IMAGE SEGMENTATION USING CLIQUE BASED SHAPE PRIOR AND THE MUMFORD SHAH FUNCTIONAL

Fredrick Park

Mathematics Department, Whittier College,
Whittier, CA 90601
Email: `fepark@whittier.edu`

ABSTRACT

A novel shape prior segmentation model is proposed that utilizes the cliques invariant signature along with a polygonal piecewise constant implementation of the Mumford-Shah Functional. The model will be shown to be useful in the context of difficult segmentation problems including and not limited to segmenting objects amidst clutter, or recognizing objects that contain components with largely varying non-uniform image intensities. In addition, the model will also be shown to be effective for image disocclusion. Lastly, the proposed model can accomplish all the aforementioned tasks both efficiently and with near automation.

***Index Terms*—** active contours, level set methods, disocclusion, Chan-Vese model, invariant signature

1 Introduction

Shape prior segmentation is an elementary problem in image processing that takes in a priori information about a known shape to reduce ambiguity in image segmentation problems. It merges many concepts from cognitive psychology, computer science, engineering, and mathematics. In general, human cognition relies on prior experience to recognize objects; particularly amidst clutter or with occlusion. Shape prior segmentation attempts to tackle a similar problem where a particular shape signature is incorporated into existing segmentation models to aid in capturing appropriate boundaries of objects of interest.

From a mathematical perspective, standard (non shape prior) image segmentation is an arduous problem since there are many different possible segmentations occurring at differing scales. Let us consider the celebrated Mumford and Shah (MS) [1] model in 2-phase form which is also known as the Chan-Vese (CV) Model [2]:

$$E_{MS}(\Sigma, c_1, c_2) = \text{Per}(\Sigma) + \lambda \int_{\Sigma} (c_1 - f)^2 d\vec{x} + \int_{\Sigma^c} (c_2 - f)^2 d\vec{x}. \quad (1)$$

Here, f is a given gray scale image, Σ denotes a region, and Σ^c the outside of that region. The key idea is to minimize the above functional (1) by matching two regions (constants) in the L^2 sense while also minimizing the perimeter of the boundaries between them. The MS model is one of the most studied and successful segmentation models in image processing. However, one particular caveat of this model, is that the functional has many local minimizers that can depend on initial starting conditions corresponding to incorrect segmentations at differing scales. Nonetheless, there are many recent convexifications of the MS/CV models, see [3, 4] and references therein.

The key difficulties in building a successful shape prior segmentation model are two-fold. Developing and/or utilizing a shape signature that is invariant under rigid motions while also possessing the ability to uniquely identifying a broad class of shapes is paramount. Many useful signatures and references thereof can be found in [5, 6, 7]. Along with invariance, the ability to incorporate such shape signatures naturally and efficiently into existing segmentation models is mandatory for applications.

The key component of the approach in this paper is to incorporate a modified two dimensional version of the cliques signature proposed by Kimmel et al. [5, 6, 7] into a polygonal implementation of the CV model. Moreover, an efficient and nearly automated numerical algorithm for successfully segmenting images in the aforementioned difficult settings will be shown. The simplified version of the cliques energy used is conducive to the polygonal form of the CV model in that it is fast, robust to noise, and easy to implement; all the while the signature strongly identifies given shapes. Numerical examples will be shown to support these claims. The contributions of this paper are summarized as:

- Incorporate the Cliques invariant signature into the CV Model for shape prior segmentation.
- Introduce a polygonal implementation of the CV Model that evolves only points on the polygonal curve.
- Introduce a fast, efficient, and nearly automated algorithm to minimize the proposed energy in the context of difficult segmentation and disocclusion problems.

This work was partially supported by grants from the NSF under contract DMS-0713767.

