An Active Contour Model without Edges

Tony Chan and Luminita Vese

Department of Mathematics, University of California, Los Angeles,
520 Portola Plaza, Los Angeles, CA 90095-1555
chan,lvese@math.ucla.edu
WWW home page: http://www.math.ucla.edu

Abstract. In this paper, we propose a new model for active contours to detect objects in a given image, based on techniques of curve evolution, Mumford-Shah functional for segmentation and level sets. Our model can detect objects whose boundaries are not necessarily defined by gradient. The model is a combination between more classical active contour models using mean curvature motion techniques, and the Mumford-Shah model for segmentation. We minimize an energy which can be seen as a particular case of the so-called minimal partition problem. In the level set formulation, the problem becomes a "mean-curvature flow"-like evolving the active contour, which will stop on the desired boundary. However, the stopping term does not depend on the gradient of the image, as in the classical active contour models, but is instead related to a particular segmentation of the image. Finally, we will present various experimental results and in particular some examples for which the classical snakes methods based on the gradient are not applicable.

1   Introduction

The basic idea in active contour models or snakes is to evolve a curve, subject to constraints from a given image u_0, in order to detect objects in that image. For instance, starting with a curve around the object to be detected, the curve moves toward its interior normal under some constraints from the image, and has to stop on the boundary of the object.

Let Ω be a bounded and open subset of $I\!\!R^2$, with $\partial\Omega$ its boundary. Let u_0 be a given image, as a bounded function defined on $\overline{\Omega}$ and with real values. Usually, $\overline{\Omega}$ is a rectangle in the plane and u_0 takes values between 0 and 255. Denote by $C(s) : [0, 1] \to I\!\!R^2$ a piecewise C^1 parameterized curve.

In all the classical snakes and active contour models (see for instance [7], [3], [9], [4]), an edge detector is used to stop the evolving curve on the boundaries of the desired object. Usually, this is a positive and regular edge-function $g(|\nabla u_0|)$, decreasing such that $\lim_{t\to\infty} g(t) = 0$. For instance,

$$g(|\nabla u_0|) = \frac{1}{1 + |\nabla G_\sigma * u_0|^2},$$

where $G_\sigma * u_0$ is the convolution of the image u_0 with the Gaussian $G_\sigma(x, y) = \sigma^{-1/2}\exp(-|x^2 + y^2|/4\sigma)$ (a smoother version of u_0). The function $g(|\nabla u_0|)$ will

be strictly positive in homogeneous regions, and near zero on the edges. The evolving curve moves by a variant of the mean curvature motion [14] with the edge-function $g(|\nabla u_0|)$ as an extra factor in the velocity.

All these classical snakes or active contour models rely on this edge-function g, depending on the gradient $|\nabla u_0|$ of the image, to stop the curve evolution. Therefore, these models can detect only objects with edges defined by gradient. Also, in practice, the discrete gradients are bounded and then the stopping function g is never zero on the edges, and the curve may pass through the boundary. On the other hand, if the image u_0 is noisy, then the isotropic smoothing Gaussian has to be strong, which will smooth the edges too. In this paper, we propose a different active contour model, without a stopping edge-function, i.e. a model which is not based on the gradient of the image u_0 for the stopping process. The stopping term is based on Mumford-Shah segmentation techniques [13]. In this way, we obtain a model which can detect contours both with or without gradient, for instance objects with very smooth boundaries or even with discontinuous boundaries. For a discussion on different types of contours, we refer the reader to [6].

The outline of the paper is as follows. In the next section we introduce our model as an energy minimization and discuss the relationship with the Mumford-Shah functional for segmentation. In Section 3, we formulate the model in terms of level set functions, compute the associated Euler-Lagrange equations, and discuss the algorithm. We end the paper validating our model by numerical results. We show in particular how we can detect contours without gradient or cognitive contours [6], for which the classical models are not applicable, and also how we can automatically detect interior contours.

Before describing our proposed model, we would like to refer the reader to the works [10] and [11] for shape recovery using level sets and edge-function, and to more recent and related works by [19], [17], and [8].

Finally, we would also like to mention [21] and [12] on shape reconstruction from unorganized points, and to the recent works [15] and [16], where a probability based geodesic active region model combined with classical gradient based active contour techniques is proposed.

