

User-Aided Boundary Delineation through the Propagation of Implicit Representations

Nikos Paragios

Real Time Vision & Modeling Department,
Siemens Corporate Research, Princeton, NJ, USA

Abstract. In this paper we introduce user-defined segmentation constraints within the level set methods. Snake-driven methods are powerful and widely explored techniques for object extraction. Level set representations is a mathematical framework technique to implement such methods. This formulation is implicit, intrinsic and parameter/topology free. Introducing shape-driven knowledge within the level set method for segmentation is a recently explored topic. User interactive constraints are more simplistic forms of prior shape knowledge. To this end, we propose a simple formulation that converts user interaction to objective function terms that aim to guide the segmentation solution through the user edits.

1 Introduction

Image segmentation approaches are either boundary or region-based. Boundary-driven techniques rely on the generation of a strength image and the extraction of prominent edges, while region-based methods rely on the homogeneity of spatially localized features and properties. Snake-driven [5] techniques is quite often the most appropriate tool to derive boundary-based methods. A curve propagation technique is a common way to implement such terms.

To this end, a parameter space that defines a curve in the image plane is considered. Then, object extraction is equivalent with finding the lowest potential of an objective function. Such a function involves internal and external terms. The internal term enforces some desired geometric characteristics of the curve, while the external term moves the curve to the desired image features. Level set methods [11] are among the most promising techniques to address such an objective in various application domains [10].

The central idea behind the level set formulation is to consider the problem in a higher dimension and represent the evolving curve as the zero-level set of an embedding function. The evolution of this function can then be derived in a straightforward manner from the original flow that guides the propagation of the curve. Such methods are implicit, intrinsic and topology free leading to a natural handling of important shape deformations.

Their main limitation is being sensitive to noise and failing to capture/encode prior knowledge shape-driven on the structure to be recovered. A geometric flow that evolves the solution closer to the prior was proposed in [7] to introduce prior shape knowledge within the segmentation process. A more elegant formulation was derived in [4,13,14] where such constraints were introduced in the form of energy components that constrain the solution space.

User-interaction [8,1] is a simplistic but important component in medical segmentation where boundary tracing tools are quite popular. It can be considered as a different form of prior knowledge to be added in the segmentation process. Recent advances in medical imaging have made automated techniques accurate enough. However, quite often clinical users have to correct their outcome. Although, level set methods are an established segmentation technique in medical imaging they do not support user interaction.

In this paper we propose a novel term to encode user-interaction within level set methods. This term is introduced in the form of an evolving shape prior [13] and transforms user-edits to level set-based propagation constraints. In section 2, we present level set methods, while interactive constraints are considered in section 3. Integration of image-driven terms and user edits is addressed in section 4. Discussion appears in section 5.

2 Level Set Representations

The level set method [11] consists of representing and evolving an evolving curve $\partial\mathcal{R}(p)$ with the zero-level set of an embedding surface $\Phi : \Omega \rightarrow \mathcal{R}$:

$$\Phi(p; t) = \begin{cases} 0 & , p \in \partial\mathcal{R}(t) \\ + \mathcal{D}((p), \partial\mathcal{R}(t)) > 0 & , p \in \mathcal{R}(t) \\ - \mathcal{D}((p), \partial\mathcal{R}(t)) < 0 & , p \in [\Omega - \mathcal{R}(t)] \end{cases} \quad (1)$$

where Ω is the image domain (bounded) and $\mathcal{D}(p, \partial\mathcal{R}(t))$ is the minimum Euclidean distance between the pixel p and the curve $\partial\mathcal{R}(t)$. The level set formulation can be considered as an optimization framework. To this end, one can define the approximations of DIRAC and HEAVISIDE distributions [15]:

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

These functions could be used to define contour-based as well as region-based energetic terms in the level set space [15]:

$$\underbrace{(i) \iint_{\Omega} H_\alpha(\Phi(p)) r(I(p)) dx dy}_{\text{regional module}}, \quad \underbrace{(ii) \iint_{\Omega} \delta_\alpha(\Phi(p)) b(I(p)) |\nabla \Phi(p)| dx dy}_{\text{boundary module}}$$

where r and g are *region* and *boundary* positive monotonically decreasing data-driven functions. The first term [i] is a grouping component that accounts for some regional properties (modulo the definition of r) of the area defined by the evolving curve. The second term [ii] is a combination of a boundary attraction term (modulo the definition of b) and a smoothness component [2,6].