The two closest related works are by Cremers et al. [8] and Dŏgam et al. [9] where both show accurate segmentations by incorporating shape priors into the MS model. The work by Cremers differs from ours in that they solve a variant of the CV model where the length term of a curve C: $L(C)$ is replaced by $\int_0^1 \left(\frac{\partial C}{\partial s}\right)^2 ds$ and they evolve a cubic spline curve. In our approach, we evolve the polygonal curve directly and solve the CV model in its original form. In the work by Dogan et al. the authors evolve a function $\vec{X} : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ such that $\vec{X}$ restricted to a curve Γ yields the (x, y) coordinates of a point on the curve. They utilize a multilevel finite element method operating on the entire grid which is significantly more costly then the proposed method which simply evolve points on a polygonal curve. Other related work concerning shape prior in deformable models using level sets can be found in [10, 11, 12], using moments based descriptors [13], and using explicit descriptors of the contour [14]. Related work on region-based active contour models using polygonal and B-spline implementations can be found in [15, 16, 17, 18]. Some related work on geometrically invariant shape signatures involving Fourier descriptors can be found in [19], Legendre moments [13], or generalized cone representation [20].

The paper is organized as follows: in section 2 we introduce the proposed model. Section 3 discusses the energy minimization of the model along with a numerical implementation and final resulting algorithm. We demonstrate the successfulness of our method via numerical experiments in section 4. In section 5, we give conclusions and future work.

2 Proposed Model

2.1 Invariant Signature: Cliques Energy

Consider the polygonal representation of a given shape S: $S = \{\vec{r}_i\}_{i=1}^N$ where, $\vec{r}_i = (x_i^r, y_i^r)$. Let d_{ij} denote the intervertex distances of a reference (i.e. library) shape that is:

$$d_{ij} = |\vec{r}_i - \vec{r}_j|\,. \tag{2}$$

Then, $[d_{ij}]$ is a symmetric matrix only depending on the given reference shape. The idea of cliques was proposed by Elad and Kimmel [7] where the authors developed a bending invariant representation which is an embedding of the geometric structure of a given surface in a small dimensional Euclidean space where geodesic distances are approximated by Euclidean ones. We consider the simplified version of the intervertex distances of uniformly distributed points on the boundary of a two dimensional shape. It is known that the above signature (2) uniquely identifies convex shapes and is useful in its unabridged form for surface classification [7] and face recognition [5, 6]. In the context of segmentation, experimentally, we have found that the signature strongly identifies non-convex shapes.

Let us now consider an evolving polygon:

$$\Sigma_p = \Sigma_p(t) := \{\vec{p}_1(t), \, \ldots, \, \vec{p}_N(t)\} \tag{3}$$

where $\vec{p}_j(t) = (x_j(t), \, y_j(t))$ denotes the vertices of the polygon, arranged in counterclockwise order, with the same number of vertices as the reference shape S. Our proposed shape signature utilizing the intervertex distances of a reference shape is then given by:

$$\inf_{\Sigma_p, \, s} \left\{ E_C(\Sigma_p, s) = \sum_{i,j=1}^{N} \left(|\vec{p}_i - \vec{p}_j|^2 - s d_{ij}^2 \right)^2 \right\}. \tag{4}$$

The (symmetric) matrix $[d_{ij}]$ is computed only once at the beginning of the segmentation process as it only depends on the given reference shape. The parameter: 's' is for scale invariance, and is to be minimized over as well. The shape signature (4) has the invariances: (a) invariance with respect to rigid motions and (b) scale invariance. For the remainder of the paper, we call (4) the 'Cliques' energy. We also remark that one we can also average shape priors with respect to the signature (4), see [21] for a related methodology.

2.2 Proposed Shape Prior Segmentation Model

We propose to utilize the cliques invariant signature (4) with a polygonal implementation of the piecewise constant Mumford Shah Functional. Let Σ_p be an evolving polygon defined in (3) then, the model has the following formulation:

$$\begin{aligned} &E(\Sigma_p, c_1, c_2, s) = E_{MS} + E_C = \\ &\text{Per}(\Sigma_p) + \lambda \int_{\Sigma_p^{in}} (c_1 - f)^2 d\vec{x} + \lambda \int_{\Sigma_p^{out}} (c_2 - f)^2 d\vec{x} \\ &+ \alpha \sum_{i,j=1}^{N} \left(|\vec{p}_i - \vec{p}_j|^2 - s d_{ij}^2 \right)^2 := E_{MSWS}. \end{aligned} \tag{5}$$