2   Description of the model

Let C be the evolving curve. We denote by c_1 and c_2 two constants, representing the averages of u_0 "inside" and "outside" the curve C.

Our model is the minimization of an energy based-segmentation. Let us first explain the basic idea of the model in a simple case. Assume that the image u_0 is formed by two regions of approximatively piecewise-constant intensities, of distinct values u_0^i and u_0^o. Assume further that the object to be detected is represented by the region with the value u_0^i and let denote his boundary by $\mathcal{C}$. Then we have $u_0 \approx u_0^i$ inside the object (inside $\mathcal{C}$) and $u_0 \approx u_0^o$ outside the object (outside $\mathcal{C}$). Now let us consider the following "fitting energy", formed by

two terms:

$$F_1(C) + F_2(C) = \int_{inside(C)} |u_0 - c_1|^2 dxdy + \int_{outside(C)} |u_0 - c_2|^2 dxdy,$$

where C is any other variable curve. We say that the boundary of the object $\mathcal{C}$ is the minimizer of the fitting energy:

$$\inf_C \left\{ F_1(C) + F_2(C) \right\} \approx 0 \approx F_1(\mathcal{C}) + F_2(\mathcal{C}).$$

This can be seen easily. For instance, if the curve C is outside the object, then $F_1(C) > 0$ and $F_2(C) \approx 0$. If the curve C is inside the object, then $F_1(C) \approx 0$ but $F_2(C) > 0$. Finally, the fitting energy will be minimized if the $C = \mathcal{C}$, i.e. if the curve C is on the boundary of the object. These remarks are illustrated in Fig. 1.

| $F_1(C) > 0$ | $F_1(C) > 0$ | $F_1(C) \approx 0$ | $F_1(C) \approx 0$ |
| $F_2(C) \approx 0$ | $F_2(C) > 0$ | $F_2(C) > 0$ | $F_2(C) \approx 0$ |
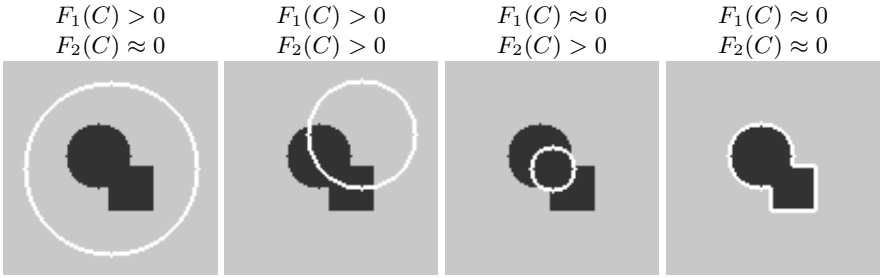
Fig. 1. Consider all possible cases in the position of the curve. The "fitting energy" is minimized only for the case when the curve is on the boundary of the object.

Therefore, in our active contour model we will minimize this fitting energy and we can add some regularizing terms, like the length of C and/or the area inside C. We introduce the energy $F(C, c_1, c_2)$ by:

$$F(C, c_1, c_2) = \mu \cdot \text{length}(C) + \nu \cdot \text{area}(insideC)$$
$$+ \lambda_1 \int_{inside(C)} |u_0 - c_1|^2 dxdy + \lambda_2 \int_{outside(C)} |u_0 - c_2|^2 dxdy,$$

where c_1 and c_2 are constant unknowns, and $\mu > 0$, $\nu \geq 0$, $\lambda_1, \lambda_2 > 0$ are fixed parameters.

In almost all our computations, we take $\nu = 0$ and $\lambda_1 = \lambda_2$. Of-course that one of these parameters can be "eliminated", by fixing it to be 1. In almost all our computations, we take $\nu = 0$ and $\lambda_1 = \lambda_2$. The area term in the energy can be used for instance when we may need to force the curve to move only inside.

In order to balance the terms and their dimensions in the energy, if d is the unit distance in the $\Omega-$plane, then μ has to be measured in units of $(size\ of\ u_0)^2 \cdot d$, and ν has to be measured in units of $(size\ of\ u_0)^2$.

