A Level Set Approach for Shape-driven Segmentation and Tracking of the Left Ventricle

Nikos Paragios

***Abstract*— Knowledge-based segmentation has been explored significantly in medical imaging. Prior anatomical knowledge can be used to define constraints that can improve performance of segmentation algorithms to physically corrupted and incomplete data. In this paper our objective is to introduce such knowledge-based constraints while preserving the ability of dealing with local deformations. Towards this end, we propose a variational level set framework that can account for global shape consistency as well as for local deformations. In order to improve performance the problems of segmentation and tracking of the structure of interest are dealt with simultaneously by introducing the notion of time in the process and looking for a solution that satisfies that prior constraints while being consistent along consecutive frames. Promising experimental results in MR and ultrasonic cardiac images demonstrate the potentials of our approach.**

***Index Terms*— Curve Propagation, Optical Flow, Level Set Technique, Prior Shape Knowledge.**

I. INTRODUCTION

Computer aided diagnosis is a growing application domain of medical image analysis. Segmentation and tracking of cardiac structures are advanced techniques used to assist physicians in various states of treatment of cardiovascular diseases.

Segmentation is an ill-posed problem. Pose and reflection properties of the object, noise from the acquisition devices are some of the factors that can interfere with the process. Medical imaging is a bounded area regarding the above conditions. The clinical user can control the acquisition process while sensor perturbations can be considered known. Last, but not least, the physical entities to be recovered are constrained to follow known topology with certain degree of variation.

Model-based segmentation methods encode such prior knowledge in the form of parametric representations. Optimal solution corresponds to the lowest potential of a an image driven objective function defined along the object representation. In the absence of prior knowledge, the use of model-free techniques based on geometric flows, statistical methods, Markov random fields, graph theory and region-growing tools were investigated.

Snake-driven approaches [5] are popular in medical image segmentation. B-Splines, deformable templates, Fourier descriptors are common ways to describe the structure of interest. Level set representations [10] in an emerging technique to represent shapes and track moving interfaces for segmentation and tracking [8]. Dealing with local deformations, multi-component structures and changes of topology are the main strengths of these representations. At the same time, accounting for prior shape knowledge - a significant limitation - was recently addressed in various forms [2], [7], [13].

Nikos Paragios is with the Real Time Vision & Modeling Department at Siemens Corporate Research, Princeton, NJ - mailto:nikos@scr.siemens.com.

Tracking is a complementary to segmentation task that involves the recovery of the structure of interest in the temporal domain. While such consideration can increase complexity, at the same time the use of dynamic information can improve segmentation performance [15]. Quite often tracking is equivalent with seeking pixel-wise correspondence, thus estimating the apparent motion. Optical flow estimation is required to establish correspondence from one frame to the next. Global motion models is a compromise between low complexity and good matching for planar objects.

In this paper we propose a level set framework for shape-driven knowledge-based segmentation and tracking that is parameter free, implicit and intrinsic. Prior shape knowledge is represented using a probabilistic level set distance map and global shape consistency is inherited to the process through a rigid registration of the evolving interface to the prior model. Visual evidence is integrated through a boundary and a region-based segmentation module while internal smoothness constraints are also imposed. Temporal information is accounted by seeking for a global motion model for the structure of interest that satisfies the visual constancy constraint. Related segmentation techniques with our approach can be found in [4], [11], [15].

The reminder of this paper is organized as follows; In section 2 we introduce the application context and the level set method. Prior shape knowledge is introduced in section 3 while in section 4 we couple segmentation and tracking using parametric motion models. Discussion is part of section 5.

II. LEFT VENTRICLE SEGMENTATION & LEVEL SET METHODS

Identifying the heart chambers, the endocardium and the epicardium is a powerful diagnostic tool. In particular the detection, segmentation and tracking of the left ventricle is of great importance because it pumps oxygenated blood out to distant tissue in the entire body. Furthermore, measuring the ventricular blood volume, wall mass, wall motion and the wall thickening properties over various stages of the cardiac cycle are components with strong diagnostic power.

Magnetic Resonance Images exhibit high quality, while involving misleading visual information, like the papillary muscles [11]. On the other hand, the signal-to-noise ratio is very high for the ultrasound modality making its direct use inappropriate for data-driven automated solutions. One can consider the use of prior knowledge to address these limitations for both modalities. Clinical expertise can be used to derive a set of training examples and a representative shape model of the structure of interest.

