

Border Detection on Short Axis Echocardiographic Views Using a Region based Ellipse-Driven framework.

Maxime Taron¹, Nikos Paragios¹, and Marie-Pierre Jolly²

¹ Ecole Nationale des Ponts et Chaussees,
Champs-sur-Marne, France

taron@cermics.enpc.fr, nikos@cermics.enpc.fr

² Imaging & Visualization Department,
Siemens Corporate Research, Princeton, NJ, USA
marie-pierre.jolly@scr.siemens.com

Abstract. In this paper, we propose a robust technique that integrates spatial and temporal information for consistent recovery of the endocardium. To account for the low image quality we introduce a local variant of the Mumford-Shah that is coupled with a model of limited parameters to describe the ventricle, namely an ellipse. The objective function is defined on the implicit space of ellipses, separates locally the blood pool from the heart wall and explores geometric constraints on the deformations of the endocardium to impose temporal consistency. Promising experimental results demonstrate the potentials of our method.

1 Introduction

Ultrasonic imaging provides time-varying two-dimensional imagery of the heart and is an important modality in medical image analysis [11] because of the low acquisition cost and the portability of the device. However, such images present low signal-to-noise (SNR) ratio resulting in poor signal quality and therefore model-based approaches are suitable to deal with such a modality.

Cardiovascular diseases are an important cause of death in the United States. Segmenting the left ventricle is a challenging component of computer aided diagnosis. Recovering the systolic and diastolic form of the endocardium is often the target of echocardiographic segmentation. Learning appropriate organ models [4] is an important pre-segmentation step for model-based techniques. The objective is to recover a compact structure that consists of limited parameters and can describe most of the examples within the training set.

Segmentation techniques for ultrasonic images consist of model-free [7, 1] and model-based approaches [3, 9]. Model-free techniques can better capture the variations of the endocardium while suffering from not being able to deal with the corrupted data. Model-based approaches are more efficient when dealing with echocardiographic images due to the high signal-to-noise ratio. One can also separate the techniques that perform filtering/segmentation in the polar [7] or in the raw space [1]. Markov random fields formulations [7], active shape and appearance models [4, 9], snakes [3], deformable models and templates [8] and level set techniques [1] are well established

techniques considered to address the segmentation of the left ventricle in echocardiographic images.

In this paper we propose a local variant of an ellipse driven Mumford-Shah framework where data terms are considered locally. Such an approach is well suited to echocardiography when data is heavily corrupted. Furthermore, the use of prior knowledge is introduced in a simple and efficient manner according to a continuous (in the temporal domain) elliptic model. Segmentation is then equivalent to the recovery of a limited set of parameters (5) that are continuous in time.

The remainder of this paper is organized as follows. In the next section, we introduce the local variant of the Mumford-Shah framework. Segmentation that involves the recovery of the ellipse parameters is presented in section 3 while temporal segmentation is considered in section 4. Results are discussed in section 5.

2 Object Extraction through Curve Propagation.

Curve evolution is a popular technique in image analysis for object extraction and segmentation. The Mumford-Shah framework [10] and the snake model [8] are the basis of a wide range of segmentation techniques. To this end, the evolution of a closed contour C is associated with the minimization of a cost function (energy) that accounts for certain image properties while respecting the internal geometric properties of the contour. In the most general case, the objective is to segment the image while restoring the signal [10]. That is equivalent to optimizing an objective function of the following form:

$$E(u, C) = E_{image}(u, C) + E_{smooth}(C) + E_{shape}(C)$$

E_{image} depends on the image features along a planar curve C that creates a partition of the domain into two regions (Ω_1 and Ω_2). This term also depends on a function u that has to be smooth on the domains Ω_1 and Ω_2 while approximating the image f . The smoothness term E_{smooth} depends on the internal curve properties. **In our case we use $E_{smooth}(C) = \nu L(C)$ according to the L^2 norm as in [6].**

Last, E_{shape} put constraints on the contour according to some prior knowledge.