Here, Σ_p^{in} and Σ_p^{out} represent the regions inside and outside the polygonal curve Σ_p respectively. Moreover, c_1 and c_2 are constants depending on the average image intensities inside Σ_p^{in} and Σ_p^{out} respectively and 's' is a scale parameter depending on the recovered segmentation feature. The parameter α dictates the strength of the shape term and lastly, 'MSWS' stands for Mumford Shah With Shape.

3 Energy Minimization and Numerics

3.1 Normal Speed of the Mumford-Shah Model

Although we will not use level sets for the numerical implementation of the proposed model (5), we will need the correct normal speed for the evolving curve in gradient descent for the final algorithm. Chan and Vese (CV)[2] proposed a level set formulation of the Piecewise Constant MS Model:

$$\begin{aligned} &E_{CV}(\phi, c_1, c_2) = \int_{\Sigma} |\nabla H_\epsilon(\phi)| \\ &+ \lambda \int_{\Sigma} H_\epsilon(\phi)(c_1 - f)^2 + (1 - H_\epsilon(\phi))(c_2 - f)^2 d\vec{x}. \end{aligned} \tag{6}$$

The regions Σ and Σ^c are now represented by a regularized Heaviside function H_ϵ of a level set function ϕ, where the zero level set $\{\phi = 0\}$ represents the boundary of Σ. The associated gradient descent flow to minimize (6) is given by:

$$\phi_t = |\nabla\phi| \left\{ \nabla \cdot \frac{\nabla\phi}{|\nabla\phi|} + \lambda \left[(c_2 - f)^2 - (c_1 - f)^2 \right] \right\} \quad (7)$$

where we read off the correct normal speed of the curve in gradient descent as the following:

$$v = \kappa + \lambda \left[(c_2 - f)^2 - (c_1 - f)^2 \right]. \quad (8)$$

Here, κ denotes the curvature of the evolving curve.

3.2 Polygonal Perimeter and Fidelity Terms

From (8), the contribution to the scalar normal speed of the perimeter term is the Euclidean curvature κ. For an evolving curve $C(p) = C(p,t)$ depending on parameter p and time variable t, $C : [0,1] \times [0,\infty) \rightarrow \mathbb{R}^2$, the gradient descent flow for minimizing the perimeter term is:

$$C_t = \kappa \vec{N} := \vec{v}^1 \quad (9)$$

We refer the reader to [22] for a detailed derivation. Now, the minimization of the proposed model (5) is to be carried out over the set Σ and the two constants c_1 and c_2. In the polygonal curve based implementation, we restrict ourselves to the case where the set Σ is the interior of a polygon of N sides Σ_p defined in (3). Let D^+ and D^- denote the forward and backward discrete difference operators: $D^+\xi_j = \xi_{j+1} - \xi_j$ and $D^-\xi_j = \xi_j - \xi_{j-1}$ respectively. The contribution of the fidelity term in the proposed model (5) to the normal velocity of the polygonal curve is:

$$\vec{v}_j^2 = \lambda \left[(c_2 - f)^2 - (c_1 - f)^2 \right] \vec{n}_j, \quad (10)$$

where $\vec{n}_j$ is the unit normal vector at the j-th vertex:

$$\vec{n}_j = \frac{1}{\sqrt{(D^+x_j)^2 + (D^+y_j)^2}} \left(-D^+y_j, D^+x_j \right). \quad (11)$$