The segmentation and tracking of the left ventricle can be viewed as a bi-modal frame partition problem. One would like to separate the endocardium from the background. We address this partition by considering a curve propagation approach. Visual (boundary and regional) terms, prior shape knowledge and internal constraints are used to derive an automated solution for detection and tracking of the left ventricle.

The level set method [10], [9] is an emerging technique for tracking moving interfaces. To this end, given a motion equation that dictates the propagation of a closed structure, one can construct a structure of a higher dimension and define a corresponding flow such that its zero level set yields always to the position of the input structure. A step further is to consider the definition of the problem and the objective function [16], [14] directly on the space of level set representations.

Towards this end, one can define the approximations of Dirac and Heaviside [16] distributions;

$$\delta_a(\phi) = \begin{cases} 0 & ,|\phi| > \alpha \\ \frac{1}{2\alpha}\left(1 + \cos\left(\frac{\pi\phi}{a}\right)\right) & ,|\phi| < \alpha \end{cases} \quad (1)$$

$$H_\alpha(\phi) = \begin{cases} 1 & ,\phi > \alpha \\ 0 & ,\phi < -\alpha \\ \frac{1}{2}\left(1 + \frac{\phi}{\alpha} + \frac{1}{\pi}\sin\left(\frac{\pi\phi}{a}\right)\right) & ,|\phi| < \alpha \end{cases} \quad (2)$$

and use them to introduce an image partition objective function. Boundary attraction as well region-consistency terms can be defined based on an evolving function Φ. The geodesic active contour [1], [6] can be used for example to perform boundary extraction.

$$E_B(\Phi) = \underbrace{\iint_\Omega \delta_\alpha(\Phi) b(|\nabla I|)|\nabla\Phi| d\Omega}_{boundary\ module} \quad (3)$$

where $b : \mathcal{R}^+ \to [0,1]$ is a monotonically decreasing function. The lowest potential of this functional corresponds to a minimal length geodesic curve attracted by the boundaries of the structure of interest. Calculus of variations will lead to a one-directional flow that aims at shrinking or expanding (mutually exclusive) the initial interface towards the object boundaries while being constrained by the curvature.

$$\frac{d}{dt}\Phi = \delta_\alpha(\phi)\text{div}\left(b(|\nabla I|)\frac{\nabla\Phi}{|\nabla\Phi|}\right) \quad (4)$$

Regional/global information can improve performance of boundary-based flows [12] that suffer of being sensitive to the initial conditions. The central idea behind this module is to use the evolving interface to define an image partition that is optimal with respect to some grouping criterion. Within level set representations such partition is natural according to the sign of the embedding function. The Heaviside function can be considered to define such partition;

$$E_R(\Phi) = \underbrace{\iint_\Omega H_\alpha(\Phi) r_O(I) d\Omega}_{endocardium} + \underbrace{\iint_\Omega (1 - H_\alpha(\Phi)) r_B(I) d\Omega}_{background} \quad (5)$$

according to some region descriptors functions $r_O : \mathcal{R}^+ \to [0,1], r_B : \mathcal{R}^+ \to [0,1]$ that are monotonically decreasing functions. Such descriptors measure the quality of matching between the observed image and the expected regional properties of the structure of interest and the background.

MR sequences of the left ventricle [Figure (1)] can be decomposed into three populations [4]: (i) the blood (bright), (ii) the muscles (gray) and (iii) the air-filled lungs (dark gray). The characteristics of these populations can be discriminated fairly well and the observed distribution (histogram) of the epicardium region is a mixture model with three components (assumed to be Gaussian) [11]. Let p_E be the endocardium density function, p_M the myocardium density function and p_B the density function of the rest of the cardiac organs (background). Then, we can write

$$p(I) = P_E\ p_E(I) + P_M\ p_M(I) + P_B\ p_B(I) \quad (6)$$

where P_E, P_M, P_B are the *a priori* probabilities for the endocardium, the myocardium and the background hypotheses. The unknown parameters of this model can be estimated using the expectation-maximization principle. The background distribution is considered to be either the distribution of the myocardium or of the other human organs. A similar analysis can be considered for ultrasonic images of the left ventricle using an exponential function for the endocardium and a Gaussian distribution for the heart walls.

The minimum of the region-driven segmentation module corresponds to the optimal grouping of the observed visual information according to some pre-defined criteria where homogeneity is a particular case. Using the calculus of variations one can recover the following motion equation;

$$\frac{d}{dt}\Phi = \delta_\alpha(\Phi)(r_O(I) - r_B(I)) \quad (7)$$

Such flow aims refers to an adaptive balloon force. It shrinks or expands the contour towards the direction that is best supported by the visual information, given the expected intensity properties of the structure of interest and the background.