We consider **the image part** of the Mumford-Shah functional to be the basis of our approach [10] :

$$E_{image}(u, C) = \frac{1}{2} \int_{\Omega} (f - u)^2 d\omega + \lambda^2 \frac{1}{2} \int_{\Omega - C} |\nabla u|^2 d\omega \quad (1)$$

The first component is a matching term between the image f and the approximation function u , while the second term enforces the regularity (smoothness) of the u function. Most of existing segmentation techniques are based on the assumption of homogeneity of visual features. Therefore, under certain conditions, u could be considered as piecewise constant. **This is equivalent to considering $\lambda \rightarrow \infty$ and** simplifies the objective function as follows :

$$E_{image}(u, C) = \frac{1}{2} \int_{\Omega} (f - u)^2 d\omega$$

The function u that minimizes the energy for a fixed contour corresponds to the average intensity within the regions Ω_1 and Ω_2 :

$$u(x) = \begin{cases} \frac{1}{|\Omega_1|} \int_{\Omega_1} f(\omega) d\omega & \text{if } \omega \in \Omega_1 \\ \frac{1}{|\Omega_2|} \int_{\Omega_2} f(\omega) d\omega & \text{if } \omega \in \Omega_2 \end{cases}$$

The assumption of homogeneity is not valid when dealing with ultrasonic images. Speckle, presence of papillary muscles, and other structures violate such assumptions. Consequently, the scenario where the two image classes are well separated using their means is unrealistic. On the other hand, such condition is valid within local image patches. The Mumford-Shah framework can now be considered in local patches. The calculation of the approximation function u only takes into account the pixel intensities in a local neighborhood.

Let ω be the radius of the local neighborhood and $\mathcal{K}_{x,y}$ be a local image patch centered at location (x, y) :

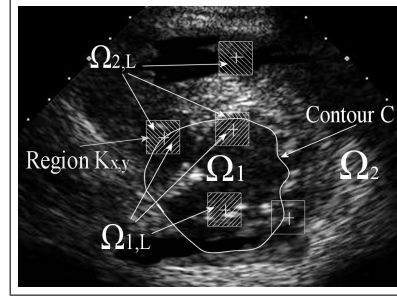
$$\mathcal{K}_{x,y} = [x - \omega, x + \omega] \times [y - \omega, y + \omega]$$

Then, using prior definitions on Ω_1 and Ω_2 , we can introduce a local partition as shown on the side figure according to:

$$\mathcal{K}_{x,y} = \Omega_{1,L}(x, y) \cup \Omega_{2,L}(x, y)$$

$$\Omega_{1,L}(x, y) = \mathcal{K}_{x,y} \cap \Omega_1$$

$$\Omega_{2,L}(x, y) = \mathcal{K}_{x,y} \cap \Omega_2$$



Using this local partition, we introduce the following functional :

$$E_{image}(u, C) = \frac{1}{2} \int \int_{\Omega} (f(x, y) - u(x, y))^2 dx dy + \frac{1}{2} \lambda \left[\int \int_{\Omega_1} \left(\int \int_{\Omega_{1,L}(X,Y)} (f(x, y) - u(X, Y))^2 dx dy \right) dX dY + \int \int_{\Omega_2} \left(\int \int_{\Omega_{2,L}(X,Y)} (f(x, y) - u(X, Y))^2 dx dy \right) dX dY \right] \quad (2)$$

Where the last two terms are optimized when u corresponds to the average image intensity over the $\Omega_{1,L}$ and $\Omega_{2,L}$. Considering the limit case when $\lambda \rightarrow \infty$, the energy can be rewritten and simplified :

$$E_{image}(u, C) = \frac{1}{2} \int \int_{\Omega} (f(x, y) - u(x, y))^2 dx dy \quad (3)$$

$$\text{with : } u(x, y) = \begin{cases} \frac{1}{|\Omega_{1,L}(x,y)|} \int \int_{\Omega_{1,L}(x,y)} f(X, Y) dX dY & \text{if } (x, y) \in \Omega_1 \\ \frac{1}{|\Omega_{2,L}(x,y)|} \int \int_{\Omega_{2,L}(x,y)} f(X, Y) dX dY & \text{if } (x, y) \in \Omega_2 \end{cases} \quad (4)$$

Such a definition makes u a piecewise smooth function on Ω_1 and Ω_2 and discontinuous across the contour C .

