4.17

Geometric Active Contours for Image Segmentation

Vicent Caselles
Universitat Pompeu-Fabra

Ron Kimmel
Technion

Guillermo Sapiro
University of Minnesota

1 Introduction

Image segmentation is one of the most fundamental, useful, and studied topics in image processing and analysis. The goal is a partition of the image into coherent regions, which is an important initial step in the analysis of the image content. For example, before a tumor is analyzed in a computed tomography (CT) volumetric image, it has to be detected and somehow isolated from the rest of the image. Before a face is recognized, it has also to be picked out from its background. In this chapter we deal with an efficient and accurate approach in image segmentation: active contours. The general idea behind this technique is to apply partial differential equations (PDEs) to deform a curve or a surface towards the boundaries of the objects of interest in the image. The deformation is driven by forces that use information about the objects of interest in the data. In particular, we describe ideas that emerged from the *geodesic active contours* framework, concentrating on some of the main models and referring to the literature for other applications. This is an example of using PDEs for image processing and analysis. In this case, such PDEs are derived as gradient-descent processes from geometric integral measures. This research field considers images as continuous geometric structures, and enables the usage of continuous mathematics like PDEs, differential geometry, and so forth. The computer image–processing algorithms are actually the numeric implementation of the resulting equations. Related examples of this general approach can be found in Chapters 3.3, 3.6, and 4.14–4.16, whereas different image-segmentation methods are described in the introductory chapters as well as in Section 4, for example, Chapters 4.10 and 4.13.

The geodesic active contour model for image segmentation was introduced about a decade ago [9–13]. It was motivated by the snake model [37] and the geometric level-sets curve evolution models [8, 48]. Related models evolved at about

the same time [42, 43, 73, 74]. Geodesic active contours play major roles in many data analysis applications beyond image segmentation. The applications are diverse, including security visual interpretation, medical imaging, and general industrial procedures like quality control and computer-aided geometric design. Techniques that emerged from this framework are considered state of the art in many disciplines and play an important part in initiatives such as the *Insight Segmentation and Registration Toolkit* project for medical image analysis (www.itk.org). A large number of academic reports and innovative ideas flourished from the philosophy that weighted arc-length parametrization is an appropriate measure for boundary detection in images. The work by Caselles et al. [9] was of course not the first to make this observation, and among the first to stress the role of geometry and Euclidean arc-length in image segmentation were probably Mumford and Shah [51]. Yet, the geodesic active contour model provides one of the first clear formulations and efficient numeric implementations for such geometric models.

The Osher-Sethian [54] level-set formulation allowed us to apply efficient topology adaptable, and stable numeric schemes by embedding an evolving boundary contour in a higher dimensional function defined on a fixed grid. We refer to Osher and Fedkiw [52] and Sethian [68] for an extensive treatment of level-set methods for tracking evolving fronts and to Dervieux and Thomassett [22, 23] for early developments of the concepts. Level sets have been previously used in a different model for image segmentation [8, 48]. The geodesic active contour and related models are examples of the very active and successful area of using geometric measures and resulting PDEs in image processing [30, 40, 53, 65, 68].

In this chapter we present our, obviously biased yet hopefully balanced, view on the area of PDE-based segmentation methods and conclude with recent results we have found to be relevant. The presentation partially follows that of Kimmel [39]. Section 2 introduces some of the mathematic notations we use in this chapter and defines the problems. In Section 3, we formulate the idea of geometric curve evolution for segmentation and review various types of variational measures (geometric functionals). These functionals describe an integral quantity defined by the curve. Our goal is to search for curves that locally minimize these geometric integral measures. Next, in Section 4, we compute the first variation of each of these functionals and comment on how to use it in a dynamic gradient-descent curve evolution setting. Section 5 gives the level-set formulation for the various curve-evolution procedures. In Section 6, we follow the results of Goldenberg et al. [28] and present an efficient numeric scheme that couples an alternating direction–implicit multiplicative scheme, with narrow band [1, 17, 61] and redistancing via the fast marching method [32, 66, 67, 70]. The scheme is unconditionally stable and thus allows large time steps for fast convergence.

It is impossible in a single chapter to cover all the important literature related to the geodesic active contour model. The interested reader is encouraged to read other texts for additional related contributions in image processing and analysis [30, 40, 53, 65, 68].

2 Mathematic Notations and Problem Formulation

Consider a gray level image as a function $I : \Omega \to \mathbb{R}^+$ where $\Omega \subseteq \mathbb{R}^2$ is the image domain, typically a rectangle. Although we present the theory for scalar planar images, the model can be easily extended to vector valued images and higher dimensions [13, 64]. The image gradient vector field is given by $\nabla I(x,y) \equiv \{I_x, I_y\}$, were we used subscripts to denote partial derivatives, e.g., $I_x \equiv \partial I(x,y)/\partial x$. We search for a contour, $C : [0, L] \to \mathbb{R}^2$, given in a parametric form $C(s) = \{x(s), y(s)\}$, where s is an arc-length parameter, and whose normal is defined by $\vec{n}(s) = \{-y_s(s), x_s(s)\}$. This contour somehow interacts with the given image, for example, a curve whose normal aligns with the gradient vector field, where the alignment of the two vectors can be measured by their inner product that we denote by $\langle \vec{n}, \nabla I \rangle$. We also use subscripts to denote full derivatives, such that the curve tangent is given by $C_s = \{x_s, y_s\} = \{dx(s)/ds, dy(s)/ds\}$. In some cases we will also use p to indicate an arbitrary (nongeometric) parametrization of the planar curve. In which case, the tangent is $C_s = C_p/|C_p|$, and the normal can be written as

$$\vec{n} = \frac{\{-y_p, x_p\}}{|C_p|},$$

where $|C_p| = \sqrt{x_p^2 + y_p^2}$. We have the well-known relationship between the arc-length s and a general arbitrary parameter p, given by

$$ds = \sqrt{dx^2 + dy^2} = \sqrt{\left(\frac{dx(p)}{dp}\right)^2 + \left(\frac{dy(p)}{dp}\right)^2}\, dp = |C_p| dp.$$

Define, as usual, κ to be the curvature of the curve C, and the curvature vector $\kappa\vec{n} = C_{ss}$ [25]. Let us recall that a Jordan curve is a plane curve that is topologically equivalent to (a homeomorphic image of) the unit circle (i.e., it is simple and closed). If C is a Jordan curve, we also define Ω_C to be the domain inside the curve C, see Fig. 1. We also assume without loss of generality that the curves are counterclockwise oriented, so that $\vec{n}$ is the inner unit normal vector.