Finally, we consider the minimization problem:

$$\inf_{C,c_1,c_2} F(C, c_1, c_2). \tag{1}$$

2.1 Relation with the Mumford-Shah functional for segmentation

The Mumford-Shah functional for segmentation is [13]:

$$F^{MS}(u, C) = \int_{\Omega \setminus C} (\alpha |\nabla u|^2 + \beta |u - u_0|^2) dx dy + \text{length}(C), \tag{2}$$

where α, β are positive parameters. The solution image u obtained by minimizing this functional is formed by smooth regions R_i and with sharp boundaries, denoted here by C.

A reduced form of this problem, as it was pointed out by D. Mumford and J. Shah in [13], is simply the restriction of F^{MS} to piecewise constant functions u, i.e. $u = c_i$ with c_i a constant, on each connected component R_i of $\Omega \setminus C$. Therefore, the constants c_i are in fact the averages of u_0 on each R_i. The reduced case is called the minimal partition problem.

Our active contour model is a particular case of the minimal partition problem, in which we look for the best approximation u of u_0, as a function taking only two values, namely:

$$u = \begin{cases} \text{average}(u_0) \text{ inside } C \\ \text{average}(u_0) \text{ outside } C, \end{cases}$$

and with one edge C, represented by the snake or the active contour.

This particular case of the minimal partition problem can be formulated and solved using the level set method [14]. This is presented in the next section.

2.2 The level set formulation of the model

In the level set method [14], an evolving curve C is represented by the zero level set of a Lipschitz continuous function $\phi : \Omega \to \mathbb{R}$. So, $C = \{(x, y) \in \Omega : \phi(x, y) = 0\}$, and we choose ϕ to be positive inside C and negative outside C. For the level set formulation of our variational active contour model we essentially follow [20]. Therefore, we replace the unknown variable C by the unknown variable ϕ and the new energy, still denoted by $F(\phi, c_1, c_2)$, becomes:

$$F(\phi, c_1, c_2) = \mu \cdot \text{length}\{\phi = 0\} + \nu \cdot \text{area}\{\phi \geq 0\}$$
$$+ \lambda_1 \int_{\phi \geq 0} |u_0 - c_1|^2 dx dy + \lambda_2 \int_{\phi < 0} |u_0 - c_2|^2 dx dy.$$

Using the Heaviside function H defined by

$$H(x) = \begin{cases} 1, \text{ if } x \geq 0 \\ 0, \text{ if } x < 0 \end{cases}$$

and the one-dimensional Dirac measure δ concentrated at 0 and defined by

$$\delta(x) = \frac{d}{dx}H(x) \text{ (in the sense of distributions)},$$

we express the terms in the energy F in the following way:

$$\begin{cases} \text{length}\{\phi = 0\} = \int_\Omega |\nabla H(\phi)| = \int_\Omega \delta(\phi)|\nabla\phi|, \\ \text{area}\{\phi \geq 0\} \quad = \int_\Omega H(\phi)dxdy, \end{cases}$$

and

$$\begin{cases} \int_{\phi \geq 0} |u_0 - c_1|^2 dxdy = \int_\Omega |u_0 - c_1|^2 H(\phi)dxdy \\ \int_{\phi < 0} |u_0 - c_2|^2 dxdy = \int_\Omega |u_0 - c_2|^2 (1 - H(\phi))dxdy. \end{cases}$$

Then the energy $F(\phi, c_1, c_2)$ can be written as:

$$F(\phi, c_1, c_2) = \mu \int_\Omega \delta(\phi)|\nabla\phi| + \nu \int_\Omega H(\phi)dxdy$$
$$+ \lambda_1 \int_\Omega |u_0 - c_1|^2 H(\phi)dxdy + \lambda_2 \int_\Omega |u_0 - c_2|^2 (1 - H(\phi))dxdy.$$

Keeping ϕ fixed and minimizing the energy $F(\phi, c_1, c_2)$ with respect to the constants c_1 and c_2, it is easy to express these constants function of ϕ by:

$$c_1(\phi) = \frac{\int_\Omega u_0 H(\phi)dxdy}{\int_\Omega H(\phi(x,y))dxdy} \qquad \text{(the average of } u_0 \text{ in } \{\phi \geq 0\}), \qquad (3)$$

$$c_2(\phi) = \frac{\int_\Omega u_0 (1 - H(\phi))dxdy}{\int_\Omega (1 - H(\phi(x,y)))dxdy} \quad \text{(the average of } u_0 \text{ in } \{\phi < 0\}). \qquad (4)$$