Minimizing the energy (3) using gradient descent results in the following evolution equation :

$$\frac{\partial C}{\partial t} = [e^+(s, t) - e^-(s, t)] \cdot n(s, t) = 0 \quad \forall s \in [0, 1].$$

Where $n(s)$ is the outer normal to the contour and $e^{-/+}$ denotes the energy density outside and inside the contour :

$$e^{-/+} = (f(C(s)) - u(C(s)))^2$$

Quite often, image derived terms are not sufficient to perform segmentation since noisy and corrupted data could lead to non optimal results. Prior knowledge on the geometric form of the structure to be recovered could address such a limitation.

Statistical analysis of a set of training examples is the most common technique to derive compact representations used to encode prior knowledge. Given a training set, we compute the statistics on the deformations of the endocardium using Principal Component Analysis (PCA) [5]. These statistics are then used to compute the likelihood of the contour C in the energy term $E_{shape}(C)$ [6]. Since the orientation and scale of the endocardium are also unknown, the energy seeks for a similarity transformation between the contour and the average model. Such an analysis could provide an adequate framework to introduce prior knowledge on the segmentation.

However, the complexity of the model and the large number of training examples are difficult issues. Instead, one can model the endocardium in short axis views as an ellipse. Such a statement is supported by the PCA where the mean model is an ellipse with a small eccentricity and the modes of variation do not significantly alter this. Therefore, it is adequate to consider short axis segmentation in the space of ellipses. Such an approach is a good compromise between complexity and fair approximation of the ventricle shape.

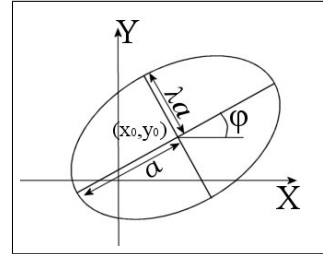
3 Ventricle Recovery through an Elliptic Structure

The segmentation according to an ellipse is equivalent to the recovery of five parameters parameters. Therefore, we consider the following space to be optimized :

$$\Theta = [a, \lambda, \phi, x_0, y_0]$$

where :

- a is the length of the main axis.
- λ is the aspect ratio between the two axes.
- ϕ is the aspect rotation between the main axis and the abscissa.
- x_0, y_0 the coordinates of the center.



Given this parameter space, the parametric and implicit form of the ellipse are defined as follows :

$$R_{\Theta}(\theta) = R_{a,\lambda,\phi,x_0,y_0}(\theta) = \begin{cases} x(\theta) = a(\cos(\theta)\cos(\phi) - \lambda\sin(\theta)\sin(\phi)) + x_0 \\ y(\theta) = a(\cos(\theta)\sin(\phi) + \lambda\sin(\theta)\cos(\phi)) + y_0 \end{cases}$$

$$F_{\Theta}(x, y) = 0 = \left(\frac{(x-x_0)\cos(\phi) + (y-y_0)\sin(\phi)}{a} \right)^2 + \left(\frac{-(x-x_0)\sin(\phi) + (y-y_0)\cos(\phi)}{\lambda a} \right)^2 - 1$$

Knowledge based segmentation is now equivalent to deforming an ellipse according to Θ , so it is attracted to the desired image features. Within such a concept, **the smoothness is ensured by the parametrisation; therefore the energy term $E_{smooth}(C)$ can**

be omitted.

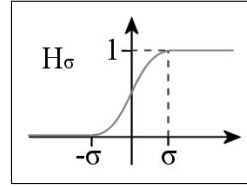
An ellipse (Θ) creates an image partition (Ω_1, Ω_2). **We use a similar energy term as in section 2, also considering the limit case in equation (2) when $\lambda \rightarrow \infty$ leading to the following simplified energy :**

$$E_1(u, \Theta) = \frac{1}{2} \int \int_{\Omega} (f(x, y) - u(x, y))^2 dx dy \quad (5)$$

where for the approximation function u an interpretation as the one presented in equation (4) is considered.

