Active Contours Without Edges

Tony F. Chan, *Member, IEEE,* and Luminita A. Vese

***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. We minimize an energy which can be seen as a particular case of the 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. We will give a numerical algorithm using finite differences. 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. Also, the initial curve can be anywhere in the image, and interior contours are automatically detected.**

***Index Terms*—Active contours, curvature, energy minimization, finite differences, level sets, partial differential equations, segmentation.**

I. 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 and has to stop on the boundary of the object.

Let Ω be a bounded open subset of $\mathbb{R}^2$, with $\partial\Omega$ its boundary. Let u_0: $\overline{\Omega} \to \mathbb{R}$ be a given image, and $C(s)$: $[0, 1] \to \mathbb{R}^2$ be a parameterized curve.

In the classical snakes and active contour models (see [9], [3], [13], [4]), an edge-detector is used, depending on the gradient of the image u_0, to stop the evolving curve on the boundary of the desired object. We briefly recall these models next.

The snake model [9] is: $\inf_C J_1(C)$, where

$$J_1(C) = \alpha \int_0^1 |C'(s)|^2\, ds + \beta \int_0^1 |C''(s)|\, ds - \lambda \int_0^1 |\nabla u_0(C(s))|^2\, ds. \quad (1)$$

Here, α, β and λ are positive parameters. The first two terms control the smoothness of the contour (the internal energy), while the third term attracts the contour toward the object in the image (the external energy). Observe that, by minimizing the energy (1), we are trying to locate the curve at the points of maxima $|\nabla u_0|$, acting as an edge-detector, while keeping a smoothness in the curve (object boundary).

A general edge-detector can be defined by a positive and decreasing function g, depending on the gradient of the image u_0, such that

$$\lim_{z\to\infty} g(z) = 0.$$

For instance

$$g(|\nabla u_0(x, y)|) = \frac{1}{1 + |\nabla G_\sigma(x, y) * u_0(x, y)|^p}, \quad p \geq 1$$

where $G_\sigma * u_0$, a smoother version of u_0, is the convolution of the image u_0 with the Gaussian $G_\sigma(x, y) = \sigma^{-1/2} e^{-|x^2+y^2|/4\sigma}$. The function $g(|\nabla u_0|)$ is positive in homogeneous regions, and zero at the edges.

In problems of curve evolution, the level set method and in particular the motion by mean curvature of Osher and Sethian [19] have been used extensively, because it allows for cusps, corners, and automatic topological changes. Moreover, the discretization of the problem is made on a fixed rectangular grid. The curve C is represented implicitly via a Lipschitz function ϕ, by $C = \{(x, y) | \phi(x, y) = 0\}$, and the evolution of the curve is given by the zero-level curve at time t of the function $\phi(t, x, y)$. Evolving the curve C in normal direction with speed F amounts to solve the differential equation [19]

$$\frac{\partial \phi}{\partial t} = |\nabla \phi| F,\ \phi(0, x, y) = \phi_0(x, y)$$

where the set $\{(x, y) | \phi_0(x, y) = 0\}$ defines the initial contour. A particular case is the motion by mean curvature, when $F = \mathrm{div}(\nabla\phi(x, y)/|\nabla\phi(x, y)|)$ is the curvature of the level-curve of ϕ passing through (x, y). The equation becomes

$$\begin{cases} \dfrac{\partial \phi}{\partial t} = |\nabla \phi| \mathrm{div}\left(\dfrac{\nabla\phi}{|\nabla\phi|}\right), & t \in (0, \infty),\ x \in \mathbb{R}^2 \\ \phi(0, x, y) = \phi_0(x, y), & x \in \mathbb{R}^2. \end{cases}$$

A geometric active contour model based on the mean curvature motion is given by the following evolution equation [3]:

$$\begin{cases} \dfrac{\partial \phi}{\partial t} = g(|\nabla u_0|)|\nabla\phi| \left(\mathrm{div}\left(\dfrac{\nabla\phi}{|\nabla\phi|}\right) + \nu\right), \\ \text{in } (0, \infty) \times \mathbb{R}^2 \\ \phi(0, x, y) = \phi_0(x, y) \text{ in } \mathbb{R}^2 \end{cases} \quad (2)$$

where

$g(\lvert\nabla u_0\rvert)$	edge-function with $p = 2$;
$\nu \geq 0$	is constant;
ϕ_0	initial level set function.

Manuscript received June 17, 1999; revised September 27, 2000. This work was supported in part by ONR under Contract N00014-96-1-0277 and NSF Contract DMS-9626755. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Robert J. Schalkoff.

The authors are with the Mathematics Department, University of California, Los Angeles, CA 90095-1555 USA (e-mail: chan@math.ucla.edu; lvese@math.ucla.edu).

Publisher Item Identifier S 1057-7149(01)00819-3.

Its zero level curve moves in the normal direction with speed $g(|\nabla u_0|)\,(curv(\phi)(x, y)+\mu)$ and therefore stops on the desired boundary, where g vanishes. The constant ν is a correction term chosen so that the quantity $(\mathrm{div}(\nabla\phi(x, y)/|\nabla\phi(x, y)|) + \nu)$ remains always positive. This constant may be interpreted as a force pushing the curve toward the object, when the curvature becomes null or negative. Also, $\nu > 0$ is a constraint on the area inside the curve, increasing the propagation speed.

Two other active contour models based on level sets were proposed in [13], again using the image gradient to stop the curve. The first one is

$$\begin{cases} \dfrac{\partial\phi}{\partial t} = |\nabla\phi| \left(-\nu + \dfrac{\nu}{(M_1 - M_2)}(|\nabla G_\sigma * u_0| - M_2)\right), \\ \phi(0, x, y) = \phi_0(x, y) \text{ in } [0, \infty) \times \mathbb{R}^2 \end{cases}$$

where ν is a constant, and M_1 and M_2 are the maximum and minimum values of the magnitude of the image gradient $|\nabla G_\sigma * u_0|$. Again, the speed of the evolving curve becomes zero on the points with highest gradients, and therefore the curve stops on the desired boundary, defined by strong gradients. The second model [13] is similar to the geometric model [3], with $p = 1$. Other related works are [14] and [15].

The geodesic model [4] is

$$\inf_C J_2(C) = 2\int_0^1 |C'(s)| \cdot g(|\nabla u_0(C(s))|)\, ds. \qquad (3)$$

This is a problem of geodesic computation in a Riemannian space, according to a metric induced by the image u_0. Solving the minimization problem (3) consists in finding the path of minimal new length in that metric. A minimizer C will be obtained when $g(|\nabla u_0(C(s)|)$ vanishes, i.e., when the curve C is on the boundary of the object. The geodesic active contour model (3) from [4] also has a level set formulation

$$\begin{cases} \dfrac{\partial\phi}{\partial t} = |\nabla\phi| \left(\mathrm{div}\left(g(|\nabla u_0|)\dfrac{\nabla\phi}{|\nabla\phi|}\right) + \nu g(|\nabla u_0|)\right), \\ \text{in } [0, \infty[\times\mathbb{R}^2 \\ \phi(0, x, y) = \phi_0(x, y) \text{ in } \mathbb{R}^2. \end{cases} \qquad (4)$$

Because all these classical snakes and active contour models rely on the edge-function g, depending on the image gradient $|\nabla u_0|$, to stop the curve evolution, these models can detect only objects with edges defined by gradient. 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, especially for the models in [3], [13]–[15]. If the image u_0 is very 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 [18]. 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 [8]). In addition, our model has a level set formulation, interior contours are automatically detected, and the initial curve can be anywhere in the image.

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. Also, we formulate the model in terms of level set functions and compute the associated Euler–Lagrange equations. In Section III we present an iterative algorithm for solving the problem and its discretization. In Section IV we validate our model by various numerical results on synthetic and real images, showing the advantages of our model described before, and we end the paper by a brief concluding section.

Other related works are [29], [10], [26], and [24] on active contours and segmentation, [28] and [11] on shape reconstruction from unorganized points, and finally the recent works [20] and [21], where a probability based geodesic active region model combined with classical gradient based active contour techniques is proposed.

II. Description of the Model

Let us define the evolving curve C in Ω, as the boundary of an open subset ω of Ω (i.e. $\omega \subset \Omega$, and $C = \partial\omega$). In what follows, $inside(C)$ denotes the region ω, and $outside(C)$ denotes the region $\Omega \setminus \overline{\omega}$.