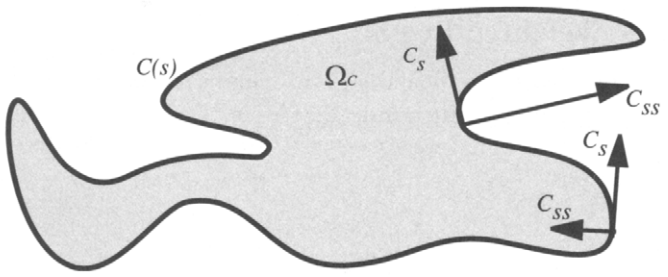

FIGURE 1 A closed curve C, with C_s the unit tangent, $\kappa\vec{n} = C_{ss}$ the curvature vector, and Ω_C the region inside the curve.

In this chapter we deal with two types of integral measures that are related via the Green theorem. The first is defined along the curve by the general form

$$E(C) = \int_0^L g(C(s))ds,$$

where g is a function of the geometry of the curve $C(s)$ and the underlying image I. Under general assumptions, this functional measures the weighted length of C (recall that s is the arc-length parametrization), where the weight is given by $g(\cdot)$. Formally, we search for the optimal planar curve C, such that

$$C = \arg\min_C E(C),$$

that is our desired geodesic. In general, we start from a specific curve and deform it to locate an extremum contour. Thus the name *geodesic active contour*.

The second functional integrates the values of the function $f(x, y)$ defined inside the curve, and is usually referred to as a region-based measure,

$$E(C) = \iint_{\Omega_C} f(x, y)dxdy,$$

where as before, Ω_C is the region inside the Jordan curve C. Under general assumptions, this is a measure of the weighted area of Ω_C, where the weight is given by the function $f(x, y)$.

3 From Edge Detectors to Geometric Evolutions

The simplest edge detectors try to locate points defined by local maxima of the image gradient magnitude. The Marr and Hildreth edges are a bit more sophisticated, and were defined as the zero crossing curves of a Laplacian of Gaussian (LoG) applied to the image [49, 50] (see also Chapter 4.14). The Marr-Hildreth edge detection and integration process can be regarded as a way to determine curves in the image plane that pass through points where the gradient is high and whose normal direction best aligns with the local edge direction as predicted by the image gradient. This observation was first made by Kimmel and Bruckstein [41] and proved to be optimal under general conditions [24].

The importance of orientation information in a variational setting for delicate segmentation tasks was recently also considered [71]: The authors proposed alignment with a general vector field as a segmentation criterion of complicated closed thin structures in three-dimensional (3D) medical images. Kimmel and Bruckstein [41] showed that the Haralick edge detector [7, 31], which is the main procedure in the Canny edge detector, can be interpreted as a solution of a two-dimensional (2D) variational principle that combines the alignment term with a topologic homogeneity measure. We will not explicitly explore this observation here and focus our attention on the geodesic active contour.

The evolution of dynamic edge integration processes and active contours started with classic snakes [37], followed by nonvariational geometric active contours [8, 48], and more recent geodesic active contours [10]. Here, we restrict our discussion to parametrization invariant (geometric) functionals that do not depend on the internal parametrization of the curve, but rather on its geometry and the properties of the image. From these functionals we extract the first variation, and use it as a gradient-descent process, also known as geometric active contour. We start by presenting a few possibilities for energies such as those defined in Section 2. The corresponding gradient descent flows lead to the geodesic active contour and recent geometric models.

3.1 Geodesic Active Contour

The geodesic active contour model [10] is defined by the functional

$$E_{GAC}(C) = \int_0^L g(C(s))ds.$$

It is an integration of an inverse edge indicator function, [i.e. any decreasing function of the modulus of the gradient, such as $g(x, y) = 1/(1 + |\nabla I|^2)$], along the contour. The search, in this case, is for a curve along which the inverse edge indicator gets the smallest possible values. This curve is a geodesic. That is, we would like to find the curve C that minimizes this functional. Modifying the function g, different results can be obtained. For example, segmentation of vector valued images [45, 64] or even solving the 3D from stereo problem [26]. This geometric energy, up to an arbitrary constant, can be obtained by manipulating the classic snakes [37] using least-action principles in physics (see also Aubert and Blanc-Feraud [2]). In addition to its use for fundamental image-processing problems, the geodesic active contour can serve as a regularization term for other variational-based segmentation formulations [41]. We point out that a well-studied example

of this functional is $g=1$, for which the functional measures the total arc-length of the curve.

3.2 Alignment Term

As pointed out by Kimmel and Bruckstein [41], g can be a function not only of the image gradient, but also of its direction. Consider the geometric functional

$$E_A(C) = -\int_0^L \langle \nabla I(x(s), y(s)), \vec{n}(s) \rangle ds,$$

or in its more "robust" form

$$E_{AR}(C) = -\int_0^L |\langle \nabla I(x(s), y(s)), \vec{n}(s) \rangle| ds,$$

where the absolute value of the inner product between the image gradient and the curve normal is our alignment measure (Fig. 2). The motivation behind E_{AR} is the fact that in many cases, the gradient direction is a good estimator for the orientation of the edge contour. The inner product gets high values if the curve normal aligns with the image gradient direction. This measure also uses the gradient magnitude as an edge indicator. Therefore, our goal would be to find curves that minimize this geometric functional, hence to maximize the alignment. This new energy can be combined with the other energies introduced in this section. Such combinations integrate additional information like the image intensity, the edge strength captured by the gradient magnitude, the edge directions captured by the gradient direction, the object area, and even shape priors. A simple penalty term is added to the variational formulation, penalizing for the discrepancy between the detected shape and the "average" shape representing the class of object of interest. Temporal changes of the intensity can be used to detect moving objects [6, 28, 56].