In order to explicitly introduce the ellipse parameters, we define a C^1 function H_{σ} (that converges toward a Heaviside distribution when $\sigma \rightarrow 0$):

$$H_{\sigma}(x) = \begin{cases} 0 & \text{if } x \in]-\infty, -\sigma[\\ 0.5(1 + \sin(\frac{\pi x}{2\sigma})) & \text{if } x \in]-\sigma, \sigma[\\ 1 & \text{if } x \in]\sigma, \infty[\end{cases}$$



Then the cost function could be rewritten :

$$E_1(u, \Theta) = \lim_{\sigma \rightarrow 0} \int_{\Omega} [H_{\sigma}(F_{\Theta} - \sigma) + H_{\sigma}(-F_{\Theta} - \sigma)] (f - u)^2 dx$$

and be differentiated leading to the following expression :

$$\frac{\partial E_1(\Theta, u)}{\partial \Theta} = \int_0^{2\pi} \frac{\partial(F_{\Theta}(R_{\Theta}(\theta)))}{\partial \Theta} \frac{1}{|\nabla F_{\Theta}(R_{\Theta}(\theta))|} k(\theta) h(\theta) d\theta$$

with $k(\theta) = ((f_{ext}(R_{\Theta}(\theta)) - u_{ext}(R_{\Theta}(\theta)))^2 - (f_{int}(R_{\Theta}(\theta)) - u_{int}(R_{\Theta}(\theta)))^2)$

and $h(\theta) = a\sqrt{(\sin(\theta))^2 + (\lambda \cos(\theta))^2}$

The calculation of the gradient of the implicit representation ($|\nabla F_{1,\Theta}(R(\theta))|$) and the derivative of $\nabla F_{1,\Theta}$ with respect to the parameters of the ellipse is straightforward.

Up to this point, regional information was considered to perform segmentation. While in general boundary information is not reliable, there are areas where the boundaries on the endocardium are visible. Towards exploring such a visual cue, a term based on gradient information can be considered. This information was considered in various papers dealing with boundary detection on short axis views after a preprocessing of the ultrasound image [12].

Let g be a monotonically decreasing function that depends on the image gradient defined at the pixel level. A common form of boundary term refers to the sum over the whole contour of the edge function g :

$$E_2(\Theta) = \int_{\theta=0}^{\theta=2\pi} g(R_{\Theta}(\theta)) d\theta \quad (6)$$

The speckle that transforms the myocardium wall into a collection of small bright dots generates strong values of the gradient magnitude almost everywhere on the heart muscle. Therefore, using the gradient magnitude to compute g could lead to inconsistent results. So, rather than seeking important gradient magnitude along the contour, we

also expect the gradient to be directed toward the outside of the ventricle as proposed in [2]. Such a condition will eliminate most of the gradients due to speckle and the ones oriented in other directions will be eliminated. We use the center of the initial contour to determine an outward unit vector \vec{n} at each point of the image. Then, we compute g as follows :

$$g(z) = \frac{1}{1 + (\nabla(G * I(z)) \cdot \vec{n})^2}$$

where G is a large Gaussian kernel that aims at removing partly the speckle noise.

The global energy for the contour is therefore composed of region based and gradient based terms :

$$E(\Theta, u) = \alpha_1 E_1(\Theta, u) + \alpha_2 E_2(\Theta)$$

A gradient descent method, starting from a manually drawn initial contour, can be used to recover the optimal ellipse parameters associated to the lowest potential of this function.

4 Temporal Segmentation.

Temporal information can be valuable in echocardiography. Quite often, parts of the endocardium border are visible in only some of the frames of the cardiac cycle. Therefore, knowledge accumulation within local time intervals can be very beneficial when dealing with corrupted data. To this end, we propose an energy to be minimized on a set of frames so that it leads to the segmentation of the endocardium on a complete sequence.