Our method 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. Let denote its boundary by C_0. Then we have $u_0 \approx u_0^i$ inside the object [or *inside* (C_0)], and $u_0 \approx u_0^o$ outside the object [or *outside* (C_0)]. Now let us consider the following "fitting" term:

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

where C is any other variable curve, and the constants c_1, c_2, depending on C, are the averages of u_0 inside C and respectively outside C. In this simple case, it is obvious that C_0, the boundary of the object, is the minimizer of the fitting term

$$\inf_C\{F_1(C) + F_2(C)\} \approx 0 \approx F_1(C_0) + F_2(C_0).$$

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 $F_2(C) > 0$. If the curve C is both inside and outside the object, then $F_1(C) > 0$ and $F_2(C) > 0$. Finally, the fitting energy is minimized if $C = C_0$, i.e., if the curve C is on the boundary of the object. These basic remarks are illustrated in Fig. 1.

In our active contour model we will minimize the above fitting term and we will add some regularizing terms, like the

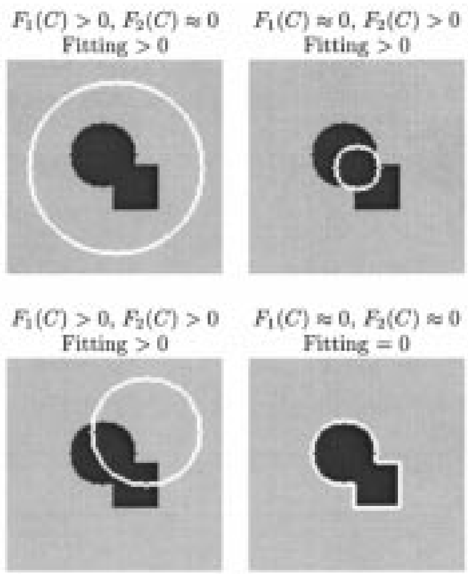


Fig. 1. Consider all possible cases in the position of the curve. The fitting term is minimized only in the case when the curve is on the boundary of the object.

length of the curve C, and (or) the area of the region inside C. Therefore, we introduce the energy functional $F(c_1, c_2, C)$, defined by

$$\begin{aligned} F(c_1, c_2, C) = &\mu \cdot \text{Length}(C) + \nu \cdot \text{Area}(inside(C)) \\ &+ \lambda_1 \int_{inside(C)} |u_0(x, y) - c_1|^2 \, dx \, dy \\ &+ \lambda_2 \int_{outside(C)} |u_0(x, y) - c_2|^2 \, dx \, dy, \end{aligned}$$

where $\mu \geq 0, \nu \geq 0, \lambda_1, \lambda_2 > 0$ are fixed parameters. In almost all our numerical calculations (see further), we fix $\lambda_1 = \lambda_2 = 1$ and $\nu = 0$.

Therefore, we consider the minimization problem:

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

Remark 1: In our model, the term Length(C) could be re-written in a more general way as $(\text{Length}(C))^p$, with $p \geq 1$. If we consider the case of an arbitrary dimension $N > 1$ (i.e., $\Omega \subset \mathbb{R}^N$), then p can have the following values: $p = 1$ for all N, or $p = N/(N-1)$. For the last expression, we use the isoperimetric inequality [7], which says in some sense that $(\text{Length}(C))^{N/(N-1)}$ is "comparable" with Area$(inside(C))$:

$$\text{Area}(inside(C)) \leq c \cdot (\text{Length}(C))^{N/(N-1)}$$

where c is a constant depending only on N.

A. Relation with the Mumford–Shah Functional

The Mumford–Shah functional for segmentation is [18]

$$\begin{aligned} F^{\text{MS}}(u, C) = &\mu \cdot \text{Length}(C) \\ &+ \lambda \int_{\Omega} |u_0(x, y) - u(x, y)|^2 \, dx \, dy \\ &+ \int_{\Omega \setminus C} |\nabla u(x, y)|^2 \, dx \, dy \end{aligned}$$

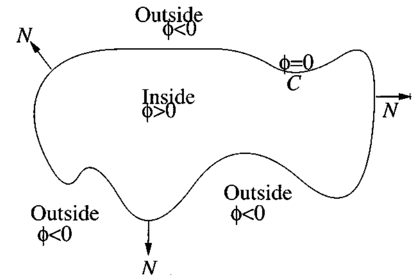


Fig. 2. Curve $C = \{(x, y): \phi(x, y) =\}$ propagating in normal direction.

where $u_0: \overline{\Omega} \to \mathbb{R}$ is a given image, μ and λ 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 is simply the restriction of F^{MS} to piecewise constant functions u, i.e., $u = constant\ c_i$ on each connected component R_i of $\Omega \backslash C$. Therefore, as it was also pointed out by D. Mumford and J. Shah [18], $c_i = average\ (u_0)$ on each connected component R_i. The reduced case is called the minimal partition problem.

Our active contour model with $\nu = 0$ and $\lambda_1 = \lambda_2 = \lambda$ 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

$$u = \begin{cases} \text{average}(u_0) \text{ inside } C \\ \text{average}(u_0) \text{ outside } C \end{cases} \tag{5}$$

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 [19]. This is presented in the next section.

B. Level Set Formulation of the Model

In the level set method [19], $C \subset \Omega$ is represented by the zero level set of a Lipschitz function $\phi: \Omega \to \mathbb{R}$, such that

$$\begin{cases} C = \partial\omega = \{(x, y) \in \Omega: \phi(x, y) = 0\}, \\ inside(C) = \omega = \{(x, y) \in \Omega: \phi(x, y) > 0\} \\ outside(C) = \Omega \setminus \overline{\omega} = \{(x, y) \in \Omega: \phi(x, y) < 0\}. \end{cases}$$

Recall that $\omega \subset \Omega$ is open, and $C = \partial\omega$. We illustrate in Fig. 2 the above assumptions and notations on the level set function ϕ, defining the evolving curve C. For more details, we refer the reader to [19].

For the level set formulation of our variational active contour model, we replace the unknown variable C by the unknown variable ϕ, and we follow [27].

Using the Heaviside function H, and the one-dimensional Dirac measure δ_0, and defined, respectively, by

$$H(z) = \begin{cases} 1, & \text{if } z \geq 0 \\ 0, & \text{if } z < 0, \end{cases} \qquad \delta_0(z) = \frac{d}{dz} H(z)$$

(in the sense of distributions), we express the terms in the energy F in the following way (see also [7]):

$$\begin{aligned}\text{Length}\{\phi=0\} &= \int_\Omega |\nabla H(\phi(x,y))|\,dx\,dy\\ &= \int_\Omega \delta_0(\phi(x,y))|\nabla\phi(x,y)|\,dx\,dy,\\ \text{Area}\{\phi\geq 0\} &= \int_\Omega H(\phi(x,y))\,dx\,dy,\end{aligned}$$

and

$$\begin{aligned}&\int_{\phi>0}|u_0(x,y)-c_1|^2\,dx\,dy\\ &\quad=\int_\Omega |u_0(x,y)-c_1|^2\,H(\phi(x,y))\,dx\,dy,\\ &\int_{\phi<0}|u_0(x,y)-c_2|^2\,dx\,dy\\ &\quad=\int_\Omega |u_0(x,y)-c_2|^2\,(1-H(\phi(x,y)))\,dx\,dy.\end{aligned}$$

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

$$\begin{aligned}&F(c_1,c_2,\phi)\\ &\quad=\mu\int_\Omega \delta(\phi(x,y))|\nabla\phi(x,y)|\,dx\,dy\\ &\qquad+\nu\int_\Omega H(\phi(x,y))\,dx\,dy\\ &\qquad+\lambda_1\int_\Omega |u_0(x,y)-c_1|^2\,H(\phi(x,y))\,dx\,dy\\ &\qquad+\lambda_2\int_\Omega |u_0(x,y)-c_2|^2\,(1-H(\phi(x,y)))\,dx\,dy.\end{aligned}$$

We note that, u as defined in (5), solution of our model as a particular case of the Mumford–Shah minimal partition problem, can simply be written using the level set formulation as

$$u(x,y)=c_1H(\phi(x,y))+c_2(1-H(\phi(x,y))),\ (x,y)\in\overline{\Omega}.$$

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

$$c_1(\phi)=\frac{\displaystyle\int_\Omega u_0(x,y)H(\phi(x,y))\,dx\,dy}{\displaystyle\int_\Omega H(\phi(x,y))\,dx\,dy} \tag{6}$$

if $\int_\Omega H(\phi(x,y))\,dx\,dy > 0$ (i.e. if the curve has a nonempty interior in Ω), and

$$c_2(\phi)=\frac{\displaystyle\int_\Omega u_0(x,y)\,(1-H(\phi(x,y)))\,dx\,dy}{\displaystyle\int_\Omega (1-H(\phi(x,y)))\,dx\,dy} \tag{7}$$

if $\int_\Omega (1-H(\phi(x,y)))\,dx\,dy > 0$ (i.e. if the curve has a nonempty exterior in Ω). For the corresponding "degenerate" cases, there are no constrains on the values of c_1 and c_2. Then, c_1 and c_2 are in fact given by

$$\begin{cases} c_1(\phi)=\text{average}(u_0)\text{ in }\{\phi\geq 0\}\\ c_2(\phi)=\text{average}(u_0)\text{ in }\{\phi<0\}.\end{cases}$$

