2|^2 dx dy. \quad (2)$$

Here σ_1 and σ_2 denote the standard deviations of gray values inside and outside the contour, respectively. It is worth mentioning that c_i and σ_i ($i = 1, 2$) are normally independent, while in echo images we have an extra relation: $c_1/\sigma_1 \approx c_2/\sigma_2$ [16]. This relation actually simplifies the evaluation of F_2 , since we can replace σ_2 with $\sigma_1 c_2/c_1$.

Recently, Paragios and Deriche [17] further unified boundary and region-based information for tracking purpose. This combination increases the robustness of the level set method. Here we adopt their method for echo image segmentation, yielding the following minimization framework

$$\min_{\mathcal{C}} \{F_1 + \lambda_1 F_2\}. \quad (3)$$

Here λ_1 is a weighting coefficient to adjust the relative importance of region information with respect to edge information.

After extracting boundaries at the coarsest scale level, we gradually interpolate them to finer levels as initial contours and refine them using local image features. In the refining process, the *active contour without edge* method (cf equation (2)) can no longer be used because the gray value distribution at lower levels of the pyramid cannot be approximated as a Gaussian. If we use simple edge-based level set methods without further constraint, the active contours may easily collapse due to noisier image features, even when the contour initials are very close to the real boundaries.

When we check the traditional multi-scale framework closely, we will find that the boundary shape at the highest pyramid level is used **only** as a initial contour at a lower pyramid level, while the shape similarity of boundaries at different scale levels is not used. Thus, we propose to add the boundary shape similarity as an additional constraint to guide the contour evolution at lower pyramid levels. Correspondingly, the minimization framework is modified as

$$\min_{\mathcal{C}} \{F_1 + \lambda_2 F_3\} \quad (4)$$

with the boundary similarity function F_3 defined as

$$F_3 = \int_0^1 D[\mu C^*(q), \mathcal{C}(q)] |C'(q)| dq. \quad (5)$$

In equation (5), μ is a scaling factor which interpolates \mathcal{C}^* to the next finer scale level, and D is the distance of the corresponding points between contours $\mu\mathcal{C}^*$ and \mathcal{C} . The parameter $\lambda_2 > 0$ in equation (4) is used to adjust the relative importance of boundary similarity function.

The above energy minimization problems can be viewed in a more general Geometric Variational Framework and can be solved using level set methods. The basic idea of 2D level set methods is to embed a 2D curve \mathcal{C} as the zeroth level set of a hyperplanar function $\phi(x, y)$ and convert the propagation of \mathcal{C} into the temporal evolution of $\phi(x, y)$. Concretely, we use a Heaviside function

$$H(z) = \begin{cases} 1 & z \geq 0 \\ 0 & z < 0 \end{cases} \quad (6)$$

and a 1D Dirac function $\delta(z) = \frac{d}{dz}H(z)$ to reformulate equations (1), (2), and (5) in the Euclidean space as:

$$E_1(\phi) = \int_x \int_y g(I(x, y)) \delta(\phi) |\nabla \phi| dx dy \quad (7)$$

$$E_2(\phi) = \int_x \int_y \frac{1}{2\sigma_1^2} |I(x, y) - c_1|^2 H(\phi) + \frac{1}{2\sigma_2^2} |I(x, y) - c_2|^2 [1 - H(\phi)] dx dy \quad (8)$$

$$E_3(\phi) = \int_x \int_y D[\mu\delta(\phi^*), \delta(\phi)] \delta(\phi) |\nabla \phi| dx dy \quad (9)$$

By replacing F_1 , F_2 , and F_3 with E_1 , E_2 , and E_3 in equations (3) and (4) and then solving their Euler-Lagrange equations with respect to ϕ , we obtain two level set evolution equations of equations (3) and (4), respectively:

$$\begin{aligned} \frac{\partial \phi}{\partial t} = & |\nabla \phi| \delta(\phi) \{ [g(-)\kappa + \nabla g \cdot \frac{\nabla \phi}{|\nabla \phi|}] \\ & - \frac{\lambda_1}{2\sigma_1^2} [|I(x, y) - c_1|^2 + \frac{c_2^2}{c_1^2} |I(x, y) - c_2|^2] \}, \end{aligned} \quad (10)$$

$$\begin{aligned} \frac{\partial \phi}{\partial t} = & |\nabla \phi| \delta(\phi) \{ [g(-)\kappa + \nabla g \cdot \frac{\nabla \phi}{|\nabla \phi|}] \\ & + \lambda_2 [D(\mu\phi_0^*, \phi_0) \kappa + \nabla D(\mu\phi_0^*, \phi_0) \cdot \frac{\nabla \phi}{|\nabla \phi|}] \}. \end{aligned} \quad (11)$$

Here the variables of the function D are represented as $\mu\phi_0^* = \{(\mu x, \mu y) | \phi^* = 0\}$ and $\phi_0 = \{(x, y) | \phi = 0\}$. For a point A on contour $\mu\phi_0^*$, a point B on ϕ_0 is defined as its corresponding point if the vector \mathbf{AB} is normal to contour $\mu\phi_0^*$.