3 User-Interactive Constraints

Segmentation techniques require often local corrections in particular when the visual information does not support the user-preferred solution. User interaction is a common technique to address this issue. One can consider the case of ultrasonic images. The low signal-to-noise ratio can lead to segmentation discrepancies. Correcting these results will take much lesser time than the complete hand drawing of the cardiac contours, which is the standard procedure.

Level set methods do perform propagation at the pixel level and therefore can account for important local deformations. On the other hand, one can claim that they are sensitive to noise. User interactive editing tools can be considered either as local or global constraints, a standard editing procedure to perform correction on the recovered solution. To this end, the user is asked to introduce some constraints on critical or miss-classified parts of the segmentation map.

We consider two forms of interaction. The first, consists of a single control point - used to correct local discrepancies -, while the second of a sequence of points that are connected. In order to derive shape constraints within a level set framework, we use linear or quadratic interpolation to converted the user edits into closed structures (shapes).

3.1 Construction of the Constraint

Quadratic interpolation aims at finding the value of a function at an unknown intermediate point given three data points. That is equivalent with fitting a parabola to the three data points $((x_{i-1}, y_{i-1}), (x_i, y_i), \text{ and } (x_{i+1}, y_{i+1}))$

$$y = Ax^2 + Bx + C$$

where A, B, and C are unknowns to be recovered. The data points, (x_{i-1}, y_{i-1}) , (x_i, y_i) , and (x_{i+1}, y_{i+1}) must all lie on the curve and are used to determine A, B, and C. The simplest method to recover these parameters is using the determinants:

$$A = \frac{\Delta_x}{\Delta}, \quad B = \frac{\Delta_b}{\Delta}, \quad C = \frac{\Delta_c}{\Delta} \quad (3)$$

where

$$\Delta = \begin{vmatrix} x_{i-1}^2 & x_{i-1} & 1 \\ x_i^2 & x_i & 1 \\ x_{i+1}^2 & x_{i+1} & 1 \end{vmatrix}, \Delta_a = \begin{vmatrix} y_{i-1} & x_{i-1} & 1 \\ y_i & x_i & 1 \\ y_{i+1} & x_{i+1} & 1 \end{vmatrix}, \Delta_b = \begin{vmatrix} x_{i-1}^2 & y_{i-1} & 1 \\ x_i^2 & y_i & 1 \\ x_{i+1}^2 & y_{i+1} & 1 \end{vmatrix}, \Delta_c = \begin{vmatrix} x_{i-1}^2 & x_{i-1} & y_{i-1} \\ x_i^2 & x_i & y_i \\ x_{i+1}^2 & x_{i+1} & y_{i+1} \end{vmatrix}$$

We consider two forms of acceptable constraints; (i) independent control points (ii) multiple connected control points provided in a clock-wise order.

3.2 User Interaction Using a Single Control Point

Quite often segmentation maps fail to capture important details due to the absence of strong visual support. One can address this limitation by enforcing the solution to go

through such a problematic segment. Within our approach, a core point (\hat{p}) provided by the user and smoothness constraints on the solution are used to correct such local discrepancies.

One can consider improving the solution locally by replacing a small segment of the actual solution with the interactive part. Therefore, given a point (\hat{p}) the curve points that lie within a distance d^1 are considered. The use of distance maps as embedding function for the level set representations, provide a straightforward manner to determine these points²

$$\mathcal{N}_{\hat{p}} = \{p_i \in \Omega : |\Phi(p) - d| < \delta\} \quad (4)$$

where $\delta \leftarrow 0$. For convex shapes and control points that lie on the object, it can be proved that $\mathcal{N}_{\hat{p}}$ consists of two points (for a reasonable small selection for d). Such assumption does not hold for any shape. Therefore, more than two points can satisfy the constraint. In order to introduce the interactive segment, only two points of the curve will be considered. We select the ones with maximum angular separation at (\hat{p}):

$$(p_l, p_r) : \underset{(p_i, p_j) \in \mathcal{N}_{\hat{p}} \times \mathcal{N}_{\hat{p}}}{\operatorname{argmax}_{\{i,j\}}} |p_i - p_j| \quad (5)$$