Remark 2: By the previous formulas, we can see that the energy can be written only function of $H(\phi)$, which is the characteristic function of the set ω. Let us denote it by χ_ω. Then we can rewrite the energy in the new form $\mathcal{F}$

$$\begin{aligned}\mathcal{F}(\chi_\omega)=\mu&\int_\Omega|\nabla\chi_\omega(x,y)|\,dx\,dy+\nu\int_\Omega\chi_\omega(x,y)\,dx\,dy\\ &+\lambda_1\int_\Omega (u_0(x,y)-c_1(\chi_\omega(x,y)))^2\,\chi_\omega(x,y)\,dx\,dy\\ &+\lambda_2\int_\Omega (u_0(x,y)\\ &-c_2(\chi_\omega(x,y)))^2\,(1-\chi_\omega(x,y))\,dx\,dy.\end{aligned}$$

Therefore, we can consider the new minimization problem

$$\inf_{\chi_\omega}\mathcal{F}(\chi_\omega),\ \chi_\omega(x,y)\in\{0,1\}\ \mathcal{L}-a.e. \tag{8}$$

among characteristic functions of sets with finite perimeter in Ω. Here, $\mathcal{L}-a.e.$ means almost everywhere with respect to the Lebesgue measure.

We expect, of course, to have existence of minimizers of the energy $F(c_1, c_2, C)$, due to several general results: our model is a particular case of the minimal partition problem, for which the existence has been proved in [18] (assuming that u_0 is continuous on $\overline{\Omega}$), and also in [16] and [17], for more general data u_0. Also, the existence for the general Mumford–Shah segmentation problem has been proved in [5]. On the other hand, it can be easily shown, by the lower-semicontinuity of the total variation $\int_\Omega|\nabla\chi_\omega|\,dx\,dy$ and classical arguments of calculus of variations, that our minimization problem (8) has minimizers (this can be an alternative proof of the existence). In this paper, the level set function ϕ is used only to represent the curve and it has many numerical advantages, but the problem could also be formulated and solved only in terms of characteristic functions.

In order to compute the associated Euler–Lagrange equation for the unknown function ϕ, we consider slightly regularized versions of the functions H and δ_0, denoted here by H_ε and δ_ε, as $\varepsilon\to 0$. Let H_ε beany $C^2\,(\overline{\Omega})$ regularization of H, and $\delta_\varepsilon = H'_\varepsilon$. We will give further examples of such approximations. Let us denote by F_ε the associated regularized functional, defined by

$$\begin{aligned}&F_\varepsilon(c_1,c_2,\phi)\\ &\quad=\mu\int_\Omega\delta_\varepsilon(\phi(x,y))|\nabla\phi(x,y)|\,dx\,dy\\ &\qquad+\nu\int_\Omega H_\varepsilon(\phi(x,y))\,dx\,dy\\ &\qquad+\lambda_1\int_\Omega |u_0(x,y)-c_1|^2\,H_\varepsilon(\phi(x,y))\,dx\,dy\\ &\qquad+\lambda_2\int_\Omega |u_0(x,y)-c_2|^2\,(1-H_\varepsilon(\phi(x,y)))\,dx\,dy.\end{aligned}$$

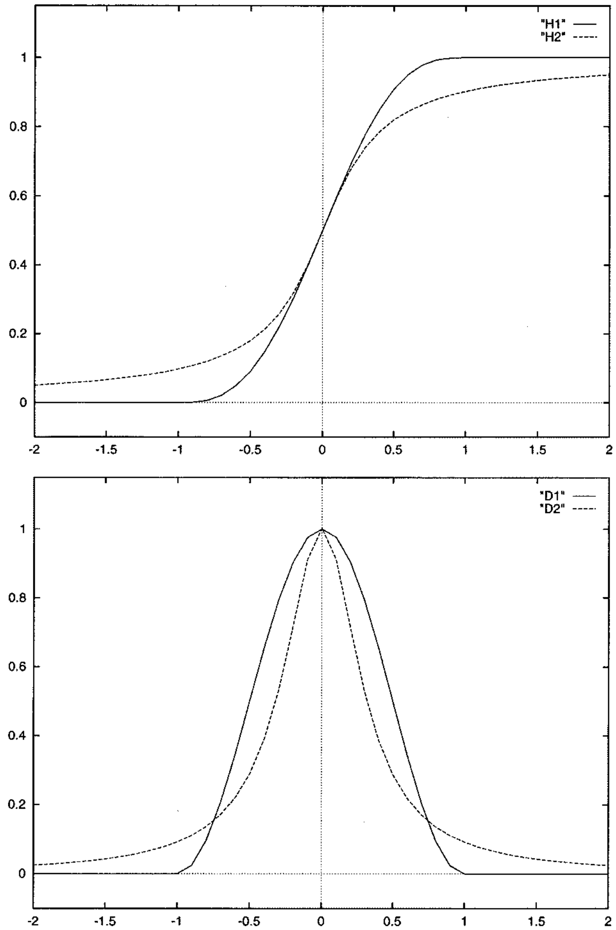


Fig. 3. Two different regularizations of the (top) heaviside function and (bottom) delta function δ_0.

Keeping c_1 and c_2 fixed, and minimizing F_ε with respect to ϕ, we deduce the associated Euler–Lagrange equation for ϕ. Parameterizing the descent direction by an artificial time $t \geq 0$, the equation in $\phi(t, x, y)$ (with $\phi(0, x, y) = \phi_0(x, y)$ defining the initial contour) is

$$\begin{aligned} \frac{\partial \phi}{\partial t} &= \delta_\varepsilon(\phi) \left[\mu \operatorname{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) - \nu - \lambda_1 (u_0 - c_1)^2 \right. \\ &\left. \qquad + \lambda_2 (u_0 - c_2)^2 \right] = 0 \text{ in } (0, \infty) \times \Omega, \\ \phi(0, x, y) &= \phi_0(x, y) \text{ in } \Omega, \\ \frac{\delta_\varepsilon(\phi)}{|\nabla \phi|} \frac{\partial \phi}{\partial \vec{n}} &= 0 \text{ on } \partial\Omega \end{aligned} \tag{9}$$

where $\vec{n}$ denotes the exterior normal to the boundary $\partial\Omega$, and $\partial\phi/\partial\vec{n}$ denotes the normal derivative of ϕ at the boundary.

III. Numerical Approximation of the Model

First possible regularization of H by $C^2(\overline{\Omega})$ functions, as proposed in [27], is

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

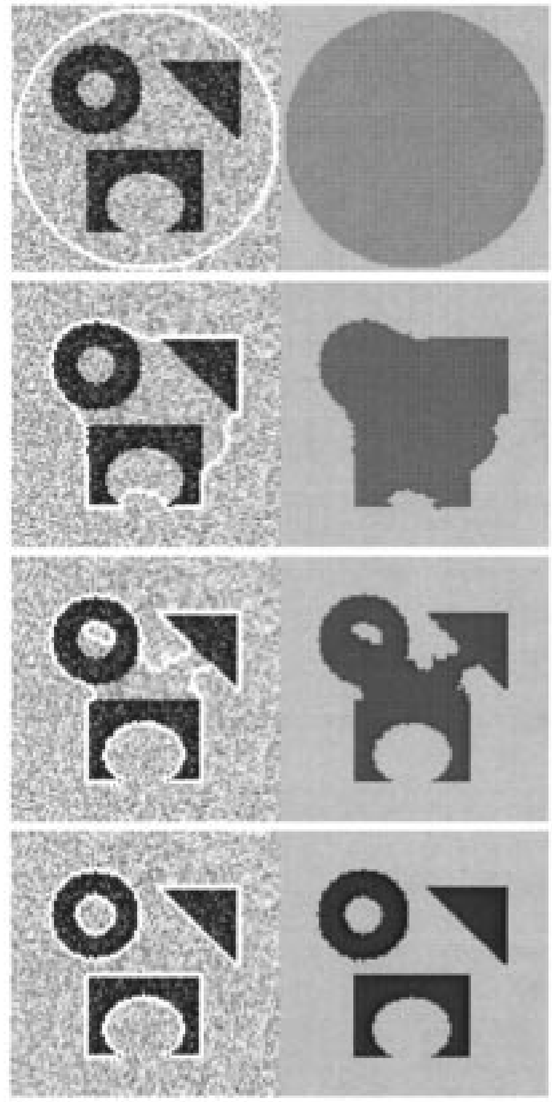


Fig. 4. Detection of different objects from a noisy image, with various shapes and with an interior contour. Left: u_0 and the contour. Right: the piecewise-constant approximation of u_0. Size $=$ 100×100, $\phi_0(x, y) = -\sqrt{(x - 50.5)^2 + (y - 50.5)^2} + 48.5$, $\mu = 0.1 \cdot 255^2$, no reinitialization, cpu $= 4.60$ s.