Finally, putting the contribution of the perimeter $\vec{v}^1$ (9) term and fidelity $\vec{v}^2$ (10) term together, and multiplying by a factor of $\sqrt{(D^+x_j)^2 + (D^+y_j)^2}$ (which improves stability and still leads to a descent direction), we obtain the following ODE system describing the time evolution of vertices $(x_j(t),\, y_j(t))$ of the polygon:

$$\begin{aligned} &\frac{d}{dt}(x_j(t),\, y_j(t)) = \\ &\left(D^- \left(\frac{D^+x_j}{\sqrt{(D^+x_j)^2 + (D^+y_j)^2}} \right), D^- \left(\frac{D^+y_j}{\sqrt{(D^+x_j)^2 + (D^+y_j)^2}} \right) \right) \\ &+ \lambda \left[(c_2 - f)^2 - (c_1 - f)^2 \right] (-D^+y_j,\, D^+x_j) \\ &:= \vec{v}_j^{per} + \vec{v}_j^{fid}. \end{aligned} \quad (12)$$

Here, $\vec{v}_j^{per}$ and $\vec{v}_j^{fid}$ denote the velocities corresponding to the perimeter and fidelity terms for the polygonal piecewise constant implementation of the CV model respectively.

3.3 Update of the constants c_1 and c_2

In the original CV model, the constants c_1 and c_2 are unknown and need to be solved for. The optimal choice of the constants is well known:

$$c_1 = \frac{\int_\Sigma f dxdy}{|\Sigma|} \quad \text{and} \quad c_2 = \frac{\int_{\Sigma^c} f dxdy}{|\Sigma^c|} \quad (13)$$

where the notation $|\Sigma|$ denotes the area of the set Σ. In our setting, the boundary of the unknown set Σ is represented explicitly as a polygonal curve Σ_p. In order to evaluate the integrations over Σ involved in these formulas with subgrid accuracy, we express the area integrals as path integrals:

$$\int_\Sigma f\, dxdy = \int_{\partial\Sigma} F\, n_x d\sigma \quad \text{where} \quad \frac{\partial F}{\partial x} = f. \quad (14)$$

Here, n_x denotes the x-component of the outer unit normal to the boundary $\partial\Sigma$ of the region Σ. The point is that there is a natural way to approximate path integrals on polygonal curves numerically. Our algorithm for computing the integral $\int_\Sigma f\, d\vec{x}$ is follows:

- Integrate f in the x-direction using the trapezoidal rule to define its primitive F on the grid so that $\frac{\partial f}{\partial x} = F$.
- Approximate the area integral as follows:

$$\begin{aligned} \int_\Sigma f\, dxdy &= \int_{\partial\Sigma} F\, n_x d\sigma \\ &\approx \sum_{j=1}^{N} \frac{[F(x_j,y_j)+F(x_{j+1},y_{j+1})]\; D^+y_j}{2\sqrt{(D^+x_j)^2+(D^+y_j)^2}}. \end{aligned} \quad (15)$$

The scheme involves interpolating F since (x_j, y_j) may not lie on grid points. With proper interpolation, the method is *second order accurate* and the primitive (i.e. F) calculation needs only to be done once and not repeated during the iterations of gradient descent. Finally, $|\Sigma|$ can also be found with sub-grid accuracy by using the above formula with $f \equiv 1$.

3.4 Variation of the Shape Prior and Scale Parameter Optimization

Variation of the shape prior (4) with respect to the vertex locations $\vec{p}_j$ of the evolving polygon is given by:

$$\begin{aligned} &\frac{\partial}{\partial \vec{p}_k} \alpha \sum_{i,j=1}^{N} \left(|\vec{p}_i - \vec{p}_j|^2 - s\, d_{i,j}^2 \right)^2 = \\ &8\alpha \sum_{j=1}^{N} \left(|\vec{p}_k - \vec{p}_j|^2 - s\, d_{kj}^2 \right) (\vec{p}_k - \vec{p}_j) \cdot \vec{p}_k. \end{aligned} \quad (16)$$

In gradient descent, the shape term contributes the velocity:

$$\vec{v}_k^1 = -\alpha \sum_{j=1}^{N} \left(|\vec{p}_k - \vec{p}_j|^2 - s d_{kj}^2 \right) (\vec{p}_k - \vec{p}_j). \quad (17)$$

In order to evolve the polygon in the normal direction to minimize the shape term, we project the velocity onto the unit normal to obtain the normal speed. Thus, the final contribution of the velocity of the shape term is:

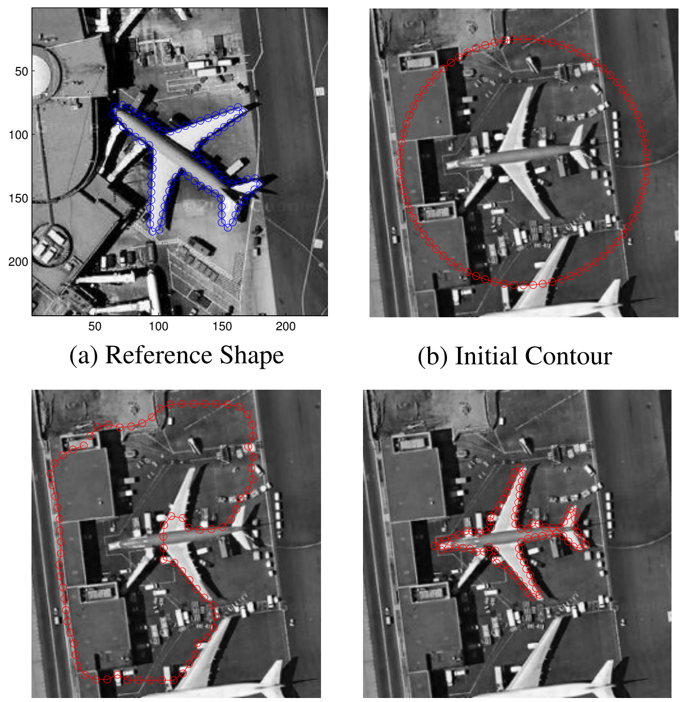
$$\vec{v}_k^{shape} = (\vec{v}_k^1 \cdot \vec{n}_k)\vec{n}_k. \quad (18)$$

Optimization with respect to the scale parameter s yields:

$$s = \frac{\sum_{i,j} |\vec{p}_i - \vec{p}_j|^2 d_{ij}^2}{\sum_{i,j} d_{ij}^4}. \quad (19)$$

3.5 Final Proposed Algorithm

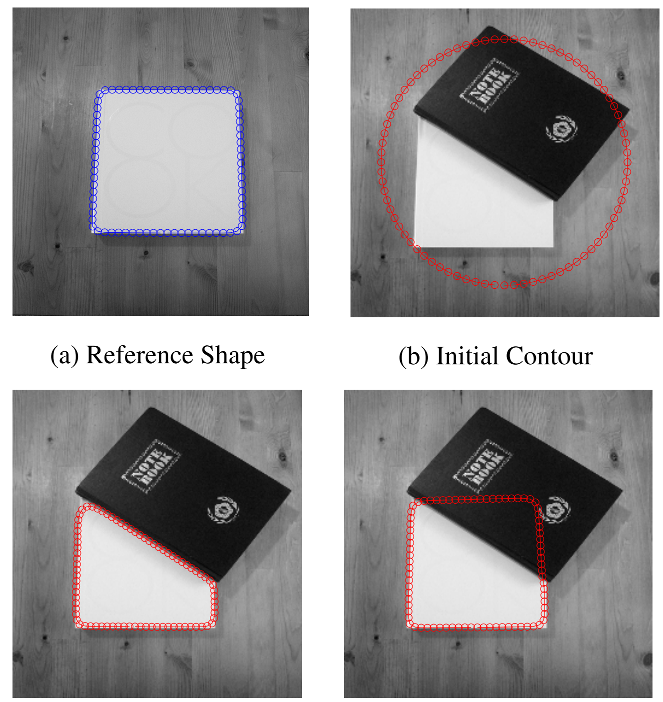
Finally, putting everything together, the complete gradient descent curve evolution equation for the proposed model is the



