HARMONIC ACTIVE CONTOURS FOR MULTICHANNEL IMAGE SEGMENTATION

Virginia Estellers[1*], *Dominique Zosso*[1†], *Xavier Bresson*[2] *and Jean-Philippe Thiran*[1]

[1] Signal Processing Laboratory, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
[2] Department of Computer Science, City University of Hong Kong, Hong Kong SAR

ABSTRACT

We propose a segmentation method based on the geometric representation of images as surfaces embedded in a higher dimensional space, handling naturally multichannel images. The segmentation is based on an active contour embedded in the image manifold, along with a set of image features. Hence, both data-fidelity and regularity terms of the active contour are jointly optimized minimizing a single Polaykov energy representing the hyper-surface of this manifold. Compared to previous methods, our approach is purely geometrical and does not require additional weighting of the energy functional to drive the segmentation to the image contours. The potential of such a geometric approach lies in the general definition of Riemannian manifolds, validating the proposed technique for scale-space methods, volumetric data or catadioptric images. We present here the segmentation technique called Harmonic Active Contours, give an implementation for multichannel images including gradient and region-based segmentation criteria and apply it to color images.

***Index Terms*—** Image segmentation, Riemannian manifolds, Active Contours, Harmonic maps

1. INTRODUCTION

Image segmentation is a fundamental step in many areas of computer vision, including object recognition, image compression and stereo vision, as it simplifies the understanding of the image from thousands of pixels to a few regions. In general terms, the goal of image segmentation is to cluster pixels into salient image regions corresponding to individual surfaces or objects with a particular characteristic. The segmentation can thus be based on different measurements taken from the image including grey level or colour values, texture features and depth or motion relative to the camera.

In natural images, however, the value of the segmenting features vary significantly over a single object, the region's contours are hard to distinguish at pixel level and the segmentation becomes an ill-posed problem. To overcome such issues, regularity constraints are usually imposed on the segmentation in terms of the region's size or the length of its contour. Image segmentation methods are then formulated as an optimization problem whose cost function includes a data and a regularity term. The optimization procedure considers the segmentation and image features as functions in a 2-dimensional space and obtains the optimal segmentation as the zero level set of the minimizing function. Active contours, for instance, define a data term encoding the image edges [1] or region-similarity information [2], while the regularity terms penalizes the length of the region's contours. In that context, we propose a new cost function based on the geometric interpretation of images as manifolds embedded in high-dimensional spaces, providing a more general framework for image segmentation and defining the optimal segmentation function as the harmonic map minimizing the hyper-surface of the manifold. The proposed technique is therefore called Harmonic Active Contour.

Harmonic Active Contours are formulated in the framework introduced by Sochen, Kimmel and Malladi in [3], which is based on geometrical ideas borrowed from high-energy physics and can be summarized in two key points. First, images are represented as Riemannian manifolds embedded in a higher dimensional spatial-feature manifold, where an adequate metric is defined. Second, the Polyakov action, a functional norm weighting embeddings in a geometric way (i.e. the representation is invariant by reparametrization), is chosen as optimization criteria and variational methods are used to minimize this action with respect to the manifold metric or embedding coordinates. The potential of this geometric framework lies in the general definition of the space-feature manifold and the choice of its metric. The features are not restricted to scalar values but include vector features encountered in color, texture or multispectral image analysis [4]. Similarly, the embedding is not limited to 2-dimensional image surfaces and generalizes naturally to n-dimensional manifolds associated to volumetric or time varying images or videos. Moreover, the choice of the metric enables the study of complex geometries inherent to scale-space methods [5] and non-flat images generated by catadioptric or omnidirectional cameras [6]. Consequently, the segmentation method we propose can be equally applied to color, texture or multispectral image features on flat or curved spaces of any dimensionality. Compared to previous segmentation methods proposed in that context [5], we provide a purely geometric approach where no additional weighting of the Polyakov action by an edge detector is required to drive the segmenting function to the image contours.

The rest of the paper is organized as follows. In section 2 we review the Beltrami framework and Polyakov action and particularize it for flat images. In section 3 we introduce the proposed embedding and define the optimal level set segmentation function as the harmonic map minimizing the hyper-surface of the manifold for different segmentation criteria. In section 4, numerical implementations for the proposed technique are given and in section 5 experimental results with color images are presented. Finally, conclusions are drawn in section 6.