In this paper, we introduce and use in our experiments the following $C^\infty(\overline{\Omega})$ regularization of H

$$H_{2,\varepsilon}(z) = \frac{1}{2} \left(1 + \frac{2}{\pi} \arctan\left(\frac{z}{\varepsilon} \right) \right).$$

These distinct approximations and regularizations of the functions H and δ_0 (taking $\delta_\varepsilon = H'_\varepsilon$) are presented in Fig. 3. As $\varepsilon \to 0$, both approximations converge to H and δ_0. A difference is that $\delta_{1,\varepsilon}$ has a small support, the interval $[-\varepsilon, \varepsilon]$, while $\delta_{2,\varepsilon}$ is different of zero everywhere. Because our energy is nonconvex (allowing therefore many local minima), the solution may depend on the initial curve. With $H_{1,\varepsilon}$ and $\delta_{1,\varepsilon}$, the algorithm sometimes computes a local minimizer of the energy,

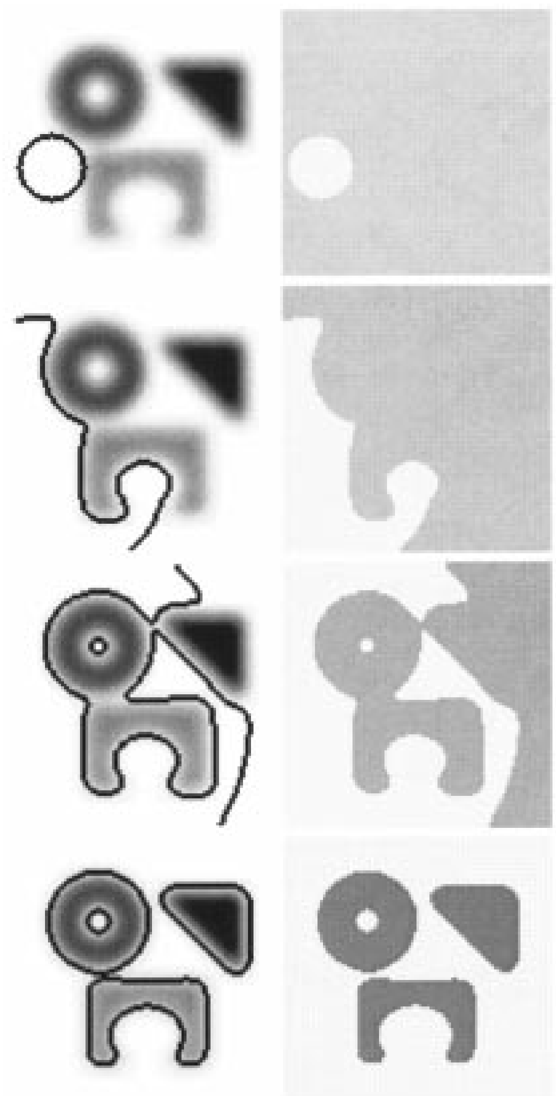


Fig. 5. Detection of three blurred objects of distinct intensities. Size $= 100 \times 100$, $\phi_0(x, y) = -\sqrt{(x-15)^2+(y-60)^2} + 12$, $\mu = 0.01 \cdot 255^2$, no reinitialization, cpu $= 48.67$ s.

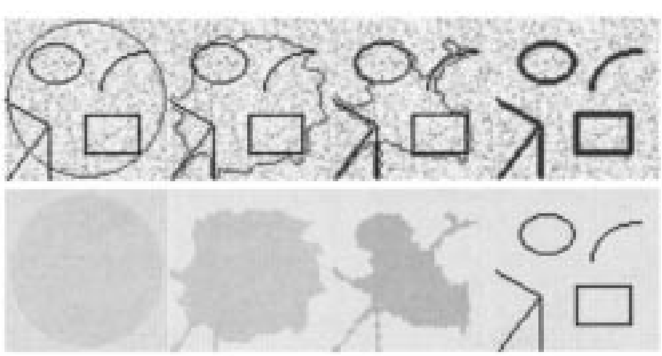


Fig. 6. Detection of lines and curves not necessarily closed. Size $= 64 \times 64$, $\phi_0(x, y) = -\sqrt{(x-32.5)^2+(y-32.5)^2} + 30$, $\mu = 0.02 \cdot 255^2$, no reinitialization, cpu $= 2.88$ s.

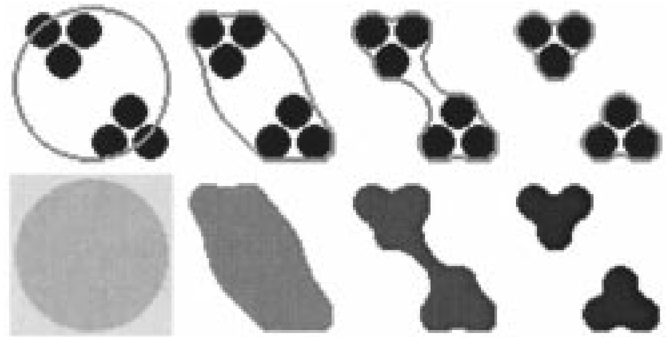


Fig. 7. Grouping based on Kanizsa's "proximity rule." Size: 64 × 64, $\phi_0(x, y) = -\sqrt{(x-32.5)^2+(y-32.5)^2} + 30$, $\mu = 2 \cdot 255^2$, no reinitialization, cpu $= 5.76$ s.

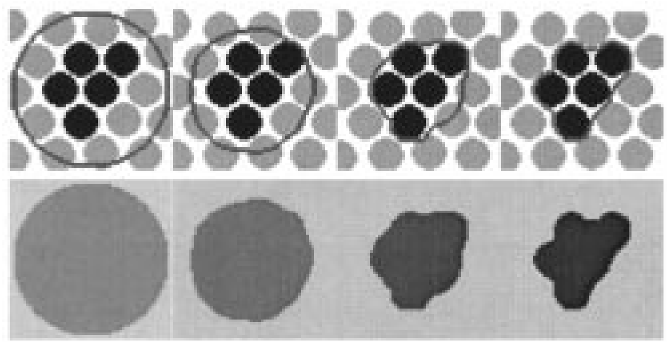


Fig. 8. Grouping based on chromatic identity. Size: 64 × 64, $\phi_0(x, y) = -\sqrt{(x-32.5)^2+(y-32.5)^2} + 30.5$, $\mu = 2 \cdot 255^2$, no reinitialization, cpu $= 5.76$ s.

while with $H_{2,\varepsilon}$ and $\delta_{2,\varepsilon}$, the algorithm has the tendency to compute a global minimizer. One of the reasons is that the Euler–Lagrange equation for ϕ acts only locally, on a few level curves around $\{\phi(x, y) = 0\}$ using $H_{1,\varepsilon}$ and $\delta_{1,\varepsilon}$; but using $H_{2,\varepsilon}$ and $\delta_{2,\varepsilon}$, the equation acts on all level curves. In this way, in practice, we can obtain a global minimizer, independently of the position of the initial curve; moreover, this allows to automatically detect interior contours (see Section IV). We mention that, in order to extend the evolution to all level sets of ϕ, another possibility is to replace $\delta_0(\phi)$ by $|\nabla\phi|$ (see [27]). In our paper, we work with $\delta_0(\phi)$, to remain close to the initial minimization problem. The problem of extending the evolution to all level sets of ϕ was solved here using the approximation $\delta_{2,\varepsilon}$ of δ_0, which is different of zero everywhere.