The next step is to perform a quadratic interpolation between (p_l, p_r, \hat{p}) and determine the interactive segment. Within the level set representations, the current position of the curve is recovered from the zero-level set of the embedding function. The curve points are four-connected (zero-crossings on the image plane), and using a simple connected component rule, we can recover them in a clock-wise order;

$$\partial\mathcal{R} = (p_0, \dots, p_r, \dots, p_l, \dots, p_N) \quad (6)$$

where p_0 is an arbitrary selected point. Towards introducing the interactive segment, one can replace the segment between p_l and p_r with the one determined by the quadratic interpolation between the control point \hat{p} and p_l and p_r ;

$$\partial\mathcal{UI} = (p_0, \dots, p_r, q_1, \dots, q_M, p_l, \dots, p_N) \quad (7)$$

One can embed such a shape in a level set function using the Euclidean distance as embedding function;

$$\Phi_C(p) = \begin{cases} 0 & , p \in \partial\mathcal{UI} \\ +\mathcal{D}(p, \partial\mathcal{UI}) > 0 & , p \in \mathcal{R}_{\mathcal{UI}} \\ -\mathcal{D}(p, \partial\mathcal{UI}) < 0 & , p \in [\Omega - \mathcal{R}_{\mathcal{UI}}] \end{cases} \quad (8)$$

Such a representation encodes the user edits in a global fashion using the existing solution in areas where user interaction is absent where one should tolerate important deviations

¹ This distance should be greater than the minimum distance between the curve and the control point; $d = D(\hat{p}, \partial\mathcal{R}(t)) + \epsilon$.

² The assumption that the control point lies on the interior of the data-driven solution has been considered within this condition to recover these points. One can easily modify the condition as follows $\mathcal{N}_{\hat{p}} = \{p_i \in \Omega : |\Phi(p) + d| < \delta\}$ to deal with control points that lie on the background.

from the constraint. We consider the distance between the control point and the image plane as an indicator for the importance of the constraint.

$$\sigma_C(p) = 1 + |p - \hat{p}|, \quad p \in \Omega \quad (9)$$

Such a measure will be small for the area around the interactive segment while being significant for the segments that are far from the user edits. The same principle can be used to account for multiple, independent user edits.

3.3 User Interaction Using Multiple Control Points

Very often, data-driven solution cannot recover a meaningful segmentation map and in order to correct errors an enormous amount of local interaction is required. Global constraints is a different form of user-interaction, that could guide the segmentation process from the very beginning. A simplistic but rather realistic scenario is the following: the user provides a minimum number of control points in a clock- wise order $(\hat{p}_1, \dots, \hat{p}_N)$ that when connected define a closed curve.

The objective is to recover a global constraint that forces the solution that be recovered to go through the control points. To this end, we wish to approximate the original function in a piecewise fashion. For any y over the entire domain of x , we must simply select the proper segment to perform the interpolation. Since the shape functions are only defined on each element we can approximate y by

$$y(x) = \alpha_{i-1}(x)y_{i-1} + \beta_i(x)y_i + \gamma_i(x)y_{i+1}$$

where $\hat{p}_i = (x_i, y_i)$ and

$$\begin{aligned} \alpha_{i-1}(x) &= \frac{(x_{i+1} - x_i)(x_{i+1} - x)(x - x_i)}{\Delta} \\ \beta_i(x) &= \frac{-(x_{i+1} - x_{i-1})(x_i - x)(x - x_{i-1})}{\Delta} \\ \gamma_{i+1}(x) &= \frac{(x_i - x_{i-1})(x_i - x)(x - x_{i-1})}{\Delta} \end{aligned}$$

This procedure can convert user interaction into a closed structure (shape) $\mathcal{R}_{\mathcal{UI}}$. The level set representation Φ_C of this structure can be considered to enforce the user input. The importance of the constraint varies across the image domain and at a given image location is inversely proportional to the minimum distance from the set of control points:

$$\sigma_C(p) = 1 + \operatorname{argmin}_i |p - \hat{p}_i|, \quad p \in \Omega, \quad i \in [1, N] \quad (10)$$

3.4 Introduction of the Constraint

The user-edits are taken into account when the evolving level set representation becomes similar to the one derived from the constraint. To this end, we consider the distance between the constraint and the evolving representation Φ .

$$E(\Phi) = \iint_{\Omega} H_{\alpha}(\Phi(x, y)) (\Phi(x, y) - \Phi_C(x, y))^2 d\Omega \quad (11)$$

that is equivalent with seeking a curve that goes through the user-defined seed points. This is done by minimizing the distance between the evolving curve and the interactive constraint.