(a) Reference Shape (b) Initial Contour

(c) Shape Param. = 0.0 (d) Shape Param. = 1.0

Fig. 1. *Segmentation Results with Learned Reference Shape*



(a) Reference Shape (b) Initial Contour

(c) Shape Param. = 0.0 (d) Shape Param. = 1.0

Fig. 2. *Segmentation Results with Learned Reference Shape*

ODE system:

$$\dot{\vec{p}}_k = \vec{v}_k^{per} + \vec{v}_k^{fid} + \vec{v}_k^{shape} \qquad (20)$$

where the velocities were obtained from the variations in the perimeter and fidelity terms (12), and shape term (18) in the proposed model. The first and second terms move the curve in a direction to minimize the perimeter and data fidelity terms, while the last shape term moves the curve to best match the reference shape. An explicit time marching scheme (iterated until steady state) can be used for the gradient descent curve evolution equation above (20) as follows:

$$\vec{p}_k^{n+1} = \vec{p}_k^{n} + dt\left(\vec{v}_k^{per} + \vec{v}_k^{fid} + \vec{v}_k^{shape}\right). \qquad (21)$$

4 Numerical Experiments

In this first example, we segment what's considered a difficult case. In Fig. 1 (a), a reference image and shape prior (in blue) are observed while in (b), a clean airplane image (to be segmented) with initial contour (red) is seen. In (b), the color of the fuselage matches the surrounding tarmac making this case arduous due to the ambiguities of the region to be segmented. Further compounding this case is the flight equipment around the plane (clutter) which has similar contrast to the planes wings. In addition, the white fuselage and wings of the adjacent plane matches the contrast of the wings of the plane to be segmented. The result from the standard (no shape) CV model is observed in Fig. 1 (c), where the plane is incorrectly segmented while in Fig. 1 (d), the shape term is set to 1.0 and the correct plane is accurately segmented despite the difficulties. Lastly, we remark that the scale parameter 's' in this experiment was manually chosen over a range of scales. In this next example we demonstrate how the proposed model can be used for image disocclusion. If we let $R \subseteq \mathbb{R}^2$ denote the region to be disoccluded, we simply make the small change in the proposed model (5) by replacing λ by $\lambda\,(1-\chi_R)$ with χ_R the indicator function of the region R. A book image is observed in Fig. 2 (a) with the reference shape shown as a blue contour. In Fig. 2 (b), the same book at a different scale with an occlusion is seen along with the initial contour in red. Here, the corner of the white book is occluded by the black one. The result from the standard CV model (no shape) is seen in 2 (c) where the corner of the book remains occluded. In turn, we use this contour as an initial contour for our shape prior model where, in Fig. 2 (c), we set the shape strength to 1.0 and the book is successfully disoccluded. The scale parameter is automatically found during the curve evolution using the condition in (19).

5 Conclusions and Future Work

We proposed an efficient and nearly automated shape prior segmentation model. Compelling numerical results illustrating the successfulness of the method in difficult settings were presented. In addition, we adapted the model for disocclusion. Future work involves finding a completely convex formulation of the proposed model.

6 References

[1] David Mumford and Jayant Shah, “Optimal approximations by piecewise smooth functions and associated variational problems,” *Communications on pure and applied mathematics*, vol. 42, no. 5, pp. 577–685, 1989.

[2] T. F. Chan and L. A. Vese, “Active contours without edges,” *Trans. Img. Proc.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.

[3] Xavier Bresson, Selim Esedoglu, Pierre Vandergheynst, Jean-Philippe Thiran, and Stanley Osher, “Fast global minimization of the active contour/snake model,” *J. Math. Imaging Vis.*, vol. 28, no. 2, pp. 151–167, June 2007.

[4] Ethan S. Brown, Tony F. Chan, and Xavier Bresson, “Completely convex formulation of the chan-vese image segmentation model,” *Int. J. Comput. Vision*, vol. 98, no. 1, pp. 103–121, May 2012.

[5] Alexander M Bronstein, Michael M Bronstein, and Ron Kimmel, “Three-dimensional face recognition,” *International Journal of Computer Vision*, vol. 64, no. 1, pp. 5–30, 2005.