To discretize the equation in ϕ, we use a finite differences implicit scheme. We recall first the usual notations: let h be the space step, Δt be the time step, and $(x_i, y_j) = (ih, jh)$ be the grid points, for $1 \le i, j \le M$. Let $\phi^n_{i,j} = \phi(n\Delta t, x_i, y_j)$ be an approximation of $\phi(t, x, y)$, with $n \ge 0$, $\phi^0 = \phi_0$. The finite differences are

$$\begin{aligned} \Delta^x_-\phi_{i,j} &= \phi_{i,j} - \phi_{i-1,j}, \quad \Delta^x_+\phi_{i,j} = \phi_{i+1,j} - \phi_{i,j}, \\ \Delta^y_-\phi_{i,j} &= \phi_{i,j} - \phi_{i,j-1}, \quad \Delta^y_+\phi_{i,j} = \phi_{i,j+1} - \phi_{i,j}. \end{aligned}$$

The algorithm is as follows (we essentially adopt the method from [23] for the discretization of the divergence operator and the iterative algorithm from [1]): knowing ϕ^n, we first compute $c_1(\phi^n)$ and $c_2(\phi^n)$ using (6) and (7), respectively. Then, we

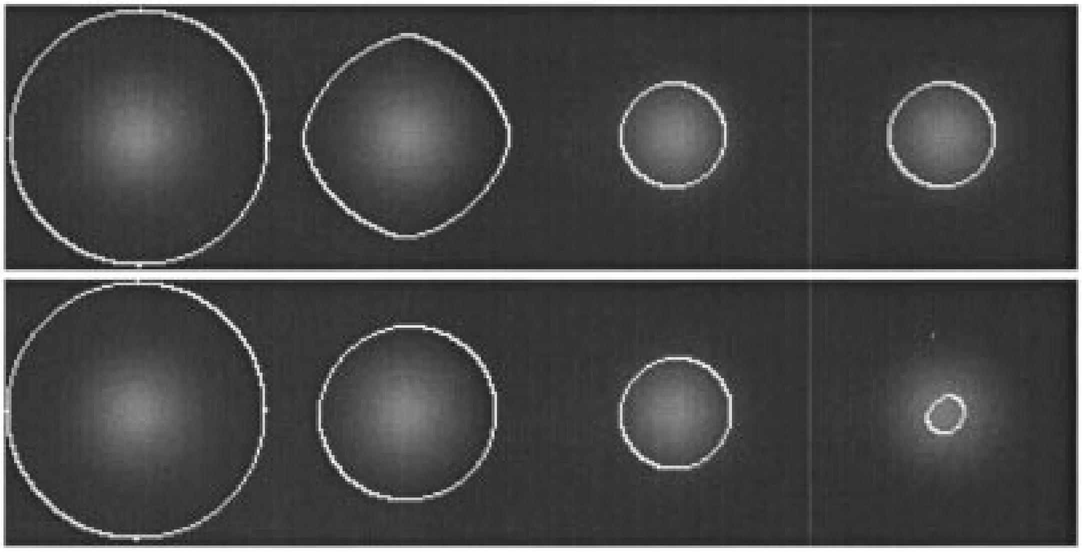


Fig. 9. Object with smooth contour. Top: results using our model without edge-function. Bottom: results using the classical model (2) with edge-function.

compute ϕ^{n+1} by the following discretization and linearization of (9) in ϕ

$$\begin{aligned}\frac{\phi_{i,j}^{n+1}-\phi_{i,j}^{n}}{\Delta t} &= \delta_h(\phi_{i,j}^n)\Bigg[\frac{\mu}{h^2}\Delta_-^x \\ &\cdot\left(\frac{\Delta_+^x\phi_{i,j}^{n+1}}{\sqrt{(\Delta_+^x\phi_{i,j}^n)^2/(h^2)+(\phi_{i,j+1}^n-\phi_{i,j-1}^n)^2/(2h)^2}}\right) \\ &+\frac{\mu}{h^2}\Delta_-^y \\ &\cdot\left(\frac{\Delta_+^y\phi_{i,j}^{n+1}}{\sqrt{(\phi_{i+1,j}^n-\phi_{i-1,j}^n)^2/(2h)^2+(\Delta_+^y\phi_{i,j}^n)^2/(h^2)}}\right) \\ &-\nu-\lambda_1(u_{0,i,j}-c_1(\phi^n))^2+\lambda_2(u_{0,i,j}-c_2(\phi^n))^2\Bigg].\end{aligned}$$

This linear system is solved by an iterative method, and for more details, werefer the reader to [1].

When working with level sets and Dirac delta functions, a standard procedure is to reinitialize ϕ to the signed distance function to its zero-level curve, as in [25] and [27]. This prevents the level set function to become too flat, or it can be seen as a rescaling and regularization. For our algorithm, the reinitialization is optional. On the other hand, it should not be too strong, because, as it was remarked by Fedkiw, it prevents interior contours from growing. Only for a few numerical results we have applied the reinitialization, solving the following evolution equation [25]:

$$\begin{cases}\psi_\tau = \operatorname{sign}(\phi(t))(1-|\nabla\psi|) \\ \psi(0,\cdot)=\phi(t,\cdot)\end{cases} \tag{10}$$

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 (10). The solution $\psi(t,\cdot)$ of (10) will have the same zero-level set as $\phi(t,\cdot)$ and away from this set, $|\nabla\psi|$ will converge to 1. To discretize the equation (10), we use the scheme proposed in [22] and [25].

Finally, the principal steps of the algorithm are:

- Initialize ϕ^0 by ϕ_0, $n=0$.
- Compute $c_1(\phi^n)$ and $c_2(\phi^n)$ by (6) and (7).
- Solve the PDE in ϕ from (9), to obtain ϕ^{n+1}.
- Reinitialize ϕ locally to the signed distance function to the curve (this step is optional).
- Check whether the solution is stationary. If not, $n=n+1$ and repeat.

We note that the use of a time-dependent PDE for ϕ is not crucial. The stationary problem obtained directly from the minimization problem could also be solved numerically, using a similar finite differences scheme.

IV. Experimental Results

We conclude this paper by presenting numerical results using our model on various synthetic and real images, with different types of contours and shapes. We show the active contour evolving in the original image u_0, and the associated piecewise-constant approximation of u_0 (given by the averages c_1 and c_2). In our numerical experiments, we generally choose the parameters as follows: $\lambda_1=\lambda_2=1$, $\nu=0$, $h=1$ (the step space), $\Delta t=0.1$ (the time step). We only use the approximations $H_{2,\varepsilon}$ and $\delta_{2,\varepsilon}$ of the Heaviside and Dirac delta functions ($\varepsilon=h=1$), in order to automatically detect interior contours, and to insure the computation of a global minimizer. Only the length parameter μ, which has a scaling role, is not the same in all experiments. If we have to detect all or as many objects as possible and of any size, then μ should be small. If we have to detect only larger objects (for example objects formed by grouping), and to not detect smaller objects (like

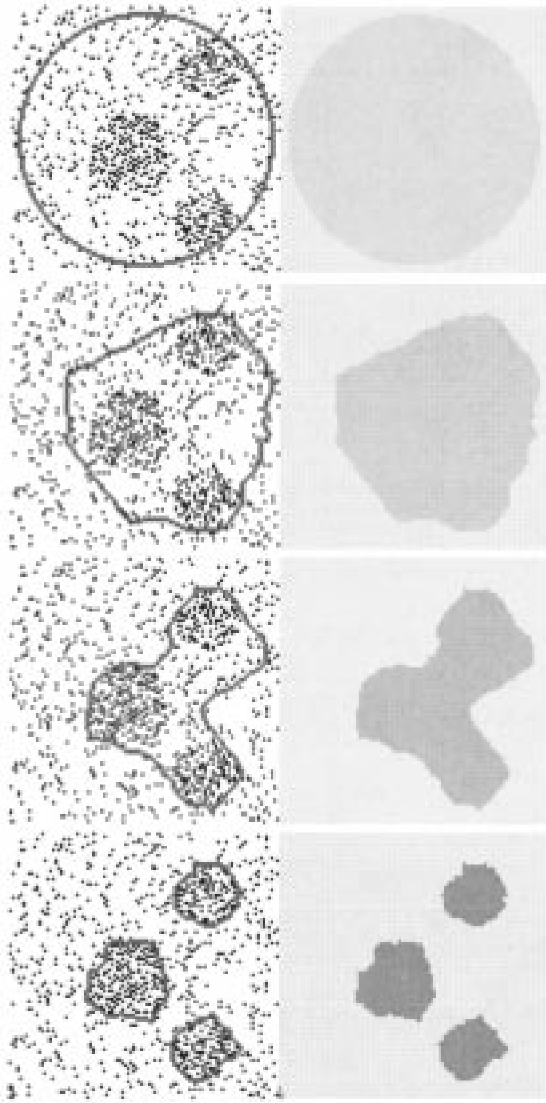


Fig. 10. Detection of a simulated minefield, with contour without gradient. Size $= 100 \times 100$, $\phi_0(x, y) = -\sqrt{(x-50.5)^2 + (y-50.5)^2} + 47$, $\mu = 0.2 \cdot 255^2$, no reinitialization, cpu $= 144.81$ s.