During the model construction, we have considered that the importance of the user-interactive is determined according to the distance from the control points. It is natural to harder enforce the constraint close to the user seeds while consider the data to guide the segmentation process when there is not input from the user. The distance between the constructed prior and the control points of the constraint can be used to implement such strategy;

$$E_{interaction}(\Phi) = \iint_{\Omega} H_{\alpha}(\Phi(x, y)) \frac{(\Phi(x, y) - \Phi_C(x, y))^2}{\sigma_C^2(x, y)} d\Omega \quad (12)$$

The user interaction is optimally considered when finding the Φ that corresponds to the lowest potential of the objective function. The calculus of variations within a gradient descent method can be used to determine the optimal flow that forces the evolving curve to respect the user-defined constraints;

$$\frac{d}{dt}\Phi = - \underbrace{2H_{\alpha}(\Phi) \frac{\Phi - \Phi_C}{\sigma_C^2}}_{user-interactive \ force} - \underbrace{\delta_{\alpha}(\Phi) \frac{(\Phi - \Phi_C)^2}{\sigma_C^2}}_{deflation \ force} \quad (13)$$

This flow consists of two terms. The first evolves the curve locally Towardss the preferred topology as defined by the user. On the other hand, the second is a constant deflation force that tends to shrink the curve and consequently minimize the objective function. Therefore, we can ignore the second term - the deflation component - and use only the first term to account for the user interaction. The proposed flow enforces the preferred topology in a qualitative fashion. The variability (σ_C) of the interaction constraint is used to down-scale the effect of the term in image locations where the user input is not strong. Propagation/segmentation in these areas will be data-driven. To this end, the next step is to define an image-based term for segmentation. In the recent years, several variational frameworks have been proposed for image segmentation [12].

4 User-Interactive Geodesic Active Regions

The geodesic active contour [2,6] can be used for example to perform boundary extraction.

$$E_{boundary}(\Phi) = \iint_{\Omega} \delta_{\alpha}(\Phi) b(|\nabla I|) |\nabla \Phi| d\Omega \quad (14)$$

where $b : \mathcal{R}^+ \rightarrow [0, 1]$ 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. 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 curve to define an image partition

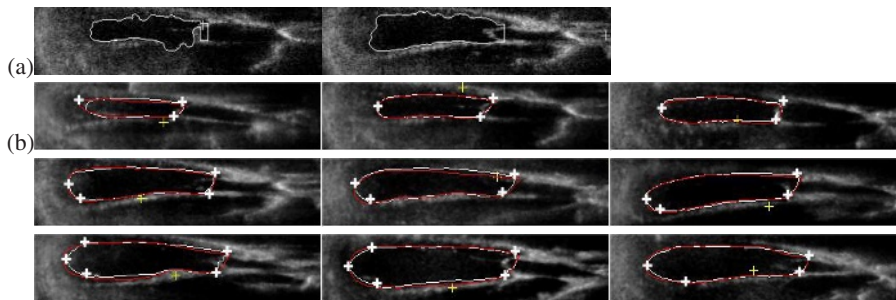


Fig. 1. User-Interactive Segmentation (white) versus Manual Segmentation (red; more that twenty edits) for the Left Ventricle in Ultrasonic polar images for different subjects. (a) Constraints on the Valve, (b) Constraints on the Valve and the Apex.

that is optimal with respect to some grouping criterion. The Mumford-Shah framework [9] has been used frequently within level set formulations [3] as global region-based grouping term using piece-wise constant functions;

$$E_{region}(\Phi) = \iint_{\Omega} H_{\alpha}(\Phi)(I - \mu_O)^2 + (1 - H_{\alpha}(\Phi))(I - \mu_B)^2 d\Omega \quad (15)$$

where μ_B, μ_O is the mean intensity for the background and the object region. The distance from the mean value is considered as a region descriptor. The mean values are dynamically updated according to the evolving segmentation map.