Integration of the boundary and the region-driven term can be considered to perform segmentation [12], namely the geodesic active region model. In the absence of noise, occlusions and corrupted visual information, such method can deal with local deformations. On the other hand, it cannot account for prior shape knowledge, deal with noisy, corrupted and occluded data.

III. Prior Shape Knowledge

Shape-driven constraints were considered within the propagation of curves in various ways. To this end, one has first to select an appropriate shape representation when introducing such constraints. Moreover, the extraction of an optimal set of parameters able to describe these constraints is to be done given a set of training examples. We consider a pixel-wise stochastic level set representation [13] to encode prior knowledge;

$$p_{M,\{x,y\}}(\phi) = \frac{1}{\sqrt{2\pi}\sigma_M(x,y)} e^{-\frac{(\phi-\Phi_M(x,y))^2}{2\sigma_M^2(x,y)}} \quad (8)$$

defined in the pixel level that consists of two unknown variables;

- The shape image Φ_M,
- The local degrees of variability image σ_M.

Distance transforms are used as embedding function in the definition of Φ_M. Such prior model also consists of a variability image that describes the confidence of the prior model. In areas where important local deformations are plausible high variability estimates are present. Variational principles according to the maximum likelihood criterion between the model and a training set are used to determine the function Φ_M and the variability estimates σ_M [13]. This model can be used within the segmentation process to enforce global shape consistency.

Let Φ be a level set representation to which we would like to introduce a global rigid-invariant shape constraint according to the model Φ_M. We assume that Φ is part of the family of shapes that consists of all possible rigid-transformations of the model. Introducing such constraint can be done by updating locally the evolving representation to meet the model properties; optimal local match. Correspondence is determined through a rigid registration.

Thus, given the current state Φ, we assume the existence of an ideal transformation A between the evolving representation and the shape model. In order to better account for the nature of the structure of interest, we assume that the optimal registration corresponds to the maximum likelihood between the representation and the model;

$$\begin{aligned} & (x,y) \rightarrow A(x,y) \\ \max\nolimits_{x,y} & \left\{ p_{M,A(x,y)} \left(\underline{s}\Phi(x,y) \right) \right\} \forall (x,y) : H_\alpha(\Phi(x,y)) \geq 0 \end{aligned}$$

where $\underline{s}$ is the scale factor of the registration model. Level Set Representations with distance transforms as embedding function are invariant to translation and rotation but not to scale variations. Their values are scaled accordingly. Consequently scale appears as a multiplicative factor in the matching process. Solving segmentation/registration now is equivalent with finding a representation Φ and a global registration model A;

$$E_M(\Phi, A) = \iint_\Omega H_\alpha(\Phi) \left[\log(\sigma_M(A)) + \frac{(s\Phi - \Phi_M(A))^2}{2\sigma_M^2(A)} \right]$$

This functional consists of two unknown variables; (i) a level set representation that is optimal when it becomes a rigid transformation of the prior model, (ii) a transformation (registration) between the evolving current representation and the model. This term is defined in a qualitative manner; model parts with low variability are more significant than the ones that undergo important local deformations.

One can integrate this module with the previously defined visual driven terms

$$E(\Phi, A) = \beta_1 \; E_B(\Phi) + \beta_2 E_R(\Phi) + E_M(\Phi, A) \quad (9)$$

where β_1, β_2 are blending parameters, leading to a data-driven segmentation approach that privileges certain prior knowledge on the structure of interest.

IV. Temporal Information & Tracking

Dynamic acquisition of scalar as well as volumetric images is a standard procedure in medical image analysis. Periodic motion in some sense is a natural constraint in cardiac imaging. Motion trajectories are important diagnostic tools and can provide better support compared to static images.

Tracking is a well explored topic in image processing and computer vision that was implicitly in many cases linked with the segmentation problem. Towards this end, one can consider segmentation known and then perform temporal tracking. The outcome of the tracking process can be used as an initial guess to the segmentation process. These two steps can alternate to deal with both tasks simultaneously. Such approach has certain strengths and important limitations.

Low complexity is the main advantage of going with a decoupled approach. At the same time though, temporal information is not exploited properly. Given a bi-modal segmentation assumption, one would like to modify the segmentation modules to also account for temporal tracking.