Keeping c_1 and c_2 fixed, and formally minimizing the energy with respect to ϕ, we obtain the Euler-Lagrange equation for ϕ (parameterizing the descent direction by an artificial time):

$$\begin{cases} \frac{\partial\phi}{\partial t} = \delta(\phi)\left[\mu\text{div}\left(\frac{\nabla\phi}{|\nabla\phi|}\right) - \nu - \lambda_1(u_0 - c_1)^2 + \lambda_2(u_0 - c_2)^2\right] \text{ in } \Omega, \\ \frac{\delta(\phi)}{|\nabla\phi|}\frac{\partial\phi}{\partial n} = 0 \text{ on } \partial\Omega. \end{cases}$$

In practice, we have to consider slightly regularized versions of the functions H and δ, denoted here by H_ε and δ_ε, such that $\delta_\varepsilon(x) = H'_\varepsilon(x)$.

A first possible regularization by C^2 and respectively C^1 functions, as proposed for instance in [20], is:

$$H_{1,\varepsilon}(x) = \begin{cases} 1 \text{ if } x > \varepsilon \\ 0 \text{ if } x < -\varepsilon \\ \frac{1}{2}\left[1 + \frac{x}{\varepsilon} + \frac{1}{\pi}\sin\left(\frac{\pi x}{\varepsilon}\right)\right] \text{ if } |x| \leq \varepsilon \end{cases}$$

and

$$\delta_{1,\varepsilon}(x) = H'_{1,\varepsilon}(x) = \begin{cases} 0 \text{ if } |x| > \varepsilon \\ \frac{1}{2\varepsilon}\left[1 + \cos\left(\frac{\pi x}{\varepsilon}\right)\right], \text{ if } |x| \le \varepsilon. \end{cases}$$

In our calculations, we use instead the following C^∞ regularized versions of H and δ, defined by:

$$H_{2,\varepsilon}(x) = \frac{1}{2}(1 + \frac{2}{\pi}\arctan(\frac{x}{\varepsilon})), \ \ \delta_{2,\varepsilon}(x) = H'_{2,\varepsilon}(x) = \frac{1}{\pi} \cdot \frac{\varepsilon}{\varepsilon^2 + x^2}.$$

As $\varepsilon \to 0$, both approximations converge to H and δ. The first approximations $H_{1,\varepsilon}$ and $\delta_{1,\varepsilon}$ are C^2 and respectively C^1 functions, with $\delta_{1,\varepsilon}$ with small compact support, arround the zero-level set. The second approximations $H_{2,\varepsilon}$ and $\delta_{2,\varepsilon}$ are both C^∞ functions, with $\delta_{2,\varepsilon}$ different of zero everywhere.

We want to formally explain here why we need to introduce the second approximations, instead of the first approximations, which have been used in previous papers (for instance in [20]). Because our energy is non-convex (allowing therefore many local minima), and because $\delta_{1,\varepsilon}$ has a very small compact support, the interval $[-\varepsilon, \varepsilon]$, the iterative algorithm may depend on the initial curve, and will not necessarily compute a global minimizer. In some of our tests using the first approximation, we obtained only a local minimizer of the energy. Using the second approximations, the algorithm has the tendency to compute a global minimizer. One of the reasons is that, the Euler-Lagrange equation acts only locally, on a few level curves arround $\phi = 0$ using the first approximation, while by the second approximation, the equation acts on all level curves, of course stronger on the zero level curve, but not only locally. In this way, in practice, we can obtain a global minimizer, independently of the position of the initial curve. Moreover, interior contours are automatically detected. We could also extend the motion to all level sets of ϕ replacing $\delta(\phi)$ in the equation by $|\nabla\phi|$ (this method is for instance used in [20]).

To discretize the equation in ϕ, we use a finite differences implicit scheme (we refer the reader to [1], for details).

We also need at each step to reinitialize ϕ to be the signed distance function to its zero-level curve. This procedure is standard (see [18] and [20]), and prevences the level set function to become too flat, or it can be seen as a L^∞ stability for ϕ and a rescaling.

This reinitialization procedure is made by the following evolution equation [18]:

$$\begin{cases} \psi_\tau = sign(\phi(t))(1 - |\nabla\psi|) \\ \psi(0, \cdot) = \phi(t, \cdot), \end{cases} \tag{5}$$

where $\phi(t, \cdot)$ is our solution ϕ at time t. Then the new $\phi(t, \cdot)$ will be ψ, such that ψ is obtained at the steady state of (5).