2. GEOMETRIC FRAMEWORK

We shortly review the geometrical framework in which images are considered as 2-dimensional Riemannian manifolds [3]. In next sections we will limit our analysis to Euclidean spaces, but we present here a general version of the framework including non-flat high-dimensional manifolds.

*This work is supported by the Swiss SNF grant number 200021-130152
†This work is supported by the National Competence Center in Biomedical Imaging (NCCBI)

An n-dimensional manifold Σ with coordinates $\boldsymbol{\sigma} = \left(\sigma^1, \ldots, \sigma^n\right)$ is embedded in an m-dimensional manifold M with coordinates $X^1, \ldots X^m$, where $m > n$. The embedding map X is an injection given by m functions of n variables $X(\boldsymbol{\sigma}) = \left(X^1(\boldsymbol{\sigma}), \ldots, X^m(\boldsymbol{\sigma})\right)$, whose Jacobian has rank m.
In order to provide the manifold with a Riemannian structure we need to define its geometry by means of a metric, i.e. a description of how to measure distances on the manifold independent of the parametrization. In Σ, the metric $[g_{\mu\nu}]$ measures the squared distance between close points p and $p + d\boldsymbol{\sigma}$ on general curvilinear and non-orthogonal coordinates as

$$ds^2 = g_{\mu\nu} d\sigma^\mu d\sigma^\nu \qquad 1 \le \mu, \nu \le n, \tag{1}$$

where Einstein's notation is assumed (summation is implied on identical sub- and super-indices). Similarly on M distances are measured with the corresponding metric $[h_{ij}]$.

Naturally we can use the metric $[h_{ij}]$ on M and the embedding map to induce the corresponding metric on Σ, i.e. impose the distance measured on M by the local coordinates to be equal to the distance measured in Σ with the embedding coordinates. This procedure, known as pull-back relation and resulting from the chain-rule, leads to the following expression for the induced metric:

$$g_{\mu\nu} = h_{ij} X^i_\mu X^j_\nu \quad \text{where} \quad X^i_\mu = \frac{\partial X^i}{\partial \sigma^\mu}. \tag{2}$$

The Polyakov action is introduced in this context as a natural generalization of the L_2-norm, providing a measure on the space of embedding maps between Riemannian manifolds. It is defined as

$$S\left(X^i, g_{\mu\nu}, h_{ij}\right) = \int \mathrm{d}^m \sigma \sqrt{g} g^{\mu\nu} X^i_\mu X^j_\nu h_{ij}(\boldsymbol{X}), \tag{3}$$

where g is the determinant of the metric tensor $[g_{\mu\nu}]$ and $[g^{\mu\nu}]$ its inverse. If we choose the induced metric in Σ, the Polyakov energy (3) shortens to

$$S\left(X^i, h_{ij}\right) = \int \mathrm{d}^m \sigma \sqrt{g} \tag{4}$$

and represents the hyper-surface of the embedded manifold. Note that the Polyakov action is parametrization invariant and depends only on the geometrical objects, not on the way we describe them.

The Euler-Lagrange equations can be applied to minimize this action with respect to the embedding coordinate X^l. For manifolds Σ and M in general, the map that minimizes the action (3) with respect to the embedding is called harmonic map and corresponds to the natural generalization of geodesic curves and minimal surfaces in higher dimensional manifolds and different embedding spaces. Harmonic maps are often used in image processing, in form of the geodesic active contour model in image segmentation, for instance. For expression (4) of the Polyakov action, the harmonic map X^l can be iteratively obtained with a gradient descent strategy leading to the following flow for X^l

$$X^l_t = -\frac{\partial S}{\partial X^l} = -\frac{1}{2\sqrt{g}}\left(\frac{\partial g}{\partial X^l} + \frac{1}{2g}\frac{\partial g}{\partial \sigma^\mu}\frac{\partial g}{\partial X^l_\mu} - \frac{\partial}{\partial \sigma^\mu}\frac{\partial g}{\partial X^l_\mu}\right). \tag{5}$$

In terms of implementation, the gradients $\frac{\partial}{\partial \sigma^\mu}$ are estimated numerically by finite differences and only explicit formulas for $\frac{\partial g}{\partial X^l}$ and $\frac{\partial g}{\partial X^l_\mu}$ must be derived explicitly for the chosen embedding and metric.