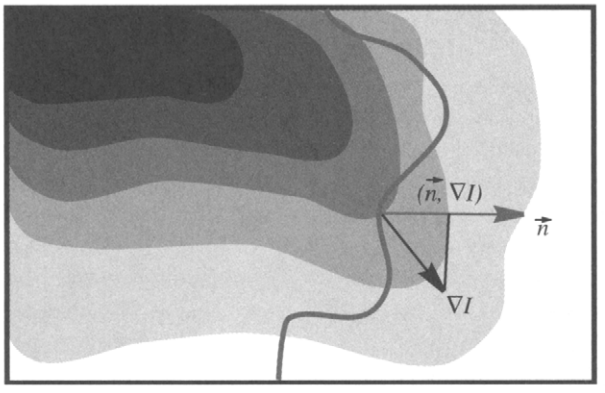

FIGURE 2 The curve C, its normal $\vec{n}$ at a specific point, and the image gradient ∇I at that point. The alignment term integrates the projection of ∇I on the normal along the curve.

3.3 Weighted Region

In some cases we would like to minimize (or maximize) an averaged quantity inside the region Ω_C defined by the Jordan curve C. In its most general setting, this weighted area measure is

$$E_W(C) = \iint_{\Omega_C} f(x,y)dxdy,$$

where $f(x,y)$ is any integrable scalar function. A simple example is $f(x,y) = 1$, for which the functional $E(C)$ measures the area inside the curve C, that is, the area of the region Ω_C that we also denote by $|\Omega_C|$. Other approaches try to find uniform regions inside C (see for example [56–59, 69, 75]) and the description in next subsection.

3.4 Minimal Variance

Chan and Vese [14] proposed a minimal variance criterion (related formulations have been proposed by Paragios and collaborators and by Yezzi and collaborators), given by

$$E_{MV}(C, c_1, c_2) = \frac{1}{2}\iint_{\Omega_C} (I(x,y) - c_1)^2 dxdy + \frac{1}{2}\iint_{\Omega \setminus \Omega_C} (I(x,y) - c_2)^2 dxdy.$$

As we will see, in the optimal case, the two constants, c_1 and c_2 are the mean intensities in the interior (inside) and the exterior (outside) of the contour C, respectively. The optimal curve would best separate the interior and exterior with respect to their relative average values. In the optimization process, we look for the best separating curve and the optimal expected values c_1 and c_2. Such optimization problems, in higher dimensions, are often encountered in color quantization and classification problems. Moreover, this formulation is simply k-means or optimal (Max-Lloyd) quantization.

To control the smoothness of their active contour, Chan and Vese [14] also included the arc-length $\int ds$ as a regularization term. Using geodesic active contours, one could actually use the more general weighted arc-length, $\int g(C(s))ds$, for which the regularization properties can be extracted from reference [10].

One could consider more generic region based measures like

$$E(C) = \iint_{\Omega_C} \|T(I(x,y)) - \bar{c}_1\| dxdy + \iint_{\Omega \setminus \Omega_C} \|T(I(x,y)) - \bar{c}_2\| dxdy,$$

where T is a general transformation of the image, the norm $\|\cdot\|$ can be chosen according to the problem in hand, and $\bar{c_1}$ and $\bar{c_2}$ are vectors of possible parameters. One such example is the robust measure

$$E_{RMV}(C) = \iint_{\Omega_C} |I(x,y) - c_1| dxdy + \iint_{\Omega \setminus \Omega_C} |I(x,y) - c_2| dxdy.$$

3.5 Intermezzo

We have shown a number of variational formulations that lead to segmentation. The basic idea is to find a curve that minimizes a given geometric energy. Next, we show how to extract the curve itself. We emphasize that other geometric measures were reported in the literature. Actually, for each application, one should modify and engineer his/her own measures that best fit the problem at hand. For example, a recent popular effort is to add information in the form of shape priors. The basic idea is to add a term that penalizes the deviation of the computed contour from an "average" shape. This is useful when working with a particular class of shapes (e.g., the human heart). Examples for shape priors can be found in the literature [20, 33, 34, 38, 44, 55, 62].

4 Calculus of Variations for Geometric Measures

Let $C : [0, L] \to \mathbb{R}^2$ be a Jordan curve. Given a curve integral of the general form

$$E(C) = \oint_C L(C_p, C) dp,$$

where p is an arbitrary parameter, we compute the first variation by

$$\frac{\delta E(C)}{\delta C}(\eta) = \frac{d}{d\epsilon} E(C + \epsilon\eta)|_{\epsilon=0} = \lim_{\epsilon \to 0} \frac{E(C + \epsilon\eta) - E(C)}{\epsilon}$$

where $\eta : [0, L] \to \mathbb{R}^2$ is a C^1 curve. The extrema of the functional $E(C)$ can be identified by the Euler-Lagrange equation $\delta E(C)/\delta C = 0$. A dynamic process known as gradient descent, which takes an arbitrary curve toward a minimum of $E(C)$, is given by the curve evolution equation

$$\frac{\partial C}{\partial t} = -\frac{\delta E(C)}{\delta C}, \tag{1}$$

where we added a virtual "time" parameter t to our curve to allow its evolution into a family of planar curves $C(p, t)$. Our hope is that this evolution process would take an almost arbitrary initial curve into a desired configuration, which gives a significant minimum of our functional. In this chapter, we restrict ourselves to closed contours. When considering open contours (or open surfaces), one should also handle the end points and add additional constraints to determine their optimal locations (see for example [5, 27, 41]). We now present a few examples of this gradient-descent flow.

Unless stated otherwise, the arc-length parameter of the curve C will be denoted by s. Recall that a function $F : \mathbb{R}^n \to \mathbb{R}^m$, $n, m \geq 1$, is said to be of class C^k if all partial derivatives up to order k are continuous.