[6] A. M. Bronstein, M. M. Bronstein, and R. Kimmel, “Expression-invariant representations of faces,” *Trans. Img. Proc.*, vol. 16, no. 1, pp. 188–197, Jan. 2007.

[7] A. Elad and R. Kimmel, “On bending invariant signatures for surfaces,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 10, pp. 1285–1295, Oct. 2003.

[8] Daniel Cremers, Timo Kohlberger, and Christoph Schnörr, “Nonlinear shape statistics in mumford-shah based segmentation,” in *Proceedings of the 7th European Conference on Computer Vision-Part II*, London, UK, UK, 2002, ECCV ’02, pp. 93–108, Springer-Verlag.

[9] Günay Doğan, Pedro Morin, and Ricardo H. Nochetto, “A variational shape optimization approach for image segmentation with a mumford-shah functional,” *SIAM J. Sci. Comput.*, vol. 30, no. 6, pp. 3028–3049, Oct. 2008.

[10] Xavier Bresson, Pierre Vandergheynst, and Jean-Philippe Thiran, “A variational model for object segmentation using boundary information and shape prior driven by the mumford-shah functional,” *Int. J. Comput. Vision*, vol. 68, no. 2, pp. 145–162, June 2006.

[11] Tony Chan and Wei Zhu, “Level set based shape prior segmentation,” in *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05) - Volume 2 - Volume 02*, Washington, DC, USA, 2005, CVPR ’05, pp. 1164–1170, IEEE Computer Society.

[12] Mikael Rousson and Nikos Paragios, “Prior knowledge, level set representations & visual grouping,” *International Journal of Computer Vision*, vol. 76, no. 3, pp. 231–243, 2008.

[13] Alban Foulonneau, Pierre Charbonnier, and Fabrice Heitz, “Affine-invariant geometric shape priors for region-based active contours,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 8, pp. 1352–1357, Aug. 2006.

[14] Muriel Gastaud, Michel Barlaud, and Gilles Aubert, “Combining shape prior and statistical features for active contour segmentation,” *IEEE TRANS. CIRCUITS SYST. VIDEO TECHNOL*, vol. 14, pp. 726–734, 2004.

[15] Christophe Chesnaud, Philippe Réfrégier, and Vlady Boulet, “Statistical region snake-based segmentation adapted to different physical noise models,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 11, pp. 1145–1157, Nov. 1999.

[16] É. Debreuve, M. Gastaud, M. Barlaud, and G. Aubert, “Using the shape gradient for active contour segmentation: From the continuous to the discrete formulation,” *J. Math. Imaging Vis.*, vol. 28, no. 1, pp. 47–66, May 2007.

[17] Frédéric Galland, Arnaud Jaegler, Marc Allain, David Savéry, and Philippe Réfrégier, “Smooth contour coding with minimal description length active grid segmentation techniques,” *Pattern Recogn. Lett.*, vol. 32, no. 5, pp. 721–730, Apr. 2011.

[18] Ganesh Sundaramoorthi and Anthony Yezzi, “Global regularizing flows with topology preservation for active contours and polygons,” *Image Processing, IEEE Transactions on*, vol. 16, no. 3, pp. 803–812, 2007.

[19] Stéphane Derrode, Mohamed Ali Charmi, and Faouzi Ghorbel, “Fourier-based invariant shape prior for snakes,” in *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*. IEEE, 2006, vol. 2, pp. II–II.

[20] Tammy Riklin-Raviv, Nahum Kiryati, and Nir Sochen, “Prior-based segmentation and shape registration in the presence of perspective distortion,” *International Journal of Computer Vision*, vol. 72, no. 3, pp. 309–328, 2007.

[21] Martin Rumpf and Benedikt Wirth, “A nonlinear elastic shape averaging approach,” *SIAM Journal on Imaging Sciences*, vol. 2, no. 3, pp. 800–833, 2009.

[22] Vicent Caselles, Ron Kimmel, and Guillermo Sapiro, “Geodesic active contours,” *Int. J. Comput. Vision*, vol. 22, no. 1, pp. 61–79, Feb. 1997.