2.1. Images as manifolds

In the following, we present some results of the application of the previous framework to image processing, where images are seen as two-dimensional manifolds embedded in higher dimensional spaces with coordinates $\boldsymbol{\sigma} = (x, y)$. We will show that the proposed framework generalizes existing image processing techniques and provides a strong theoretical frame to derive new geometric approaches.

In their seminal work [3, 4], Sochen, Kimmel and Malladi embed grey and color images in the feature space $\left(x, y, I^1, \ldots, I^l\right)$, where $I^i = I^i(x, y)$ is the grey or color intensity value for pixel (x, y). They consider the metric tensor $[h_{ij}] = diag(1, 1, \beta, \ldots, \beta)$, with $\beta > 0$ a constant controlling the scale of the feature dimension independently of the spatial dimensions, and induce the corresponding metric on Σ. In the grey level case, the Polyakov action reduces to a regularization term on the intensity pixel values given by $\int_\Sigma \sqrt{1 + \beta |\nabla I|^2}$, where the metric's parameter β allows us to consider different norms. If $\beta \to \infty$, the summand 1 in the Polyakov energy becomes negligible, and the energy approaches the TV-norm commonly used in image denoising. On the other hand, if $\beta \to 0$ the minimizing flow approaches the isotropic heat diffusion. In the case of color images, the Polyakov action includes a term controlling the coupling of the different color channels and trying to align their contours.

In the case of image segmentation, this geometric framework has been used to introduce equivalent regularization constraints in the function $\phi(x, y)$ defining the level set segmentation function [5]. The embedding map is then naturally given by $(x, y, \phi(x, y))$ and a weighting of the Polyakov action is introduced in order to attract the contours of $\phi(x, y)$ to the desired gradients of the image. It is in fact a generalization of the geodesic active contours of Caselles [1], where the curve defining the segmentation is considered a 1-dimensional manifold embedded in the 2-dimensional image space and the Polyakov action, measuring here the length of the segmenting curve, is weighted by an edge detector function. A similar geometric approach and weighted Polyakov action has also been used for image registration [7].

It is to note, however, that no purely geometrical framework for segmentation has been defined yet as the proposed embeddings include either information on the image features $I(x, y)$ or the segmenting function $\phi(x, y)$ and require an additional weighting of the Polyakov action to drive the level set function to the image contours. The proposed Harmonic Active Contours address these limitations by defining a new embedding map and corresponding metric including both image features and segmenting function in the manifold geometry. The need of a edge detector is here suppressed as the minimization of the manifold surface directly aligns the contour of the level set with the main image contours.

3. PROPOSED EMBEDDING AND METRIC

In this work we define a novel image segmentation scheme using a purely geometric approach, where data and regularization terms are naturally included in the embedding and Riemannian metric. Indeed, the Polyakov energy provides both a regularity constraint on the segmenting level set function and a data term, represented as the coupling of the gradients in the surface element, which attracts the contours of the level set function to the image contours.

In our technique, the segmentation is defined by the zero level set of a function $\phi(\boldsymbol{\sigma})$ defined on the image manifold Σ. The segmentation criterion is given by the contours of features $f^1, \ldots, f^k$, also defined on Σ and which might eventually depend on ϕ i.e $f^i = f^i(\boldsymbol{\sigma}, \phi)$ for $1 \le i \le k$.

A new embedding map M is defined to include both image features and segmentation function ϕ in the image manifold. It is given by

$$X : \left(\sigma^1, \ldots, \sigma^n\right) \mapsto \left(\sigma^1, \ldots, \sigma^n, f^1, \ldots, f^k, \phi\right). \quad (6)$$

In M we consider the metric $[h_{i,j}] = diag(1, \ldots, 1, \alpha_1, \ldots, \alpha_k, \beta)$, which induces the following metric on Σ

$$g_{\mu\nu} = \delta_{\mu\nu} + \alpha_i f^i_\mu f^i_\nu + \beta \phi_\mu \phi_\nu, \quad (7)$$

where $\delta_{\mu\nu}$ is the Kronecker delta. Note that in the proposed embedding and metric, both the image features and segmenting functions are included in the induced manifold metric with terms weighted α_i and β respectively. Previous geometric schemes would only include one of those terms and are therefore forced to weight the Polyakov action by an edge detector in order to drive the segmenting function to the images contours. We avoid it by introducing an additional dimension on the embedding, which however does not affect the dimension of our minimization function, as the intrinsic dimension of the image manifold Σ does not depend on the embedding.