Lemma 1 *Assume that $g : \mathbb{R}^2 \to \mathbb{R}$ is of class C^1. The geodesic active contour model is*

$$E_{GAC}(C) = \oint_C g(C(s)) ds,$$

for which the first variation is given by

$$\frac{\delta E_{GAC}(C)}{\delta C} = -(\kappa g - \langle \nabla g, \vec{n} \rangle)\vec{n}.$$

Proof: Let $\eta : [0, L] \to \mathbb{R}^2$ be a closed curve of class C^1, and $\epsilon \in \mathbb{R}$. Using that

$$\frac{d}{d\epsilon} |C_s + \epsilon\eta_s||_{\epsilon=0} = \langle C_s, \eta_s \rangle, \tag{2}$$

we compute

$$\begin{aligned}
\frac{d}{d\epsilon} E_{GAC}(C + \epsilon\eta)|_{\epsilon=0} &= \frac{d}{d\epsilon}|_{\epsilon=0} \oint g(C(s) + \epsilon\eta(s))|C_s + \epsilon\eta_s| \, ds \\
&= \oint \langle \nabla g(C(s)), \eta(s) \rangle \, ds + \oint g(C(s)) \langle C_s, \eta_s \rangle \, ds \\
&= \oint \langle \nabla g(C(s)), \eta(s) \rangle \, ds - \oint \langle \nabla g(C(s)), C_s \rangle \langle C_s, \eta \rangle \, ds \\
&\quad - \oint g(C(s)) \langle C_{ss}, \eta \rangle \, ds \\
&= \oint \langle \nabla g(C(s)), C_s^{\perp} \rangle \langle C_s^{\perp}, \eta \rangle \, ds - \oint g(C(s)) \langle C_{ss}, \eta \rangle \, ds
\end{aligned}$$

where the third equality is a result of integration by parts.

If we set $g = 1$, the gradient descent curve evolution result given by $C_t = -\delta E_{GAC}(C)/\delta C$ is the well-known curvature flow, $C_t = \kappa\vec{n}$ or equivalently $C_t = C_{ss}$, also known as the geometric heat equation. The geodesic active contour model assumes that $g(x) = \tilde{g}(|\nabla I(x)|)$, $x \in \mathbb{R}^2$, where $\tilde{g}$ is a decreasing function of the modulus of the gradient.

Let R denote the counterclockwise rotation matrix of angle $\pi/2$, and, for any matrix A, let A^t denote its transposed.

Lemma 2 *Let $F:\mathbb{R}\to\mathbb{R}$ be a function of class C^2. Given the vector field $\vec{V}(x,y)=\{u(x,y),v(x,y)\}$, we define the alignment measure,*

$$E_A(C)=\oint_C F(\langle\vec{V}(C(s)),\vec{n}\rangle)ds$$

for which the first variation is given by

$$\frac{\delta E(C)}{\delta C}=F'(\langle\vec{V}(C),\vec{n}\rangle)(\mathrm{div}V)(C)\vec{n}-(F'(\langle\vec{V}(C),\vec{n}\rangle))_sR^tV(C)$$
$$-[(G(\langle\vec{V}(C),\vec{n}\rangle))_sC_s+G(\langle\vec{V}(C),\vec{n}\rangle)k\vec{n}]. \tag{3}$$

Typical examples of F are $F(z)=-|z|^p$, $p\geq 1$. Observe that when $1\leq p<2$, F is of class C^2 on $\mathbb{R}\setminus\{0\}$, but with a proper adaptation of the proof, one can still perform the computations below. The minus sign accounts for maximizing the alignment while minimizing E_A. Here we give an alternative proof to that of reference [41].

Proof: Let $\eta:[0,L(C)]\to\mathbb{R}^2$ be a closed curve of class C^1, and $\epsilon\in\mathbb{R}$. Notice that, if both ϵ and η are not zero, the curve $C+\epsilon\eta$ is not parameterized by arc-length anymore. Using (2) and

$$\frac{d}{d\epsilon}\frac{C_s^\perp+\epsilon\eta_s^\perp}{|C_s^\perp+\epsilon\eta_s^\perp|}|_{\epsilon=0}=\eta_s^\perp-\langle C_s^\perp,\eta_s^\perp\rangle C_s^\perp$$

we compute

$$\begin{aligned}
&\frac{d}{d\epsilon}E(C+\epsilon\eta)|_{\epsilon=0}\\
&=\frac{d}{d\epsilon}|_{\epsilon=0}\oint_C F\left(\left\langle\vec{V}(C+\epsilon\eta),\frac{C_s^\perp+\epsilon\eta_s^\perp}{|C_s^\perp+\epsilon\eta_s^\perp|}\right\rangle\right)|C_s+\epsilon\eta_s|\,ds\\
&=\oint_C F'(\langle\vec{V}(C),\vec{n}\rangle)\langle D\vec{V}(C)(\eta),C_s^\perp\rangle\,ds\\
&\quad+\oint F'(\langle\vec{V}(C),\vec{n}\rangle)\langle\vec{V}(C),\eta_s^\perp\rangle\,ds\\
&\quad-\oint F'(\langle\vec{V}(C),\vec{n}\rangle)\langle\vec{V}(C),C_s^\perp\rangle\langle C_s^\perp,\eta_s^\perp\rangle\,ds\\
&\quad+\oint F(\langle\vec{V}(C),\vec{n}\rangle)\langle C_s,\eta_s\rangle\,ds.
\end{aligned}$$

Denote $G(z)=F(z)-zF'(z)$. Then, using $\vec{n}=C_s^\perp$ and $\langle C_s^\perp,\eta_s^\perp\rangle=\langle C_s,\eta_s\rangle$, after integration by parts, we obtain

$$\begin{aligned}
\frac{d}{d\epsilon}E(C+\epsilon\eta)|_{\epsilon=0}&=\oint_C F'(\langle\vec{V}(C),\vec{n}\rangle)\langle D\vec{V}(C)^t(\vec{n}),\eta\rangle\,ds\\
&\quad-\oint[(F'(\langle\vec{V}(C),\vec{n}\rangle))_s\langle R^tV(C),\eta\rangle\\
&\quad+F'(\langle\vec{V}(C),\vec{n}\rangle)\langle R^tD\vec{V}(C)(C_s),\eta\rangle]\,ds\\
&\quad-\oint[(G(\langle\vec{V}(C),\vec{n}\rangle))_s\langle C_s,\eta\rangle\\
&\quad+G(\langle\vec{V}(C),\vec{n}\rangle)\langle C_{ss},\eta\rangle]\,ds\\
&=\oint_C F'(\langle\vec{V}(C),\vec{n}\rangle)[\langle D\vec{V}(C)^t(\vec{n}),\eta\rangle\\
&\quad-\langle R^tD\vec{V}(C)(C_s),\eta\rangle]\,ds\\
&\quad-\oint(F'(\langle\vec{V}(C),\vec{n}\rangle))_s\langle R^tV(C),\eta\rangle\\
&\quad-\oint[(G(\langle\vec{V}(C),\vec{n}\rangle))_s\langle C_s,\eta\rangle\\
&\quad+G(\langle\vec{V}(C),\vec{n}\rangle)k\langle\vec{n},\eta\rangle]\,ds.
\end{aligned}$$