3   Experimental results

We present here numerical results using our model. For the examples in Figures 2-5, we show the image and the evolving contour (top), together with the

piecewise-constant approximations given by the averages c_1 and c_2 (bottom). In all cases, we start with a single initial closed curve. We choose the level set function ϕ to be positive "inside" the initial curve, and negative "outside" the initial curve, but in our model this choice is not important. We could consider the opposite signs, and the curve would still be attracted by the object. Also, the position of the initial curve is not important.

In Fig. 2 we show how our model can detect contours without gradient or cognitive contours (see [6]) and an interior contour automatically, starting with only one initial curve. This is obtained using our second approximations for H and δ. In Fig. 3 we consider a very noisy image. Again the interior contour of the torus is automatically detected.

In Fig. 4 we validate our model on a very different problem: to detect features in spatial point processes in the presence of substantial cluster. One example is the detection of minefields using reconnaissance aircraft images that identify many objects that are not mines. These problems are for instance solved using statistical methods (see for instance [5] and [2]). By this application, we show again that our model can be used to detect objects or features with contours without gradient. This is not possible using classical snakes or active contours based on the gradient.

We end the paper with results on two real images (Fig. 5 and 6.), illustrating all the properties of our model: detecting smooth boundaries, scaling role of the length term in the size of the detected objects, and automatic change of topology.