The heart sits in a sac (the pericardium) filled with serous liquid; therefore the movement relatively to the other organs is restricted. Moreover, the effort made by the myocardium occurs simultaneously in every directions, so the global movement of the heart relative to the transducer (fixed during the recording process) will remain negligible compared with the inner movements of the heart muscles. Consequently, the coordinate of the center as well as the orientation of the ellipse can be considered constant over short sequences of frames. Moreover, the eccentricity of the ellipses approximating the endocardium varies very little so that the scale factor will be considered as the only temporal variable.

We consider the vectors of parameters Θ_t , defining an elliptic contour on the frame at time t . The motion constraints mean that the parameters $[\lambda, \phi, x_0, y_0]$ are constant over a set of frames and the only parameter depending on time is the length of the main axis a_t :

$$\Theta_t = [a_t, \lambda, \phi, x_0, y_0] = [a_t, \Theta^*]$$

Then one can define the following energy computed on a sequence of frames :

$$E(\Theta) = \sum_{\tau=t-\Delta t}^{\tau=t+\Delta t} \{\alpha_1 E_{1,\tau}(a_\tau, \Theta^*) + \alpha_2 E_{2,\tau}(a_\tau, \Theta^*)\} \quad (7)$$

That consists of $5 + 2\Delta t$ parameters. The minimization of such a functional using a gradient descent leads to the segmentation of the endocardium for $2\Delta t + 1$ frames.

$E_{1,\tau}$ is an energy term similar to the one presented in (5) for the frame at time τ . Therefore, it depends on the ellipse with parameters Θ_τ and the approximation function u . Assuming that the contour is tracking some features, one can also consider that these features have about the same intensity on consecutive frames. Therefore, the approximation function u combined information from neighboring frames:

$$\frac{\partial E_1(\Theta, u)}{\partial \Theta_t^*} = \sum_{\tau=t-\Delta t}^{\tau=t+\Delta t} \int_0^{2\pi} \frac{\partial(F_{\Theta_\tau}(R_{\Theta_\tau}(\theta)))}{\partial \Theta_t^*} \frac{1}{|\nabla F_{\Theta_\tau}(R_{\Theta_\tau}(\theta))|} g_\tau(\theta) h_\tau(\theta) d\theta$$

with $h_\tau(\theta) = a_\tau \sqrt{(\sin(\theta))^2 + (\lambda \cos(\theta))^2}$
and $g_\tau(\theta) = [f_{ext,\tau}(R_{\Theta_\tau}(\theta)) - v_{ext,\tau}(\theta)]^2 - [f_{int,\tau}(R_{\Theta_\tau}(\theta)) - v_{int,\tau}(\theta)]^2$
 $v_{side,\tau}(\theta) = 0.2u_{side,\tau-1}(R_{\Theta_{\tau-1}}(\theta)) + 0.6u_{side,\tau}(R_{\Theta_\tau}(\theta)) + 0.2u_{side,\tau+1}(R_{\Theta_{\tau+1}}(\theta))$
 $v_{side,\tau}$ is calculated on frame τ with *side* either equal to *ext* or *int*.

In practice, an initial ellipse is positioned in the vicinity of the endocardium in one of the frames (t) of the cardiac cycle. Then, segmentation is performed within a temporal window of five frames. Upon convergence of the process, the temporal window is shifted and centered at the next frame. The segmentation parameters on the previous step are used to initialize the process, and a new solution is recovered for frame $t + 1$. The process is repeated until the entire clip is segmented.

5 Discussion and Perspectives

The proposed technique was considered to segment several short axis echocardiographic clips. Preliminary results - as shown in figure 1 - demonstrate the potentials of the method. From a set of 15 patients, the method recovered successfully the endocardium in 5 cases. While a more in depth validation is required, we conclude that the use an ellipse is a good compromise between complexity and performance. In addition, local image-driven terms improve significantly the segmentation results. However, more complicated statistical methods are required to capture the complex structure of the visual features, a promising direction to be explored. Moreover, introducing visual motion and tracking in a coherent manner along with existing geometric consistency will further improve the results and provide the clinical user with tools for better diagnosis.