Thus, the first variation of $E_A(C)$ is

$$\begin{aligned}
\frac{\delta E(C)}{\delta C}&=F'(\langle\vec{V}(C),\vec{n}\rangle)[D\vec{V}(C)^t(\vec{n})-R^tD\vec{V}(C)(C_s)]\\
&\quad-(F'(\langle\vec{V}(C),\vec{n}\rangle))_sR^tV(C)\\
&\quad-[(G(\langle\vec{V}(C),\vec{n}\rangle))_sC_s+G(\langle\vec{V}(C),\vec{n}\rangle)k\vec{n}].
\end{aligned}$$

Now, observe that $D\vec{V}(C)^t(\vec{n})-R^tD\vec{V}(C)(C_s)=D\vec{V}(C)^tRC_s-R^tD\vec{V}(C)(C_s)=(\mathrm{div}V)(C)\vec{n}$, which gives (3).

Notice that the second term in the right-hand side of (3) can be further expanded as ([41])

$$\begin{aligned}
&(F'(\langle\vec{V}(C),\vec{n}\rangle))_sR^tV(C)\\
&\quad=(F'(\langle\vec{V}(C),\vec{n}\rangle))_s[\langle V(C),R\vec{n}\rangle\vec{n}+\langle V(C),RC_s\rangle C_s]\\
&\quad=F''(\langle\vec{V}(C),\vec{n}\rangle)[\langle V(C)_s,\vec{n}\rangle\langle V(C),C_s\rangle\\
&\qquad-\langle V(C),C_s\rangle^2]\vec{n}+\text{tangential components.}
\end{aligned}$$

An important example of (3) corresponds to the choice $\vec{V}=\nabla I$, for which the formula for the first variation is

$$\frac{\delta E(C)}{\delta C}=F'(\langle\nabla I(C),\vec{n}\rangle)\Delta I(C)\vec{n}-(F'(\langle\nabla I(C),\vec{n}\rangle))_sR^t\nabla I(C)$$
$$-[(G(\langle\nabla I(C),\vec{n}\rangle))_sC_s+G(\langle\nabla I(C),\vec{n}\rangle)k\vec{n}]$$

where $\Delta I = I_{xx} + I_{yy}$ is the image Laplacian.

The robust alignment term

$$E_{AR}(C) = -\oint_C |\langle \vec{V}, \vec{n} \rangle| ds,$$

corresponds to $F(z) = -|z|$ which is C^2 in $\mathbb{R} \setminus \{0\}$. In this case $G(z) = 0$ and the first variation takes the form

$$\frac{\delta E_{AR}(C)}{\delta C} = -\operatorname{sign}(\langle \nabla I(C), \vec{n}(s) \rangle) \Delta I(C) \vec{n} - (\operatorname{sign}(\langle \nabla I(C), \vec{n}(s) \rangle))_s \nabla I(C)^{\perp}.$$

If $\langle \nabla I(C), \vec{n}(s) \rangle$ has constant sign, then the second term vanishes and we obtain

$$\frac{\delta E_{AR}(C)}{\delta C} = -\operatorname{sign}(\langle \nabla I(C), \vec{n}(s) \rangle) \Delta I(C) \vec{n}. \tag{4}$$

Thus, in this case, the Euler-Lagrange equation $\frac{\delta E_{AR}(C)}{\delta C} = 0$ gives a variational explanation of the Marr-Hildredth edge detector which is defined by the zero-crossings of the Laplacian as first indicated in Kimmel and Bruckstein [41].

Lemma 3 *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be an integrable function. The weighted region functional*

$$E_W(C) = \iint_{\Omega_C} f(x,y) dx dy,$$

yields the first variation

$$\frac{\delta E_W(C)}{\delta C} = -f(x,y)\vec{n}.$$

Proof: Following [75], we define the two functions $P(x,y)$ and $Q(x,y)$, such that $P_y(x,y) = -\frac{1}{2} f(x,y)$ and $Q_x = \frac{1}{2} f(x,y)$. We readily have that $f(x,y) = Q_x - P_y$. Next, using Green's theorem we can write

$$\begin{aligned} E(C) &= \iint_{\Omega_C} f(x,y) dx dy = \iint_{\Omega_C} (Q_x - P_y) dx dy \\ &= \int_C (P dx + Q dy) = \int_C (P x_s + Q y_s) ds \\ &= \int_C \langle \{-Q, P\}, \vec{n} \rangle ds, \end{aligned}$$

and the weighted region energy is expressed as an alignment energy with $\vec{V} = \{-Q, P\}$ and $F(z) = z$. Using Lemma 2, we obtain

$$\frac{\delta E(C)}{\delta C} = \operatorname{div}(\{-Q, P\})\vec{n} = -(Q_x - P_y)\vec{n} = -f\vec{n}.$$

This term is sometimes called the *weighted area* [75] term, and for f constant, its resulting variation is known as the "balloon" [18] force. If we set $f = 1$, the gradient descent curve evolution process is the constant flow. It generates offset curves via $C_t = \vec{n}$, and its efficient implementation is closely related to Euclidean distance maps [15, 21] and fast marching methods [32, 66, 67, 70].

Lemma 4 *The Chan-Vese minimal variance criterion [14] is given by*

$$E_{MV}(C, c_1, c_2) = \frac{1}{2} \iint_{\Omega_C} (I - c_1)^2 dx dy + \frac{1}{2} \iint_{\Omega \setminus \Omega_C} (I - c_2)^2 dx dy,$$

for which the first variation is

$$\begin{aligned} \frac{\delta E_{MV}}{\delta C} &= -(c_2 - c_1)\Big(I - \frac{c_1 + c_2}{2}\Big)\vec{n} \\ \frac{\delta E_{MV}}{\delta c_1} &= \iint_{\Omega_C} (c_1 - I) dx dy \\ \frac{\delta E_{MV}}{\delta c_2} &= \iint_{\Omega \setminus \Omega_C} (c_2 - I) dx dy. \end{aligned}$$

Proof: Using Lemma 3, we have the first variation given by

$$\begin{aligned} \frac{\delta E_{MV}}{\delta C} &= -\frac{1}{2}\big((I - c_1)^2 - (I - c_2)^2\big)\vec{n} \\ &= -\frac{1}{2}\big(I^2 - 2Ic_1 + c_1^2 - I^2 + 2Ic_2 - c_2^2\big)\vec{n} \\ &= -\Big((c_2 - c_1)I - \frac{(c_1 + c_2)(c_2 - c_1)}{2}\Big)\vec{n} \\ &= -(c_2 - c_1)\Big(I - \frac{c_1 + c_2}{2}\Big)\vec{n}. \end{aligned}$$

The optimal c_1 and c_2, extracted from $\delta E_{MV}/\delta c_1 = 0$ and $\delta E_{MV}/\delta c_2 = 0$, are the average intensities of the image inside and outside the contour, respectively.