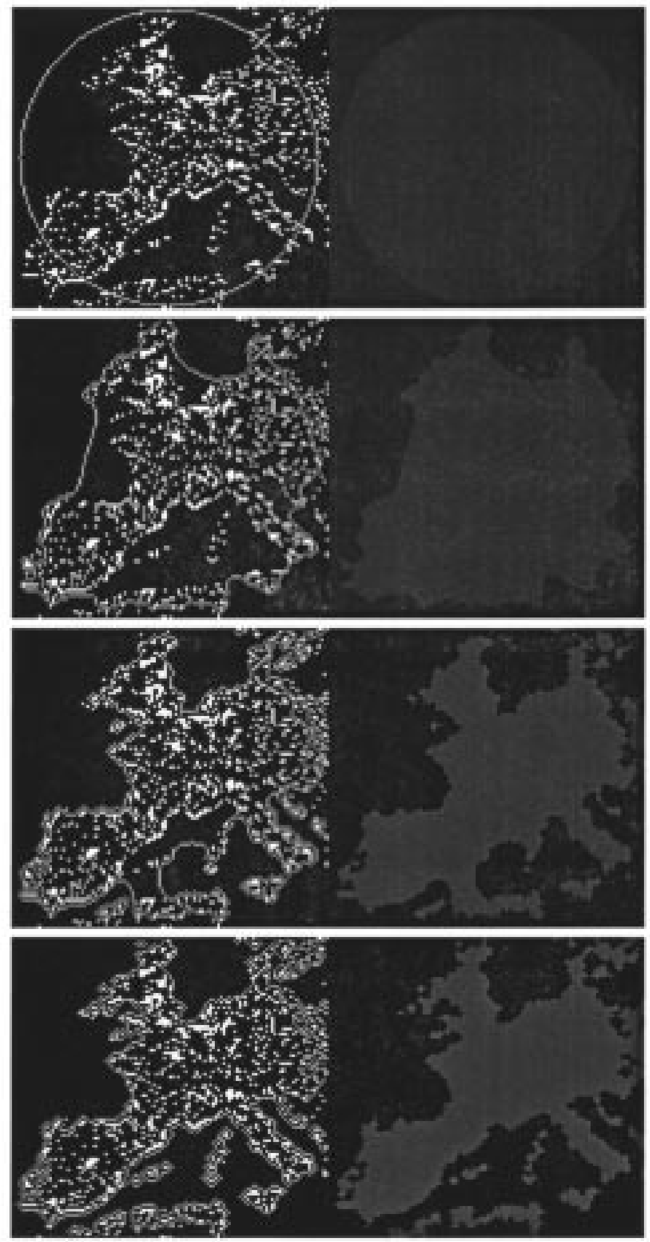


Fig. 11. Europe night-lights. Size $= 118 \times 113$, $\phi_0(x, y) = -\sqrt{(x-59.)^2 + (y-57.)^2} + 55$, $\mu = 0.05 \cdot 255^2$, five iterations of reinitialization, cpu $= 32.74$ s.

points, due to the noise), then μ has to be larger. We will give the exact value of μ each time, together with the initial level set function ϕ_0, and the cpu time, in seconds, of our calculations, performed on a 140 MHz Sun Ultra 1 with 256 MB of RAM.

In Fig. 4, we show how our model works on a noisy synthetic image, with various shapes and an interior contour, which is automatically detected, without considering a second initial curve. Due to the level set implementation, the model allows automatical change of topology.

In Fig. 5, we show that our model can detect different objects of different intensities, and with blurred boundaries. Again, the interior contour of the torus is automatically detected. This is also due to the fact that the velocity has a global dependence, and the curve is automatically attracted toward the objects. In this example we also show that the initial curve does not necessarily surround the objects.

In Fig. 6, we show how we can detect lines and curves (not necessarily closed) in a noisy image. The final level set function is zero on the curves and negative outside the curves.

In the next examples (Figs. 7 and 8) we consider images with "contours without gradient" or "cognitive contours" (see [8]). We also illustrate here the role of the length term as a scale parameter: if μ is small, then also smaller objects will be detected; if μ is larger, then only larger objects are detected, or objects formed by grouping. In Fig. 7, we show that our algorithm can detect objects defined by grouping according to Kanizsa's "proximity rule." In Fig. 8 we show how the grouping is based on the chromatic resemblance or identity, among objects of the same shape.

We next consider an image with very smooth contours. In Fig. 9 top, we show results obtained using our model, while in Fig. 9 bottom, we show the results obtained with a classical active contour model based on the edge-function $g(|\nabla u_0|)$ [here

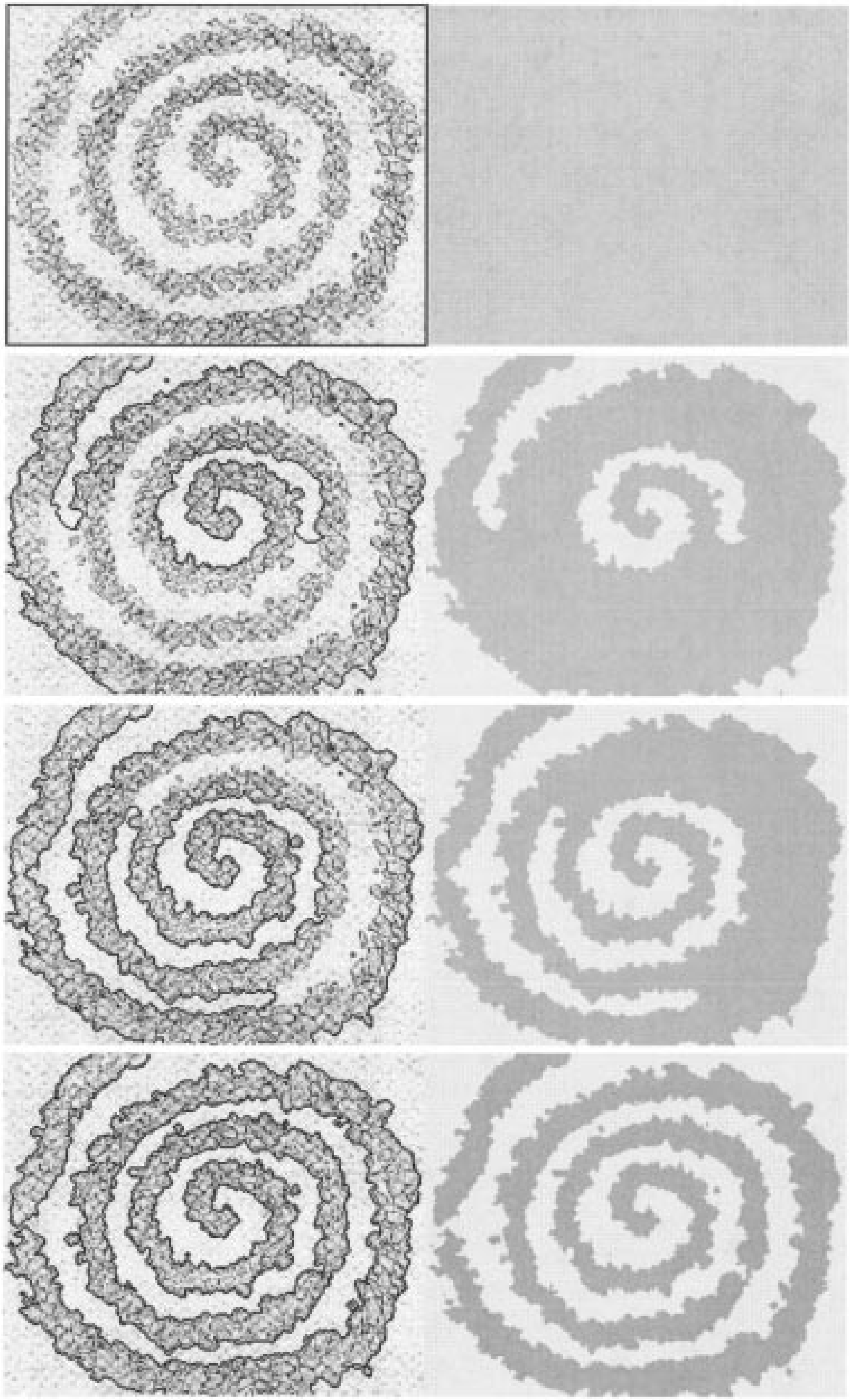


Fig. 12. Spiral from an art picture. Size $= 234 \times 191$, $\mu = 0.000\,003\,3 \cdot 255^2$, five iterations of reinitialization, cpu $= 108.85$ s.

the geometric model (2)], by which the curve cannot detect the smooth boundary.