3 Experiment

The experimental images are obtained using HP Sonos 5500 imaging system with 3D omniprobe transducer. Each 3D+T data set has 13 to 17 frames between end-diastole

(full expansion) and end-systole (full contraction). The rotation interval of the transducer is 5 degrees per slice. In the experiments, only 2DE images with closed LV boundaries are chosen to simplify the implementation of level set methods. Image intensities, gradients, and boundary similarity measures are equally normalized to the interval $[0,1]$ in order to simplify the setting of weighting parameters. We set $\lambda_1/2\sigma_1^2 = 1$ and $\lambda_2 = 0.2$ for all experiments.

For validation, we use 24 2DE images with closed boundary contours from five different sequences in four different 3D+T datasets. They cover different slice angles and different frames between end-diastole and end-systole. As no gold standard is available, three experienced observers are asked to manually segment the LV boundary in each 2DE image. These results are used as reference to judge if the performance of our automatic segmentation algorithm is within the variation range of manual segmentation results. Here we use the mean absolute distance (MAD) (cf [18] for details) to compare the shape difference between two contours. For each 2DE image, the MADs of every two manual contours and the MADs between the automatic contour and each manual contour are estimated. Thus, we have three manual MADs and three automatic-manual MADs for each 2DE image. Then, we choose five 2DE images from one sequence with the best quality (sequence 1 in table 1) and six 2DE images from one other sequence with the worst quality (sequence 2 in table 1) as two testing groups. The mean values and the standard deviations of manual MADs and of automatic-manual MADs are separately calculated in each group, as shown in the first two rows in table 1. The same process takes place for all 24 2DE images. Table 1 shows that the standard deviation of MADs using automatic segmentation is in the same range of that using manual segmentation. One exception is the MAD comparison of the good sequence (row 1), where the standard deviation of MADs using automatic segmentation is larger than that using manual segmentation. This indicates that the importance of shape constraint is reduced when the image quality is adequately good. But in general, the results in table 1 validate that the performance of our framework is comparable to that of manual segmentation.

Table 1. The performance comparison between manual segmentation results and the results using combinative multi-scale level set framework. *Manual MADs* describes the shape distance of different manual contours, while *Automatic-Manual MADs* denotes the difference between our automatic segmentation results and manual ones.

	Manual MADs	Automatic-Manual MADs
sequence 1 (mean/deviation [pixel])	1.447 / 0.199	1.302 / 0.242
sequence 2 (mean/deviation [pixel])	2.370 / 0.794	1.958 / 0.683
all sequences (mean/deviation [pixel])	1.859 / 0.672	1.643 / 0.503

In images at the coarsest scale, the combinative level set method using both region and edge information is superior to the edge-based level set (figure 2). The edge-based active contour model cannot stop at the blurred boundary, while the combinative framework can.

In figure 3, we pass the boundary shapes from level 1 back to original images as contour initials. Traditional multi-scale analysis frameworks (such as solely edge-based

level set method) do not use the boundary shape constraint and they do not get reasonable results, even when the initial contours are very close to the true ones. The performance of the level set method is improved after using boundary shape constraint, which restricts the evolution region of the active contour and therefore guarantees the correct convergence of the level set method.

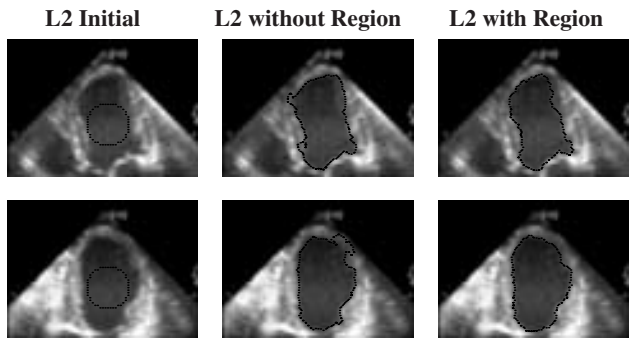


Fig. 2. Performance of the level set method at pyramid level 2. The active contours are marked in black. **Left:** The initial level set. **Middle:** Solely edge-based level set cannot provide accurate boundaries. **Right:** Segmentation results using edge and region-based combinative level set.

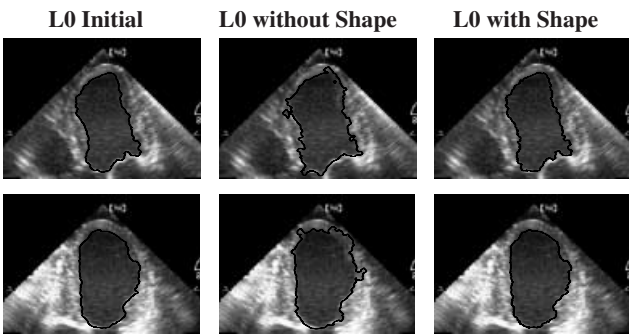


Fig. 3. **Left:** Interpolated boundary initials at level 0. They are very close to the actual boundaries. **Middle:** Solely edge-based level set cannot provide accurate boundaries. **Right:** Segmentation results using edge level set method constrained by boundary similarity.