One could recognize the variational interpretation of segmentation by the threshold $(c_1 + c_2)/2$ given by the Euler Lagrange equation $\delta E_{MV}/\delta C = 0$.

Finally, in a similar way, we obtain the first variation of the robust minimal deviation measure E_{RMV}.

Lemma 5 *The robust minimal total deviation criterion is given by*

$$E_{RMV}(C, c_1, c_2) = \iint_{\Omega_C} |I - c_1| dx dy + \iint_{\Omega \setminus \Omega_C} |I - c_2| dx dy,$$

for which the first variation is

$$\frac{\delta E_{RMV}}{\delta C} = -(|I - c_1| - |I - c_2|)\vec{n}$$
$$\frac{\delta E_{RMV}}{\delta c_1} = \iint_{\Omega_C} \text{sign}(c_1 - I)dxdy$$
$$\frac{\delta E_{RMV}}{\delta c_2} = \iint_{\Omega \setminus \Omega_C} \text{sign}(c_2 - I)dxdy,$$

where $\text{sign}(r) = +1$ *if* $r > 0$; -1 *if* $r < 0$; *and it can be any value in* $[-1, 1]$ *if* $r = 0$.

Heuristically, the optimal value of c_1 is the value of $I(x, y)$ in Ω_C that splits its area into two equal parts. From obvious reasons we refer to this value as the median of I in Ω_C, or formally,

$$c_1 = \text{median}_{\Omega_C} I(x, y).$$

In a similar way, the optimal value of c_2 is

$$c_2 = \text{median}_{\Omega \setminus \Omega_C} I(x, y).$$

The computation of c_1 and c_2 can be efficiently implemented via the intensity histograms in the interior or the exterior of the contour. One rough discrete approximation is the median of the pixels inside or outside the contour.

The robust minimal deviation term should be preferred when the penalty for isolated pixels with wrong affiliation is insignificant. The minimal variance measure penalizes large deviations in a quadratic fashion and would tend to oversegment such events or require large regularization that could oversmooth the desired boundaries. For an example of coupling these terms with the alignment terms for 3D thin-structure detection in medical images we refer to Holtzman-Gazit et al. [35].

4.1 Another Intermezzo

We reviewed a number of possible curve flows to deform a given initial curve toward a steady-state contour that defines the boundaries between objects, a process known as image segmentation. These flows are defined via gradient descent obtained from the geometric energies presented in the previous section. We next review the level-set framework, which allows stable and efficient implementation of such flows.

5 Gradient Descent in Level Set Formulation

We have just reviewed some fundamental curve evolution/flows that are useful for image segmentation. The next step is the implementation of such flows. For this, we embed a closed curve in a higher dimensional $\phi(x, y)$ function, which implicitly represents the curve C as a zero set [i.e., $C = \{\{x, y\}: \phi(x, y) = 0\}$]. This way, the well-known Osher-Sethian [54] level-set method can be used to implement the curve propagation toward its optimal location. Figure 3 presents two planar curves, and a single function which is an implicit representation for both curves.

Given the curve evolution equation $C_t = \gamma\vec{n}$, its implicit level set evolution equation reads

$$\phi_t = \gamma|\nabla\phi|.$$

For that, one assumes that the Jordan curve $C(s, t)$ is a level set of the evolving function $\phi(t, x, y)$, $(x, y) \in \mathbb{R}^2$. To fix ideas, let us assume that $C(t)$ is the zero level set of $\phi(t, x, y)$,

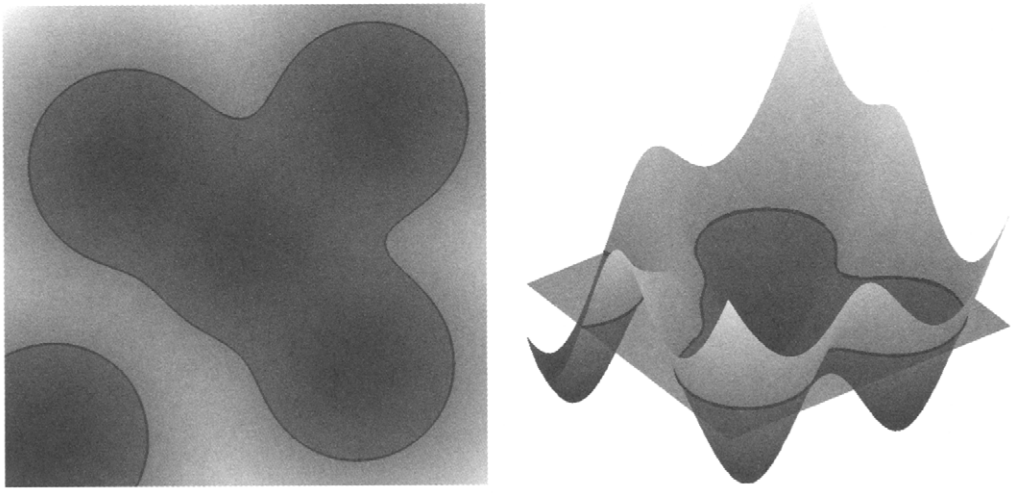

FIGURE 3 Left: Two given disconnected planar curves. Right: An implicit representation of the two curves as a single level set of a smooth function rendered by a perspective projection of the function. The level set, plotted as black contours, is the intersection of the function with a plane.

and $\phi(t,x,y)$ is negative inside the zero level-set, and positive outside (in some cases, the signed distance function is a preferred choice). Thus, we have

$$\phi(t, C(t)) = 0.$$

Differentiating the above identity with respect to t we obtain

$$\phi_t + \langle \nabla\phi, C_t \rangle = 0.$$

With the orientation convention of Section 2 we have the relation $\vec{n} = -\nabla\phi/|\nabla\phi|$, hence

$$\phi_t = -\langle \nabla\phi, C_t \rangle = -\langle \nabla\phi, \gamma\vec{n} \rangle = \gamma \left\langle \nabla\phi, \frac{\nabla\phi}{|\nabla\phi|} \right\rangle = \gamma|\nabla\phi|.$$

More on the rigorous equivalence between both flows and the particular relevance for the geodesic active contours can be found in Caselles and colleagues [10, 12], where the reader is also referred to theoretical results on existence, uniqueness, and consistency of the geodesic active contours flow.