In Fig. 10, we validate our model on a very different problem: to detect features in spatial point processes in the presence of

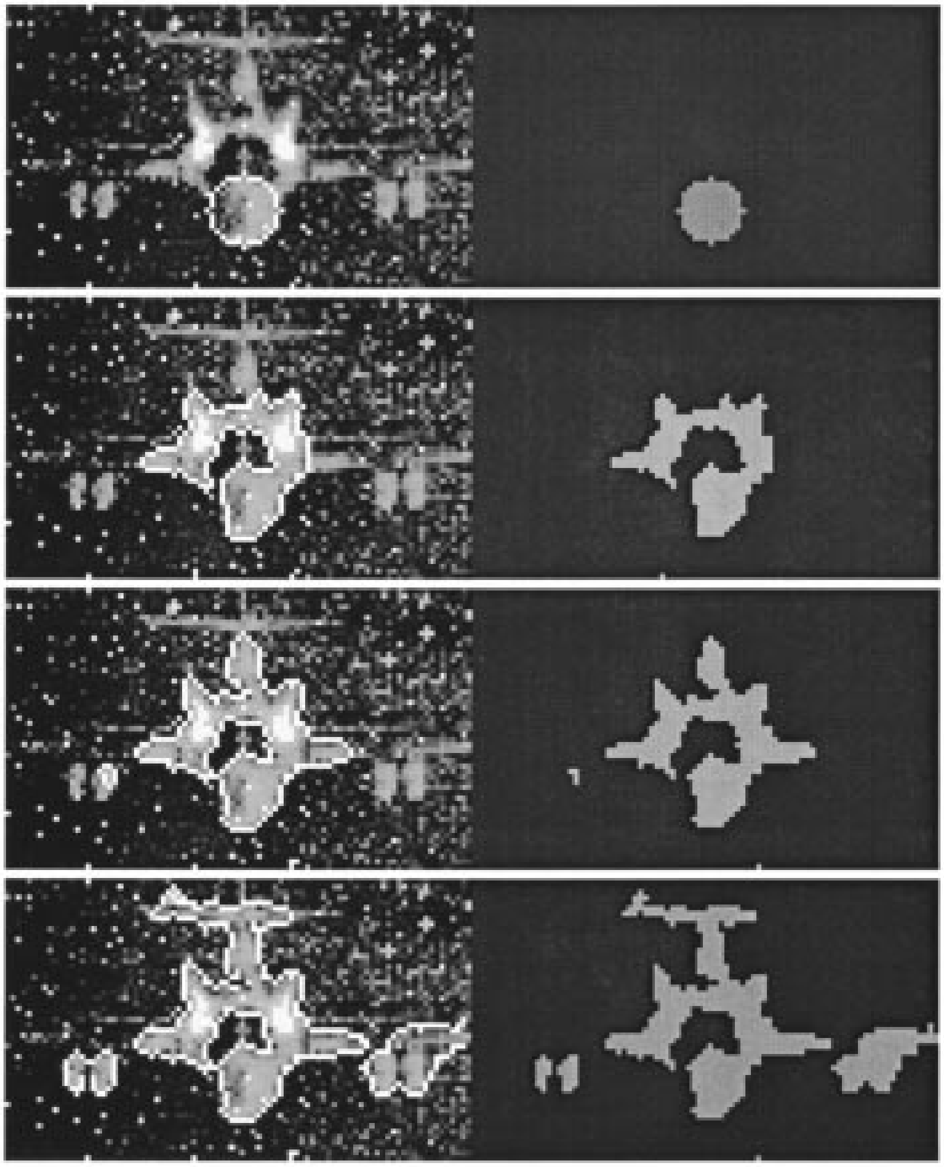


Fig. 13. Detection of the contours of a plane from a noisy image. Size $= 87 \times 53$, $\Delta t = 0.01$, $\mu = 0.17 \cdot 255^2$, $\phi_0(x,\, y) = -\sqrt{(x-45)^2 + (y-39)^2} + 6$, one iteration of reinitialization, cpu $= 2.87$ s.

substantial cluster. One application is the detection of minefields using reconnaissance aircraft images that identify many objects that are not mines. These problems are usually solved using statistical methods (see [6] and [2]). By this application, we show again how 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. A similar application is presented in Fig. 11, where the white points are Europe night-lights.

We also show examples on real noisy images, with different types of contours or shapes, illustrating all the advantages of our model: the ability of detecting smooth boundaries, scale adaptivity, automatic change of topology, and robustness with respect to noise.

In Fig. 12, we consider an art picture from the *Los Angeles Times* by Brian Forrest. Here, $p = 2$ from *Remark 1*, Section II [we have $(\text{Length}(C))^p$ in the energy, with $p = 2$ instead of $p = 1$]. The initial curve is the boundary of the image. After a time, a curve in the middle of the image appears and expands until merges with the initial evolving curve.

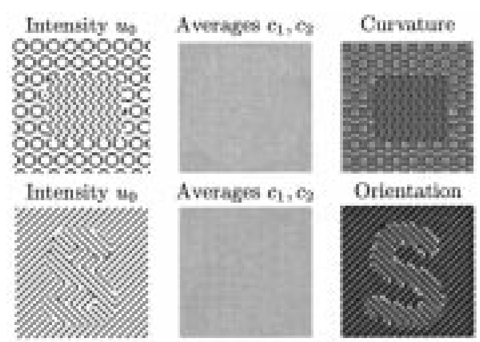


Fig. 14. Examples of images for which the averages "inside" and "outside" the objects are the same.

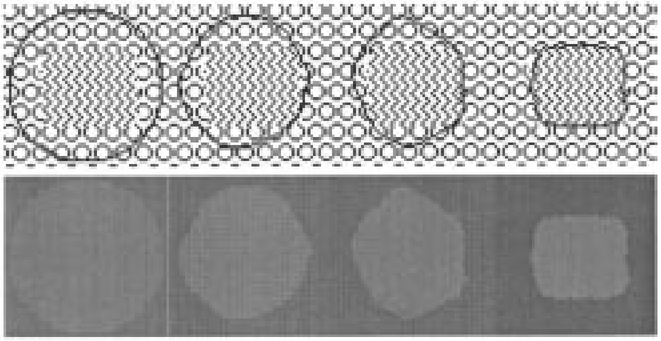


Fig. 15. Grouping based on shape identity. In our model, we replaced u_0 from Fig. 14 top left, by the curvature of the level curves of u_0 (Fig. 14 top right). Size $= 64 \times 64, \mu = 0.05 \cdot 255^2, \phi_0(x, y) = -\sqrt{(x-32.5)^2 + (y-32.5)^2} + 30.5$, 5 iterations of reinitialization, cpu $= 10.20$ s.

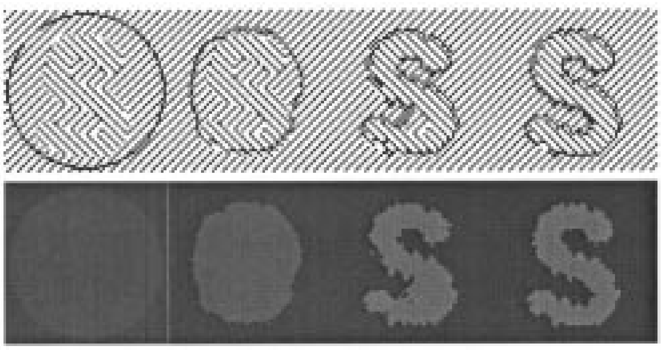


Fig. 16. Grouping based on orientation identity. In our model, we replaced u_0 from Fig. 14 bottom left, by the orientation of the normal to the level curves of u_0 (Fig. 14 bottom right). Size: 64×64, $\mu = 0.025 \cdot 255^2$, $\nu = 0.02 \cdot 255^2$, $\phi_0(x, y) = -\sqrt{(x-32.5)^2 + (y-32.5)^2} + 30$, five iterations of reinitialization, cpu $= 10.25$ s.

Finally, in Fig. 13, the algorithm detects the contours of a plane from a real noisy image.

Of course that our model has its limitations. For instance, it will be interesting to extend the model to the general case of the Mumford–Shah functional.