References

1. E. Angelini, A. Laine, S. Takuma, J. Holmes, and S. Homma. LV volume quantification via spatio-temporal analysis of real-time 3D echocardiography. *IEEE Transactions on Medical Imaging*, 20:457–469, 2001.
2. S. Birchfield. Elliptical head tracking using intensity gradient and color histograms. *IEEE Computer Vision, Pattern Recognition*, 1998.
3. V. Chalana, D.T. Linker, D.R. Haynor, and Y. Kim. A multiple active contour model for cardiac boundary detection on echocardiographic sequences. *MedImg*, 15(3):290–298, June 1996.
4. T. Cootes, A. Hill, C. Taylor, and J. Haslam. Use of Active Shape Models for Locating Structures in Medical Imaging. *Image Vision and Computing*, 12:355–366, 1994.

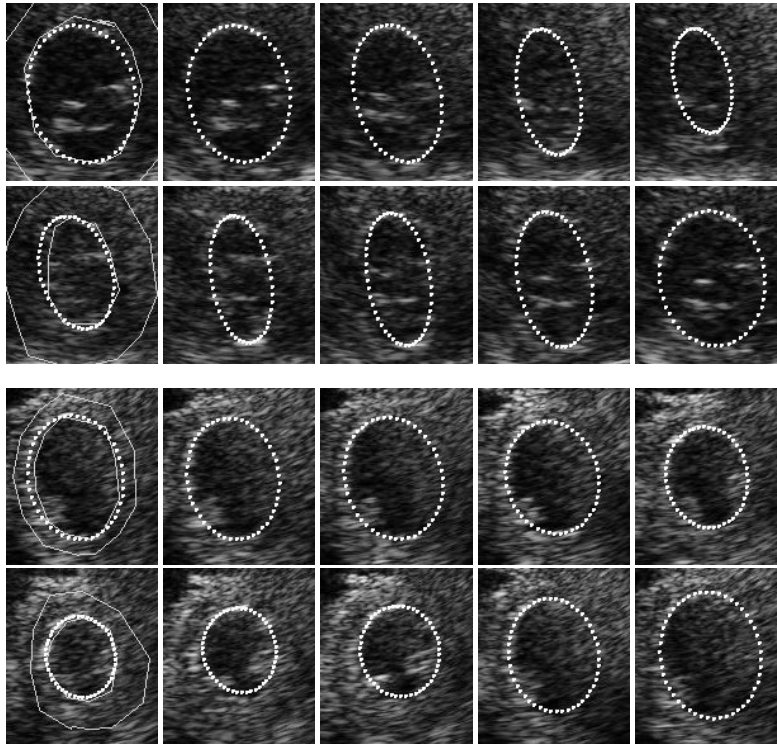


Fig. 1. Segmentation using the ellipse model on a complete cardiac cycles for two different patients. The dotted ellipse shows the recovered endocardium. The thin line shows the ground truth for both endocardium and epicardium on two frames as drawn by an expert.

5. T.F. Cootes, C.J. Taylor, D.H. Cooper, and J. Graham. Active shape models, their training and applications. *Computer vision and image understanding*, 61(1):38–59, 1995.
6. D. Cremers, F. Tischhuser, J. Weickert, and C. Schnrr. Diffusion snake : Introducing statistical shape knowledge into the mumford-shah functional. *IJCV*, 3:295–313, 2002.
7. I. Herlin and N. Ayache. Feature extraction and analysis methods for sequences of ultrasound images. In *European Conference in Computer Vision*, pages 43–55, 1992.
8. M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active Contour Models. *International Journal of Computer Vision*, 1:321–332, 1988.
9. S.C. Mitchell, J.G. Bosch, B.P.F. Lelieveldt, R.J. van der Geest, J.H.C Reiber, and M. Sonka. 3-d active appearance models: Segmentation of cardiac mr and ultrasound images. *IEEE Trans. Med. Img.*, 21:1167–1178, 2003.
10. D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math.*, 42:577–685, 1989.
11. C. Rumack, S. Wilson, and W. Charboneau. *Diagnostic Ultrasound*. Mosby, 1998.
12. Y. Kim V. Chalana, D. R Haynor. Left ventricular detection from short axis echocardiograms : the use of active contour models. *SPIE*, 2167:787–798, 1994.