In the following paragraphs we limit ourselves to planar images, i.e $\boldsymbol{\sigma} = (x, y)$, but the proposed method can be easily extended to higher image dimensions and scale spaces. The determinant of the metric tensor is then computed as $g = g_{11}g_{22} - (g_{12})^2$ and the squared hyper-surface element corresponds to

$$g = 1+\alpha_i|\nabla f^i|^2+\beta|\nabla\phi|^2+\frac{1}{2}\alpha_i\alpha_j\left[\nabla f^i, \nabla f^j\right]^2+\alpha_i\beta\left[\nabla f^i, \nabla\phi\right]^2, \quad (8)$$

where $\left[\nabla f^i, \nabla f^j\right] = f^i_x f^j_y - f^i_y f^j_x$ is the magnitude of the cross product of the vectors ∇f^i and ∇f^j. In the surface element of the defined manifold, therefore, the terms $\left[\nabla f^i, \nabla f^j\right]$ and $\left[\nabla f^i, \nabla\phi\right]$ measure the coupling of the different features f^i and the segmentation function ϕ.

The optimal segmentation function ϕ corresponds then to the harmonic map minimizing the Polyakov action defined by the induced metric, considering $f^1, \ldots, f^k$ as fixed functions of ϕ. Indeed, as can be seen in equation (8), minimization of the Polyakov energy aligns the gradients of the level set function ϕ with the gradients of the embedded features, eliminating the necessity of an edge detector function, while keeping the level set variations small. The trade-off between gradient fidelity and level set regularity is here controlled by the ratio of the metric parameters β versus α_i.

In the following, we particularize these expressions defining different features f^i to obtain both contour- and region-based segmentation.

3.1. Contour and region-based segmentation

Purely contour-based segmentation is obtained by choosing the features to be local image descriptors: $f^i = I^i(\boldsymbol{\sigma})$ for $i = 1, \ldots, l$ each channel on the image. In the easiest case, we simply embed the pixel's grey-level or color intensities, but more elaborate features such as semi-local texture descriptors, Wavelet or Gabor coefficients could be used.

At the same time, in order to detect objects that are not clearly defined by closed gradients, we introduce region-based features. We adopt the criterion proposed by Chan and Vese [2] and try to approximate the image by a two-valued piece-wise constant function. The two image regions corresponds to $\phi < 0$ and $\phi > 0$ and are characterized by specific mean values μ^i_+, μ^i_- of the local features, which are given by

$$\mu^i_+ = \frac{1}{m_+}\int_\Sigma I^i(\boldsymbol{\sigma})\, H(\phi(\boldsymbol{\sigma})) \quad \text{with} \quad m_+ = \int_\Sigma H(\phi(\boldsymbol{\sigma}))$$
$$\mu^i_- = \frac{1}{m_-}\int_\Sigma I^i(\boldsymbol{\sigma})\, H(-\phi(\boldsymbol{\sigma})) \quad \text{with} \quad m_- = \int_\Sigma H(-\phi(\boldsymbol{\sigma})), \quad (9)$$

where H corresponds to the Heaviside function. We can define then region features measuring the distance between the local image features and each region of its piece-wise decomposition by

$$f^j = \begin{cases} \left(I^i(\boldsymbol{\sigma}) - \mu^i_+\right)^2 & j = i + l,\ 1 \le i \le l \\ \left(I^i(\boldsymbol{\sigma}) - \mu^i_-\right)^2 & j = i + 2l,\ 1 \le i \le l \end{cases}. \quad (10)$$

In the proposed geometric framework we can easily combine the contour and region-based segmentations by including both contour and region-based features in the embedding. We simplify the metric then to $\alpha_i = \frac{\alpha_1}{s^i}$ for the contour features and $\alpha_i = \frac{\alpha_2}{v^i}$ for the region ones, where $s^i = \max I^i - \min I^i$ and v^i variance of I^i normalize the feature range of each channel. We have now three parameters controlling the segmentation: α_1 controls the contour criterion associated to the edge detector of the image features, α_2 controls the region criterion corresponding to a piece-wise constant decomposition of the image and β regularizes the segmentation function ϕ.