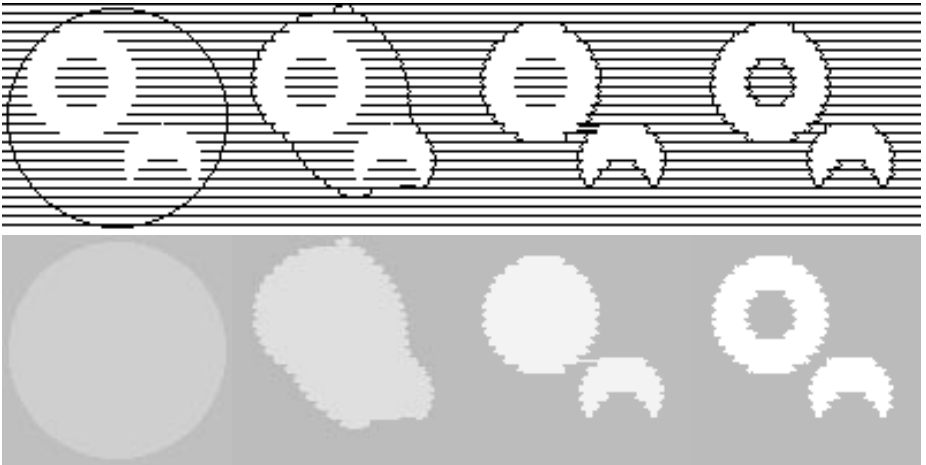
Fig. 2. Detection of different objects in a synthetic image, with various convexities and with an interior contour, which is automatically detected. Here we illustrate the fact that our model can detect edges without gradient. Top: u_0 and the contour. Bottom: the piece-wise constant approximation of u_0.
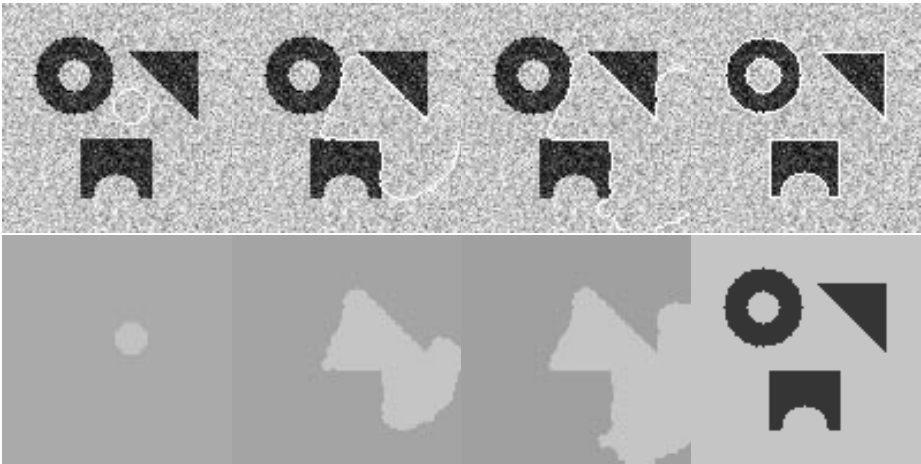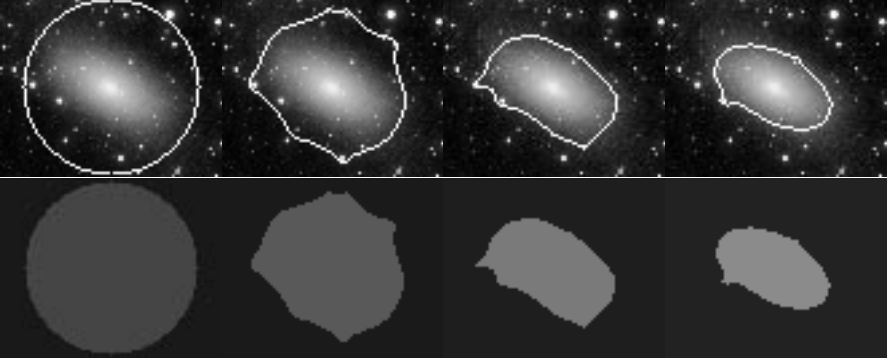
Fig. 3. Results for a very noisy image, with the initial curve not surrounding the objects. Top: u_0 and the contour. Bottom: the piece-wise constant approximation of u_0.
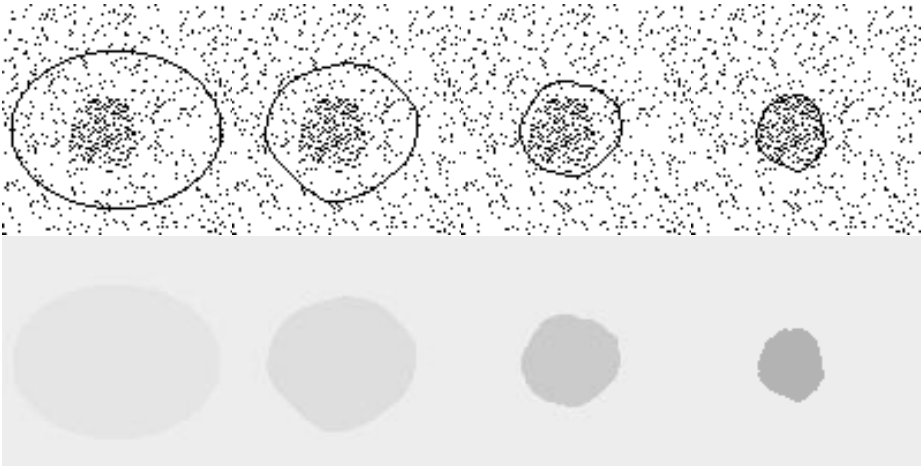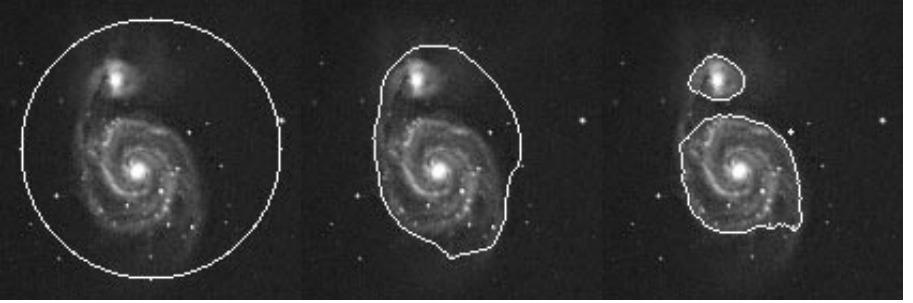

Fig. 4. Detection of a simulated minefield, with contour without gradient.

Fig. 5. Detection of a galaxy with very smooth boundaries.

Fig. 6. Detection of the contours of a galaxy.

4   Concluding remarks

In this paper we proposed an active contour model based on Mumford-Shah segmentation techniques and level set methods. Our model is not based on an edge-function, like in the classical active contour models, to stop the evolving curve on the desired boundary. We do not need to smooth the initial image, even if it is very noisy and in this way, the locations of boundaries are very well detected. Also, we can detect objects whose boundaries are not necessarily defined by gradient or with very smooth boundaries. The model automatically detects interior contours, starting with only one initial curve. The initial curve does not necessarily start around the objects to be detected. Finally, we validated our model by various numerical results.