4 Conclusion and Future Work

In shape-based approaches, the shape information has been learned via off-line training. In our approach, the boundary information is extracted automatically at the highest scale level of the pyramid. No interactive effort is needed, except the initialization step of the level set algorithm. In our algorithm, we use the boundary shape **not only** as initial contours at finer scales, **but also** as additional constraint to guide contour evolutions.

One limitation of the level set method is that it assumes a boundary is a closed curve or surface, while some long-axis echo images may not have a closed contour. The relaxation of this assumption needs to be studied in the future work. We also plan to extend our algorithm to 3D surface segmentation for real 3D echo volume data.

References

1. Xiao, G., Brady, M., Noble, J., Zhang, Y.: Segmentation of Ultrasound B-Mode images with intensity inhomogeneity correction. *IEEE Trans. Medical Imaging* **21**(1) (2002) 48–57
2. Dias, J., Leitão, J.: Wall position and thickness estimation from sequences of echocardiographic images. *IEEE Trans. Medical Imaging* **15**(1) (1996) 25–38
3. Evans, A., Nixon, M.: Biased motion-adaptive temporal filtering for speckle reduction in echocardiography. *IEEE Trans. Medical Imaging* **15**(1) (1996) 39–50
4. Zong, X., Laine, A., Geiser, E.: Speckle reduction and contrast enhancement of echocardiograms via multiscale nonlinear processing. *IEEE Trans. Medical Imaging* **17**(4) (1998) 532–540
5. Angelini, E.D., Laine, A.F., Takuma, S., Holmes, J.W., Homma, S.: LV volume quantification via spatiotemporal analysis of real-time 3-d echocardiography. *IEEE Trans. Medical Imaging* **20** (2001) 457–469
6. Bosch, J., Mitchell, S., Lelieveldt, B., Nijland, F., Kamp, O., Sonka, M., Reiber, J.: Automatic segmentation of echocardiographic sequences by active appearance motion models. *IEEE Trans. Medical Imaging* (2001)
7. Hao, X., Bruce, C., Pislaru, C., Greenleaf, J.: Segmenting high-frequency intracardiac ultrasound images of myocardium into infarcted, ischemic, and normal regions. *IEEE Trans. Medical Imaging* **20**(12) (2001) 1373–1383
8. Mulet-Parada, M., Noble, J.: 2D+T acoustic boundary detection in echocardiography. In: *Medical Image Computing and Computer-Assisted Intervention (LNCS 1496)*, Cambridge, MA, USA, October 11–13, Springer Verlag (1998) 806–813
9. Brandt, E., Wigström, L., Wranne, B.: Segmentation of echocardiographic image sequences using spatio-temporal information. In: *Medical Image Computing and Computer-Assisted Intervention (LNCS 1496)*, Cambridge, UK, September 19–22, Springer Verlag (1999) 410–419
10. Mikić, I., Krucinski, S., Thomas, J.: Segmentation and tracking in echocardiographic sequences: Active contours guided by optical flow estimates. *IEEE Trans. Medical Imaging* **17**(2) (1998) 274–284
11. Chen, Y., Thiruvankadam, S., Tagare, H., Huang, F., Wilson, D., Geiser, E.: On the incorporation of shape priors into geometric active contours. *IEEE Workshop on Variational and Level Set Methods in Computer Vision* **1** (2001)
12. Rohling, R., Gee, A., Berman, L.: A comparison of freehand three-dimensional ultrasound reconstruction techniques. *Medical Image Analysis* **3**(4) (1999) 339–359
13. Burt, P.J., Adelson, E.H.: The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications* **31** (1983) 532–540
14. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic active contours. *International Journal of Computer Vision* **22**(1) (1997) 61–79
15. Chan, T., Vese, L.: Active contours without edges. *IEEE Trans. Image Processing* **10**(2) (2001) 266–277
16. Tao, Z., Beatty, J., Jaffe, C., Tagare, H.: Gray level models for segmenting myocardium and blood in cardiac ultrasound images. Technical report, Department of Diagnostic Radiology, Yale University (2002)
17. Paragios, N., Deriche, R.: Unifying boundary and region-based information for geodesic active tracking. In: *IEEE Conf. Computer Vision and Pattern Recognition*. Volume II., Fort Collins, CO, June 23–25 (1999) 300–305
18. Chalana, V., Linker, D., Haynor, D., Kim, Y.: A multiple active contour model for cardiac boundary detection on echocardiographic sequences. *IEEE Trans. Medical Imaging* **15**(3) (1996) 290–298