3.2. Level Set Constraint

Adopting a level set method to represent the region contours as the zero level set of the function ϕ, we must constrain this function to be a level set function during its evolution. A common practice in this case is to perform reinitialization with existing fast-marching or level set methods [8], that is, periodically stopping the evolution of ϕ and reshaping the degraded function as a signed distance function. Signed distance functions are characterized by the property $|\nabla\phi| = 1$, ensuring that the level set function is smooth and not too steep or too flat. In practice, signed distance functions lead to a stable evolution of the level set function and accurate numerical computations.

In general, however, reinitialization is not a desirable approach as it might move the zero level set away from the expected position and slows down the optimization process. To avoid it, we adopt the variational formulation proposed in [9] and introduce a distance regularization term $R(\phi)$ driving the evolution of ϕ to a signed distance function, reducing the need of reinitialization.

The evolution of the segmenting function ϕ is now derived from the minimization of the energy functional

$$E(\phi) = S(\phi) + \mu R(\phi) \quad \text{with} \quad R(\phi) = \int_\Sigma \frac{1}{2}(1 - |\nabla\phi|)^2, \quad (11)$$

where $S(\phi)$ is the previous Polyakov action and μ a positive constant. Introducing the level set term $R(\phi)$ encourages the property $|\nabla\phi| = 1$ characteristic to signed distance functions and avoids the use of reinitialization techniques.

4. NUMERICAL IMPLEMENTATION

The evolution of the segmenting function ϕ is derived as a gradient flow that minimizes the combined energy functional (11), including both the Polyakov action and the level set penalty. The combined flow presents thus two terms $\phi_t = -\frac{\partial S}{\partial\phi} - \mu\frac{\partial R}{\partial\phi}$.

For the level set constraint, the flow corresponds to the following divergence operator

$$\frac{\partial R}{\partial \phi} = -\text{div}\left(\left(1 - \frac{1}{|\nabla \phi|}\right)\nabla \phi\right). \tag{12}$$

While for the flow associated to the Polyakov action $\frac{\partial S}{\partial \phi}$, we use expression (5), where $X^l = \phi$ and $\frac{\partial g}{\partial X^l}$ and $\frac{\partial g}{\partial X^l_\mu}$ correspond to

$$\begin{aligned}\frac{\partial g}{\partial \phi} =& 2\alpha_i\left(f^i_x + \alpha_j\left[\nabla f^i, \nabla f^j\right] f^j_y + \beta\left[\nabla f^i, \nabla \phi\right]\phi_y\right)\frac{\partial f^i_x}{\partial \phi}\\ &+ 2\alpha_i\left(f^i_y - \alpha_j\left[\nabla f^i, \nabla f^j\right] f^j_x - \beta\left[\nabla f^i, \nabla \phi\right]\phi_x\right)\frac{\partial f^i_y}{\partial \phi}\end{aligned} \tag{13}$$

and

$$\frac{\partial g}{\partial \phi_\mu} = 2\beta\left(\left(1 + \alpha_i\left(f^i_\nu\right)^2\right)\phi_\mu - \alpha_i f^i_\mu f^i_\nu \phi_\nu\right). \tag{14}$$

In particular, for the segmenting features given in expression (10), $\frac{\partial f^i_\mu}{\partial \phi}$ takes the form (15), where δ is the Dirac distribution.

$$\frac{\partial f^j_\mu}{\partial \phi} = \begin{cases} 0 & 1 \le j \le l \\ -\frac{2I^i_\mu}{m_+}\int_\Sigma \left(I^i(\boldsymbol{\sigma}) - \mu^i_+\right)\delta\left(\phi\left(\boldsymbol{\sigma}\right)\right) & j = i + l, 1 \le i \le l \\ \frac{2I^i_\mu}{m_-}\int_\Sigma \left(I^i(\boldsymbol{\sigma}) - \mu^i_-\right)\delta\left(\phi\left(\boldsymbol{\sigma}\right)\right) & j = i + 2l, 1 \le i \le l \end{cases} \tag{15}$$

As it is common in level set methods, in the numerical implementation we have substituted the Heaviside and Dirac distributions by smooth approximations H_ϵ and δ_ϵ verifying $H'_\epsilon = \delta_\epsilon$ [2].