On the other hand, there are objects which cannot be detected using the intensity average only. For instance, we show in Fig. 14 two such examples, together with the averages inside and outside the objects, which are practically the same (left and middle). One way to overcome this difficulty, would be to use other informations from the initial image u_0, like the curvature (see Fig. 14 top right), or the orientation of level sets (see Fig. 14 bottom right). In this framework, we refer the reader to [12].

For the results from Figs. 15 and 16, we replaced in our model u_0 by $curvature(u_0) = \text{div}(\nabla u_0/|\nabla u_0|)$, and by $orientation(u_0) = \tan^{-1}(u_{0,y}/u_{0,x})$ respectively (the angle of the normal to the level curves). Other discriminants may be considered.

V. Concluding Remarks and Discussions

In this paper, we proposed an active contour model based on Mumford–Shah segmentation techniques and the level set method. Our model is not based on an edge-function to stop the evolving curve on the desired boundary. Also, 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 and preserved. By our model, we can detect objects whose boundaries are not necessarily defined by gradient or with very smooth boundaries, for which the classical active contour models are not applicable. Finally, we can automatically detect interior contours starting with only one initial curve. The position of the initial curve can be anywhere in the image, and it does not necessarily surround the objects to be detected. We validated our model by various numerical results.

Acknowledgment

The authors would like to thank J.-M. Morel and G. Sapiro for suggesting us the references on Kanisza's work [8] and Paragios–Deriche work [20], [21], respectively, to R. Fedkiw, M. J. Black, and S. Ruuth for valuable conversations, and especially to S. Osher for the very useful discussions in the "level set collective" at UCLA. Finally, they wish to thank all three unknown referees for their general and detailed comments and suggestions, which helped very much to improve the presentation of the paper.

References

[1] G. Aubert and L. Vese, "A variational method in image recovery," *SIAM J. Numer. Anal.*, vol. 34, no. 5, pp. 1948–1979, 1997.

[2] S. Byers and A. Raftery, "Nearest-neighbor cluster removal for estimating features in spatial point processes," *J. Amer. Statist. Assoc.*, vol. 93, no. 442, pp. 577–584, 1998.

[3] V. Caselles, F. Catté, T. Coll, and F. Dibos, "A geometric model for active contours in image processing," *Numer. Math.*, vol. 66, pp. 1–31, 1993.

[4] V. Caselles, R. Kimmel, and G. Sapiro, "On geodesic active contours," *Int. J. Comput. Vis.*, vol. 22, no. 1, pp. 61–79, 1997.

[5] G. Dal Maso, J. M. Morel, and S. Solimini, "A variational method in image segmentation: existence and approximation results," *Acta Matematica*, vol. 168, pp. 89–151, 1992.

[6] A. Dasgupta and A. Raftery, "Detecting features in spatial point processes with cluster via model-based clustering," *J. Amer. Statist. Assoc.*, vol. 93, no. 441, pp. 294–302, 1998.

[7] L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions*. Boca Raton, FL: CRC, 1992.

[8] G. Kanizsa, *La Grammaire du Voir. Essais sur la perception*: Diderot Editeur, Arts et Sciences, 1997.

[9] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.*, vol. 1, pp. 321–331, 1988.

[10] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzy, "Gradient flows and geometric active contour models," in *Proc. Int. Conf. Computer Vision*, Cambridge, MA, 1995, pp. 810–815.

[11] M.-S. Lee and G. Medioni, "Inferred descriptions in terms of curves, regions and junctions from sparse, noisy binary data," in *Proc. IEEE Int. Symp. Computer Vision*, Coral Gables, FL, 1995, pp. 73–78.

[12] C. Lopez and J. M. Morel, "Axiomatization of shape analysis and application to texture hyper discrimination," in *Proc. Trento Conf. Surface Tension Movement Mean Curvature*, Berlin, Germany, 1992.

[13] R. Malladi, J. A. Sethian, and B. C. Vemuri, "A topology independent shape modeling scheme," in *Proc. SPIE Conf. Geometric Methods Computer Vision II*, vol. 2031, San Diego, CA, 1993, pp. 246–258.

[14] ——, "Evolutionary fronts for topology-independent shape modeling and recovery," in *Proc. 3rd Eur. Conf. Computer Vision* Stockholm, Sweden, 1994, vol. 800, pp. 3–13.

[15] R. Malladi, J. A. Sethian, and B. C. Vemuri, "Shape modeling with front propagation: A level set approach," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, pp. 158–175, Feb. 1995.

[16] J. M. Morel and S. Solimini, *Segmentation of Images by Variational Methods: A Constructive Approach*. Madrid, Spain: Revista Matematica Universidad Complutense de Madrid, 1988, vol. 1, pp. 169–182.

[17] ——, *Segmentation d'Images par Méthode Variationnelle: Une Preuve Constructive d'Existence*. Paris, France: C.R. Acad. Sci. Paris, 1989, vol. 308, pp. 465–470.

[18] D. Mumford and J. Shah, “Optimal approximation by piecewise smooth functions and associated variational problems,” *Commun. Pure Appl. Math*, vol. 42, pp. 577–685, 1989.

[19] S. Osher and J. A. Sethian, “Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton–Jacobi Formulation,” *J. Comput. Phys.*, vol. 79, pp. 12–49, 1988.

[20] N. Paragios and R. Deriche, “Geodesic Active Regions for Texture Segmentation,” INRIA RR-3440, 1998.

[21] ——, “Geodesic active regions for motion estimation and tracking,” INRIA RR-3631, 1999.

[22] E. Rouy and A. Tourin, “A viscosity solutions approach to shape-from-shading,” *SIAM J. Numer. Anal.*, vol. 29, no. 3, pp. 867–884, 1992.

[23] L. Rudin, S. Osher, and E. Fatemi, “Nonlinear total variation based noise removal algorithms,” *Phys. D*, vol. 60, pp. 259–268, 1992.

[24] K. Siddiqi, Y. B. Lauziére, A. Tannenbaum, and S. W. Zucker, “Area and length minimizing flows for shape segmentation,” *IEEE Trans. Image Processing*, vol. 7, pp. 433–443, Mar. 1998.

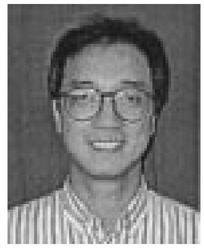
[25] M. Sussman, P. Smereka, and S. Osher, “A level set approach for computing solutions to incompressible two-phase flow,” *J. Comput. Phys.*, vol. 119, pp. 146–159, 1994.

[26] C. Xu and J. L. Prince, “Snakes, shapes and gradient vector flow,” *IEEE Trans. Image Processing*, vol. 7, pp. 359–369, Mar. 1998.

[27] H.-K. Zhao, T. Chan, B. Merriman, and S. Osher, “A variational level set approach to multiphase motion,” *J. Comput. Phys.*, vol. 127, pp. 179–195, 1996.

[28] H.-K. Zhao, S. Osher, B. Merriman, and M. Kang, “Implicit, nonparametric shape reconstruction from unorganized points using a variational level set method,” UCLA CAM Rep. 98-7, 1998.

[29] S. C. Zhu, T. S. Lee, and A. L. Yuille, “Region competition: unifying snakes, region growing, energy/bayes/MDL for multi-band image segmentation,” in *Proc. IEEE 5th Int. Conf. Computer Vision*, Cambridge, MA, 1995, pp. 416–423.



Tony F. Chan (M’98) received the B.S. degree in engineering and the M.S. degree in aerospace engineering, both in 1973, from the California Institute of Technology, Pasadena, and the Ph.D. degree in computer science from Stanford University, Stanford, CA, in 1978.

He is currently the Department Chair of the Department of Mathematics, University of California, Los Angeles, where he has been a Professor since 1986. His research interests include PDE methods for image processing, multigrid, and domain decomposition algorithms, iterative methods, Krylov subspace methods, and parallel algorithms.

Luminita A. Vese received the M.S. degree in mathematics from the University of Timisoara, Romania, in 1993 and the M.S. and Ph.D. degrees in applied mathematics, both from the University of Nice-Sophia Antipolis, France, in 1992 and 1996, respectively. Her Ph.D. subject was variational problems and partial differential equations for image analysis and curve evolution.

She held teaching and research positions at the University of Nice, Nice, France, (1996–1997), the University of Paris IX, Paris, France, (1998), and the Department of Mathematics, University of California (UCLA), Los Angeles (1997–2000). Currently she is computational and applied mathematics Assistant Professor at UCLA. Her research interests include problems of curve evolution and segmentation using PDEs, variational methods, and level set methods.