Those familiar with the optical flow problem in image analysis could easily recognize this derivation. There is an interesting relation between the classic optical flow problem and the level-set method. Level-set formulation for the evolution of a family of embedded curves can be interpreted as a dynamic image synthesis process. On one hand, optical flow in image analysis is a search for the motion of features in the image. These two inverse problems share the same fundamental equation. Computing a vector field that represents the flow of the gray level sets from a given sequence of images is known as the *normal flow* computation. Next, at a higher level of image understanding, the motion field of objects in an image is known as the *optical flow*. On the other hand, in the level-set formulation, the goal is to compute the dynamics of an arbitrary image, in which one level set represents a specific curve, from a given motion vector field of that specific level set. The image in this case is an implicit representation of its level sets, whereas the vector field itself could be extracted from either the geometric properties of the level sets, or from another image or external data.

As a first example, consider the gradient descent flow [(see (1)] for the special case of the robust alignment term (4) from [41], given by

$$C_t = \text{sign}(\langle \nabla I, \vec{n} \rangle)\Delta I \vec{n}.$$

The corresponding level set evolution is

$$\phi_t = -\text{sign}(\langle \nabla I, \nabla\phi \rangle)\Delta I|\nabla\phi|.$$

We can add the geodesic active contour term, the threshold term, or its dynamic expectation version defined by the minimal variance criterion. The optimization problem takes the form of

$$\arg\min_{C, c_1, c_2} E(C, c_1, c_2),$$

for the functional

$$\begin{aligned} E(C,c_1,c_2) &= E_{AR}(C,c_1,c_2) + \alpha E_{GAC}(C) + \beta E_{MV}(C) \\ &= -\oint_C |\langle \nabla I, \vec{n} \rangle| ds + \alpha \oint_C g(C(s)) ds \\ &\quad + \frac{\beta}{2}\left(\iint_{\Omega_C} (I - c_1)^2 dxdy + \iint_{\Omega \setminus \Omega_C} (I - c_2)^2 dxdy \right), \end{aligned}$$

where α and β are positive constants, and, in practice, α is small so that the geodesic active contour term is used mainly for regularization. Using the computed first variations of all energy terms, the gradient descent flow is

$$\begin{aligned} C_t = \big[&\text{sign}(\langle \vec{n}, \nabla I \rangle)\Delta I + \alpha\big(g(x,y)\kappa - \langle \nabla g, \vec{n} \rangle\big) \\ &+ \beta(c_2 - c_1)(I - (c_1 + c_2)/2)\big]\vec{n}. \end{aligned}$$

$$c_1 = \frac{1}{|\Omega_C|} \iint_{\Omega_C} I(x,y) dxdy$$

$$c_2 = \frac{1}{|\Omega \setminus \Omega_C|} \iint_{\Omega \setminus \Omega_C} I(x,y) dxdy,$$

where $|\Omega_C|$ denotes the area of the regions Ω_C. One could recognize the relation to the Max-Lloyd quantization method, as the simplest implementation for this system is a sequential process that involves a change of the segmentation set followed by an update of the mean representing each set. The difference is the additional penalties and resulting forces we use for the smoothness and alignment of the boundary contours.

The level-set formulation of the curve evolution equation is

$$\begin{aligned} \phi_t = \Big[&-\text{sign}(\langle \nabla\phi, \nabla I \rangle)\Delta I + \alpha \text{div}\left(g(x,y) \frac{\nabla\phi}{|\nabla\phi|} \right) \\ &+ \beta(c_2 - c_1)\left(I - \frac{c_1 + c_2}{2} \right) \Big] |\nabla\phi|. \end{aligned}$$

Efficient solutions for this equation can use additive operator splitting (AOS) [46, 47, 72], ADI, or locally one-dimensional (LOD) methods [39], coupled with a narrow band approach [1, 16, 61] (see [28] for the geodesic active contours case). In the next section we use a simple first-order, implicit alternating direction–multiplicative scheme.

The following table summarizes some of the functionals reported, the resulting first variation for each functional, and the level set formulations for these terms.

Measure	$E(C)$	$-\delta E/\delta C$	$-\delta E/\delta C$ in level-set form
Weighted area	$\iint_{\Omega_C} f(x,y)dxdy$	$f(x,y)\vec{n}$	$f(x,y)\lvert\nabla\phi\rvert$
Minimal variance	$\iint_{\Omega_C}(I-c_1)^2dxdy+$ $\iint_{\Omega\setminus\Omega_C}(I-c_2)^2dxdy$	$(c_2-c_1)\times$ $(I-(c_1+c_2)/2)\vec{n}$	$(c_2-c_1)\times$ $(I-(c_1+c_2)/2)\lvert\nabla\phi\rvert$
GAC	$\oint_C g(C(s))ds$	$(\kappa g-\langle\nabla g,\vec{n}\rangle)\vec{n}$	$\mathrm{div}g\frac{\nabla\phi}{\lvert\nabla\phi\rvert}\lvert\nabla\phi\rvert$
Alignment	$-\oint\lvert\langle\nabla I,\vec{n}\rangle\rvert ds$	$(\langle\nabla I,\vec{n}\rangle)\Delta I\vec{n}$	$-\mathrm{sign}(\langle\nabla I,\nabla\phi\rangle)\Delta I\lvert\nabla\phi\rvert$

GAC, geodesic active contour.