5. EXPERIMENTAL RESULTS

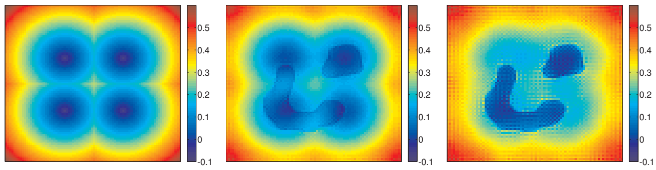
We present results with the proposed method for grey and color images. The value of the region and gradient parameters were roughly adapted to the image properties, while the level set and regularity constraints were fixed to $\beta = 10^{-4}$ and $\mu = 10^{-2}$. The segmenting function ϕ was always initialized as a signed distance function with no prior knowledge about the image, see figure 1.

The method was first applied to a simple grey-level image of two cells with equal weight for the region and gradient features $\alpha_1 = \alpha_2 = 1$. Figure 1 shows the evolution of ϕ for 500 iterations of the gradient descent step, where we observe that ϕ evolves from a signed distance function to a level set function with gradients defined by the original image. We also test our method with natural color images, where two color regions can be distinguished, see figure 2. For the airplane image more weight was given to gradient features with $\alpha_1 = 2$ and $\alpha_2 = 1$, whereas the region term dominated the segmentation of the beans' image with $\alpha_1 = 1$ and $\alpha_2 = 5$. The obtained segmentations were equally successful with grey and color images, experimentally validating the proposed segmentation technique for single and multichannel images.

6. CONCLUSIONS AND FUTURE WORK

We presented a general method for image segmentation which can naturally handle multichannel images and non-Euclidean spaces. The proposed Harmonic Active Contours are developed by a purely geometric criterion minimizing the hyper-surface of a Riemannian manifold containing data and regularity terms on its metric, and avoiding the necessity of an edge detector defined in high dimensional or non Euclidean spaces.

In this paper numerical implementations are proposed for multichannel image segmentation and successful experimental results are presented for flat color images. Future work, therefore, includes the extension of the proposed technique to spherical images and faster implementations of the numerical schemes for color and texture segmentation.



Initial ϕ. ϕ after 250 iterations. Final ϕ.

Fig. 1. Evolution of ϕ for 'cells' images.

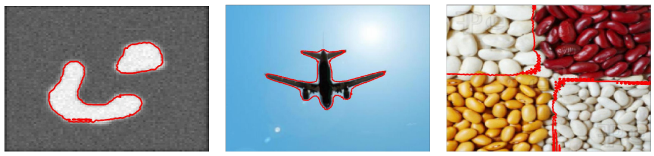


Fig. 2. Segmentation obtained as the level set of ϕ.

7. REFERENCES

[1] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International Journal on Computer Vision*, vol. 22, pp. 61–79, 1997.

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

[3] N. Sochen, R. Kimmel, and R. Malladi, "A general framework for low level vision," *IEEE Trans. Image Process.*, vol. 7, no. 3, pp. 310 –318, 1998.

[4] R. Kimmel, R. Malladi, and N. Sochen, "Images as embedded maps and minimal surfaces: Movies, color, texture, and volumetric medical images," *International Journal of Computer Vision*, vol. 39, pp. 111–129, 2000.

[5] X. Bresson, P. Vandergheynst, and J-P Thiran, "Multiscale active contours," *International Journal of Computer Vision*, vol. 70, pp. 197–211, 2006.

[6] I. Bogdanova, X. Bresson, J.-P. Thiran, and P. Vandergheynst, "Scale space analysis and active contours for omnidirectional images," *IEEE Trans. Image Process.*, vol. 16, no. 7, pp. 1888 –1901, 2007.

[7] D. Zosso, X. Bresson, and J.-P. Thiran, "Geodesic active fields - a geometric framework for image registration," *IEEE Trans. Image Process.*, , no. 99, 2010.

[8] S. Osher and N. Paragios, "Level set methods," in *Geometric Level Set Methods in Imaging, Vision, and Graphics*, pp. 3–20. 2003.

[9] C. Li, C. Xu, C. Gui, and M.D. Fox, "Level set evolution without re-initialization: A new variational formulation," *Proceedings of CVPR*, vol. 1, pp. 430–436, 2005.