References

1. Aubert, G., Vese, L.: A variational method in image recovery. SIAM J. Num. Anal. **34/5** (1997) 1948-1979.
2. Byers, S., Raftery, A.: Nearest-Neighbor Clutter Removal for Estimating Features in Spatial Point Processes. Journal of the American Statistical Association **93/442** (1998) 577-584.
3. Caselles, V., Catté, F., Coll, T., Dibos, F.: A geometric model for active contours in image processing. Numerische Mathematik **66** (1993) 1-31.
4. Caselles, V., Kimmel, R., Sapiro, G.: On geodesic active contours. Journal of Computer Vision **22(1)** (1997) 61-79.
5. Dasgupta, A., Raftery, A.: Detecting Features in Spatial Point Processes With Clutter via Model-Based Clustering. Journal of the American Statistical Association **98/441** (1998) 294-302.
6. Kanizsa, G.: La Grammaire du Voir. Essais sur la perception. Diderot Editeur, Arts et Sciences (1997).
7. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. International Journal of Computer Vision **1** (1988) 321-331.
8. Marquina, A., Osher, S.: Explicit Algorithms for a New Time Dependent Model Based on Level Set Motion for Nonlinear Deblurring and Noise Removal. UCLA CAM Report **99-55** (1999).
9. Malladi, R., Sethian, J.A., and Vemuri, B.C.: A Topology Independent Shape Modeling Scheme. Proc. SPIE Conf. on Geometric Methods in Computer Vision II, San Diego, **2031** (1993), 246-258.
10. Malladi, R., Sethian, J.A., and Vemuri, B.C.: Evolutionary Fronts for Topology-Independent Shape Modeling and Recovery. Proceedings of the Third European Conference on Computer Vision, LNCS **800** (1994), Stockholm, Sweden, 3-13.
11. Malladi, R., Sethian, J.A., Vemuri, B.C.: Shape Modeling with Front Propagation: A Level Set Approach. IEEE Transactions on Pattern Analysis and Machine Intelligence, **17**, No. 2 (1995), 158-175.
12. Lee, M.S., Medioni, G: Inferred Descriptions in Terms of Curves, Regions and Junctions from Sparse, Noisy Binary Data. Proceedings of the IEEE International Symposium on Computer Vision, Coral Gable, Florida (1995), 73-78.
13. Mumford, D., Shah, J.: Optimal approximation by piecewise smooth functions and associated variational problems. Comm. Pure Appl. Math. **42** (1989) 577-685.
14. Osher, S., Sethian, J. A.: Fronts Propagating with Curvature-Dependent Speed: Algorithms Based on Hamilton-Jacobi Formulation. Journal of Computational Physics **79** (1988) 12-49.
15. Paragios, N., and Deriche, R.: Geodesic Active Regions for Texture Segmentation. INRIA Research Report **3440** (1998).
16. Paragios, N., and Deriche, R.: Geodesic Active Regions for Motion Estimation and Tracking. INRIA Research Report 3631 (1999).
17. Siddiqi, K., Lauzière, Y. B., Tannenbaum, A., and Zucker, S. W.: Area and Length Minimizing Flows for Shape Segmentation. IEEE Transactions on Image Processing (Special Issue) **7/3** (1998) 433-443.
18. Sussman, M., Smereka, P., Osher, S.: A Level Set Approach for Computing Solutions to Incompressible Two-Phase Flow. UCLA CAM Report **93-18** (1993).
19. Xu, C., Prince, J. L.: Snakes, Shapes and Gradient Vector Flow. IEEE Transactions on Image Processing (Special Issue) **7/3** (1998) 359-369.

20. Zhao, H. K., Chan, T., Merriman, B., Osher, S.: A Variational Level Set Approach to Multiphase Motion. Journal of Computational Physics **127** (1996) 179-195.
21. Zhao, H.-K., Osher, S., Merriman, B. and Kang, M.: Implicit, Nonparametric Shape Reconstruction from Unorganized Points Using A Variational Level Set Method. UCLA CAM Report **98-7** (1998) (revised February 1999).