Visual constancy in the temporal domain is a common tracking constraint. In medical image analysis, one can control the acquisition devices and avoid global changes of the illumination. Reflections properties of the tissue being mapped do not change and the hypothesis of constant depth for the structure of interest can be valid. Then tracking refers to an optimal transformation T for the structure of interest between the two images I_t, I_{t+1} that satisfies the visual constancy constraint;

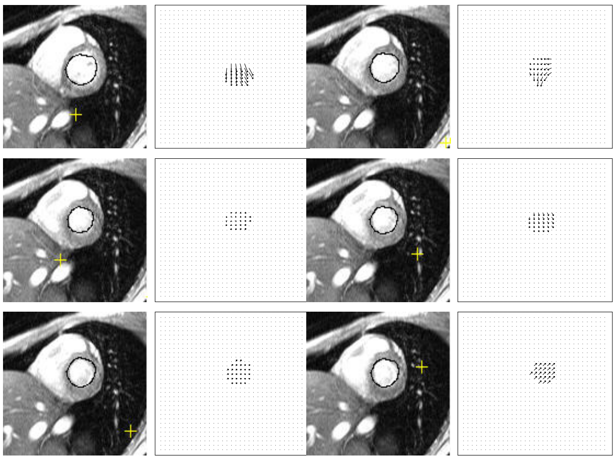
$$\begin{aligned} & (x,y) \rightarrow T(x,y) \\ I_t(x,y) \approx I_{t+1}(T(x,y)), \quad & \forall (x,y) : H_\alpha(\phi_t(x,y)) \geq 0 \end{aligned} \quad (10)$$

where this transformation can be either a parametric motion model (rigid, affine, projective) or a local deformation field $(u(x,y), v(x,y))$. A cost functional like the sum of square differences between between the image intensities can be used to recover the optimal transformation. Under the assumption of a global motion model, we can then write;

$$E(\phi_t, \phi_{t+1}, T) = \iint_\Omega H_\alpha(\phi_t) \left(I_t - I_{t+1}(T) \right)^2 d\Omega \quad (11)$$

Meaningful correspondences between the two images are considered. Using such functional one can guarantee an one-to-one correspondences from frame t to frame $t+1$. However, the same criterion is not valid for the inverse procedure. In [3] the use of direct as well as inverse transformation was proposed for volume registration. Furthermore, least square estimators like the sum of square differences are sensitive to the presence of outliers. Therefore, one can replace the least square estimator with a more robust norm ρ that refers to a bounded error function;

$$\begin{aligned} E_T(\Phi_t, \Phi_{t+1}, T) = & \iint_\Omega H_\alpha(\Phi_t) \rho \left(I_t - I_{t+1}(T) \right) d\Omega \\ & + \iint_\Omega H_\alpha(\Phi_{t+1}) \rho \left(I_t(T^{-1}) - I_{t+1} \right) d\Omega \end{aligned} \quad (12)$$



(a) (b) (c) (d)

Fig. 1. Cardiac Segmentation/Tracking for Magnetic Resonance Images presented in raster-scan format. (a) input t frame, (c) segmented t frame, (b-d) motion estimation (flow) for the structure of interest between frame t and frame $t+1$ up-scaled four times for demonstration purposes.

where T^{-1} is the inverse transformation between I_t and I_{t+1}. Calculus of variations within a gradient descent method can be used to recover the optimal estimates of the motion parameters. We consider a rigid transformation to capture the motion of the endocardium. One can integrate the tracking component to the segmentation process;

$$\begin{aligned} E(\Phi_t, \Phi_{t+1}, A_t, A_{t+1}, T)) = \beta_1\ E_B(\Phi) + \beta_2 E_R(\Phi) \\ + \beta_3 E_M(\Phi, A) + E(\Phi_t, \Phi_{t+1}, A_t, A_{t+1}, T) \end{aligned} \quad (13)$$

where segmentation and (optical flow) tracking are performed by evolving a initial curve according to the observed visual information while respecting some global shape consistency. One can ignore the propagation term that is derived from the motion component of the objective function since is a constant shrinking force. Within the proposed framework, improvements on the segmentation result lead to a better tracking and vice-versa.