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 be efficient and deal with local deformations. One can also integrate the visual terms with the user-interactive constraint when available as follows;

$$E(\Phi) = w_1 E_{boundary}(\Phi) + w_2 E_{region}(\Phi) + w_3 E_{interaction}(\Phi)$$

The calculus of variations as shown earlier for each component separately, will provide a curve propagation flow that integrates visual support and user interaction. Modification of the user preferences can update the constraint on-the-fly. To this end, first a solution derived from the visual terms is recovered and then the user introduce seeds (points) for corrections. Such interaction is then converted to propagation force and refines the segmentation map Towards the user preferred solution with minimal edits. It is important to note that user interaction is introduced in the form of soft-to-hard constraint. The final solution is an equilibrium between the user edits and the solution provided by the data.

5 Conclusions

In summary, we have proposed a framework for user-interaction within the propagation of curves using level set representations. Segmentation techniques based on the propagation of curves are very popular in image processing and computer vision. Level set methods

is an emerging formulation to implement these techniques with certain strengths as well as some limitations.

Important local deformations as well as topological changes can be captured by these techniques. At the same time, they refer to an implicit geometry where local properties of the evolving curve can be easily determined. Although, some of the limitations of these methods - like their inability to account for prior knowledge - have been dealt with, to our knowledge user interaction has not been addressed. Our approach converts interactive editing into propagation constraints that force the solution to respect the user edits. The construction of the such constraints is simple and does not require additional computational resources.

Encouraging experimental results [Figure (1)] demonstrate the potentials of our method for interactive segmentation³. To this end, we have considered a medical example, the segmentation of the left ventricle in polar [Figure (1)] for ultrasonic images. This modality suffers from high signal-to-noise ratio and visual support is not sufficient to provide accurate segmentation results. Global interactive constraints have been used to improve segmentation performance of the polar domain (5 points; 2 points on the valve and 3 points on the apex).

References

1. W. Barret and E. Mortensen. Interactive Live-Wire Boundary Extraction. *Medical Image Analysis*, 1:331–341, 1997.
2. V. Caselles, R. Kimmel, and G. Sapiro. Geodesic Active Contours. In *IEEE ICCV*, pages 694–699, Boston, USA, 1995.
3. T. Chan and L. Vese. A Level Set Algorithm for Minimizing the Mumford-Shah Functional in Image Processing. In *IEEE Workshop on Variational and Level Set Methods*, pages 161–168, 2001.
4. Y. Chen, H. Thiruvankadam, H. Tagare, F. Huang, and D. Wilson. On the Incorporation of Shape Priors into Geometric Active Contours. In *IEEE VLSP*, pages 145–152, 2001.
5. M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active Contour Models. In *IEEE ICCV*, pages 261–268, 1987.
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, Boston, USA, 1995.
7. M. Leventon, E. Grimson, and O. Faugeras. Statistical Shape Influence in Geodesic Active Contours. In *IEEE CVPR*, pages I:316–322, 2000.
8. J. Liang, T. McInerney, and D. Terzopoulos. United Snakes. In *IEEE ICCV*, pages 993–940, 1999.
9. D. Mumford and J. Shah. Boundary detection by minimizing functionals. In *IEEE CVPR*, pages 22–26, San Francisco, USA, 1985.
10. S. Osher and N. Paragios. *Geometric Level Set Methods in Imaging, Vision and Graphics*. Springer Verlag, 2003.
11. 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.
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.

³ The user edits correspond to the crosses that appear in the image.

13. M. Rousson and N. Paragios. Shape Priors for Level Set Representations. In *ECCV*, pages II:78–93, Copenhagen, Denmark, 2002.
14. A. Tsai, A. Yezzi, W. Wells, C. Tempny, D. Tucker, A. Fan, A. Grimson, and A. Willsky. Model-based Curve Evolution Technique for Image Segmentation. In *IEEE CVPR*, volume I, pages 463–468, 2001.
15. 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.