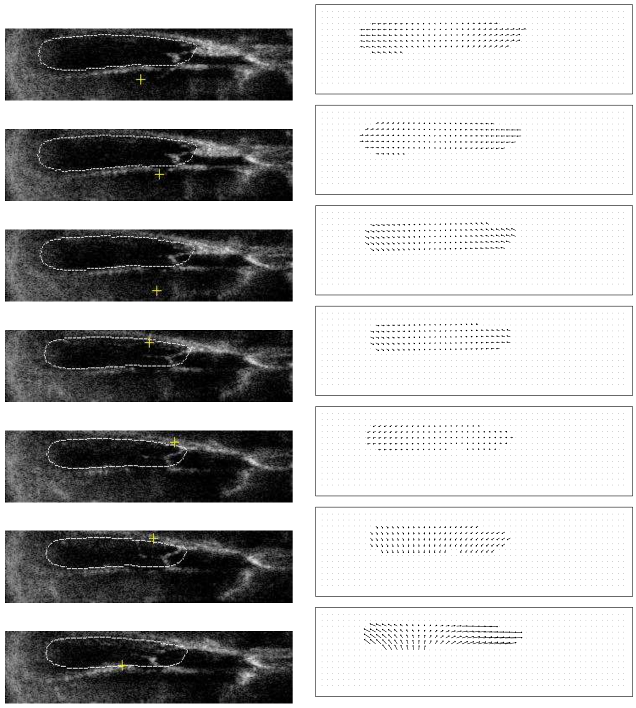
V. Discussion

In this paper we have proposed a shape-driven variational framework for knowledge-based segmentation and tracking. Our approach integrates visual information with shape constraints in the spatial and temporal domain to deal simultaneously with these tasks.

Promising experimental results using cardiac [Figure (1,2)] MRI and ultrasound were obtained. The 3D implementation of our approach is under investigation. Non-parametric shape representations within the space of distance transform is a step forward for our approach. Changes of topology is a strength of level set representations. Our approach can detect single objects with complex topology but cannot recover structures of different topology that are connected.

Acknowledgments

The author would like to thank Mikael Rousson for the implementation of the shape prior term and Marie-Piere Jolly from Siemens Corporate Research for fruitful discussions.



(a) (b)

Fig. 2. Cardiac Segmentation/Tracking for Ultrasonic Images. (a) segmented t frame, (b) motion estimation (flow) between frame t and frame $t+1$ up-scaled four times for demonstration purposes.

References

[1] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic Active Contours. *IJCV*, 22:61–79, 1997.
[2] Y. Chen, H. Thiruvenkadam, H. Tagare, F. Huang, and D. Wilson. Using Prior Shapes in Geometric Active Contours in a Variational Framework. *IJCV*, 50:315–328, 2002.
[3] G. Christensen, A. Joshi, and M. Miller. Volumetric Geometric Transformations for Mapping Anatomy. *IEEE TMI*, 16:864–877, 1997.
[4] M.-P. Jolly. Combining Edge, Region, and Shape Information to Segment the Left Ventricle in Cardiac MR Images. In *MICCAI*, pages 482–490, 2001.
[5] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active Contour Models. *IJCV*, 1:321–332, 1988.
[6] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi. Gradient flows and geometric active contour models. In *IEEE ICCV*, pages 810–815, 1995.
[7] M. Leventon, E. Grimson, and O. Faugeras. Statistical Shape Influence in Geodesic Active Controus. In *IEEE CVPR*, pages I:316–322, 2000.
[8] R. Malladi, J. Sethian, and B. Vemuri. Shape Modeling with Front Propagation: A Level Set Approach. *IEEE PAMI*, 17:158–175, 1995.
[9] S. Osher and N. Paragios. *Geometric Level Set Method in Imaging, Vision and Graphics*. Springer, 2003.
[10] S. Osher and J. Sethian. Fronts propagating with curvature-dependent speed : Algorithms based on the Hamilton-Jacobi formulation. *Journal of Computational Physics*, 79:12–49, 1988.
[11] N. Paragios. A Variational Approach for the Segmentation of the Left Ventricle in Cardiac Image Analysis. *IJCV*, 50:345–362, 2002.
[12] N. Paragios and R. Deriche. Geodesic Active Regions: A New Framework to Deal with Frame Partition Problems in Computer Vision. *Journal of Visual Communication and Image Representation*, 13:249–268, 2002.
[13] M. Rousson and N. Paragios. Shape Priors for Level Set Representations. In *ECCV*, pages II:78–93, Copenhagen, Denmark, 2002.
[14] L. Vese and T. Chan. A Multiphase Level Set Framework for Image Segmentation Using the Mumford and Shah Model. *IJCV*, 50:271–293, 2002.
[15] A. Yezzi, L. Zollei, and T. Kapur. A Variational Framework for Joint Segmentation and Registration. In *IEEE MMBIA*, pages 44–51, 2001.
[16] H-K. Zhao, T. Chan, B. Merriman, and S. Osher. A variational Level Set Approach to Multiphase Motion. *Journal of Computational Physics*, 127:179–195, 1996.