6 Efficient Numeric Schemes

Weickert et al. [72] used an unconditionally stable, and thus efficient, numeric scheme for nonlinear inhomogeneous isotropic image diffusion known as *additive operator splitting* (AOS) that was first introduced by Lu et al. [46, 47] and has some nice symmetry and parallel properties. To obtain a fast numeric implementation, Goldenberg et al. [28] coupled the AOS with fast marching on regular grids [15, 32, 67, 70], with multiresolution [59], and with the narrow band approach [1, 16, 48, 61]. Here, following Kimmel [39] and Kimmel and Bruckstein [41], we present an extension of these efficient numeric methods for the geodesic active contour [10] presented by Goldenberg et al. [28], for the variational edge integration models [41] and minimal variance [14]. We review efficient numeric schemes and modify them to solve the level-set formulation of edge integration and object segmentation problem in image analysis.

Let us analyze the following level set formulation,

$$\phi_t=\left(\alpha\mathrm{div}\left(g(x,y)\frac{\nabla\phi}{|\nabla\phi|}\right)+\eta(\phi,\nabla I)\right)|\nabla\phi|,$$

$$\eta(\phi,\nabla I)=-\mathrm{sign}(\langle\nabla I,\nabla\phi\rangle)\Delta I+\beta(c_2-c_1)\left(I-\frac{c_1+c_2}{2}\right).$$

If we assume $\phi(x,y;t)$ to be a distance function of its zero set, then, we could approximate the short time evolution of the above equation by setting $|\nabla\phi|=1$. The short time evolution for a distance function ϕ is thereby[1]

$$\begin{aligned}\phi_t&=\alpha\mathrm{div}(g(x,y)\nabla\phi)+\eta(\phi,\nabla I)\\&=\alpha\frac{\partial}{\partial x}\left(g(x,y)\frac{\partial}{\partial x}\phi\right)+\alpha\frac{\partial}{\partial y}\left(g(x,y)\frac{\partial}{\partial y}\phi\right)+\eta(\phi,\nabla I).\end{aligned}$$

[1]As it evolves, ϕ immediately departs from being a distance function of its zero set. We can ignore this second-order effect, while redistancing every numeric time step.

Note, that when using a narrow band around the zero set to reduce computational complexity on sequential computers, the distance from the zero set needs to be recomputed to determine the width of the band at each iteration. Thus, there is no additional computational cost in simplifying the model while considering a distance map rather than an arbitrary smooth function. We thereby enjoy both the efficiency of the simplified, almost linear model, and the low computational cost of the narrow band.

Denote the operators

$$A_1=\frac{\partial}{\partial x}g(x,y)\frac{\partial}{\partial x}$$
$$A_2=\frac{\partial}{\partial y}g(x,y)\frac{\partial}{\partial y}.$$

Using these notations we can write the evolution equation as

$$\phi_t=\alpha(A_1+A_2)\phi+\eta(\phi,\nabla I).$$

Next, we approximate the time derivative using forward approximation $\phi_t\approx(\frac{\phi^{n+1}-\phi^n}{\tau})$, that yields the explicit scheme

$$\begin{aligned}\phi^{n+1}&=\phi^n+\tau(\alpha(A_1+A_2)\phi^n+\tau\eta(\phi^n,\nabla I))\\&=(\mathcal{I}+\tau\alpha(A_1+A_2))\phi^n+\tau\eta(\phi^n,\nabla I),\end{aligned}$$

where, after sampling the x,y plane, $\mathcal{I}$ is the identity matrix and I is our input image. The operators A_l become matrix operators, and ϕ^n is represented as a vector in either column or row stack, depending on the acting operator. This way, the operators A_l are tri-diagonal, which makes its inverse computation fairly simple using Thomas algorithm. Note that in the explicit scheme there is no need to invert any operator, yet the numeric time step is bounded for stability.

Let us first follow [72], and use a simple discretization for the A_l, $l\in\{1,2\}$ operators. For a given row, let

$$\frac{\partial}{\partial x}\left(g(x)\frac{\partial}{\partial x}\phi\right)\approx\sum_{j\in\mathcal{N}(i)}\frac{g_j+g_i}{2h^2}(\phi_j-\phi_i),$$

where $\mathcal{N}(i)$ is the set $\{i-1, i+1\}$, representing the two horizontal neighbors of pixel i, and h is the space between neighboring pixels. The elements of A_1 are thereby given by

$$a_{ij} = \begin{cases} \frac{g_i + g_j}{2h^2} & j \in \mathcal{N}(i) \\ -\sum_{k \in \mathcal{N}(i)} \frac{g_i + g_k}{2h^2} & j = i \\ 0 & \text{else.} \end{cases}$$

To construct an unconditionally stable scheme we use a locally one-dimensional (LOD) scheme adopted for our problem. We use the following scheme

$$\phi^{n+1} = \prod_{l=1}^{2} (\mathcal{I} - \tau \alpha A_l)^{-1} (\phi^n + \tau \eta(\phi^n, \nabla I)).$$

In one dimension it is also known as fully implicit, or backward Euler scheme. It is a first-order implicit numeric approximation, since we have that

$$\begin{aligned} &(\mathcal{I} - \tau A_1)^{-1} (\mathcal{I} - \tau A_2)^{-1} (\phi + \tau \eta) \\ &\quad = \left(\mathcal{I} - \tau A_1 - \tau A_2 + O(\tau^2)\right)^{-1} (\phi + \tau \eta) \\ &\quad = \phi + \tau (A_1 + A_2)\phi + \tau \eta + O(\tau^2), \end{aligned}$$

where we applied the Taylor series expansion in the second equality. First-order accuracy is sufficient, as our goal is the steady-state solution. We also included the weighted region–balloon, minimal variance, and the alignment terms within this implicit scheme, while keeping the first-order accuracy and stability properties of the method. The operators $\mathcal{I} - \tau A_l$ are positive definite symmetric operators, which make the implicit process unconditionally stable, using either the above multiplicative or the additive (AOS) schemes. If we have an indication that the contour is getting closer to its final destination, we could switch to an explicit scheme for the final refinement steps with a small time step. In this case, the time step should be within the stability limits of the explicit scheme. In our implementation we also use a multi–resolution pyramidal approach, where the coarse grid still captures the details of the objects we would like to detect, for MATLAB code of the LOD scheme we refer to Kimmel [40].

7 Examples

We now present a number of examples of the above-described segmentation geometric flows.

Figure 4 shows segmentation results for a relatively clean image. In this example, the alignment and minimal variance terms were tuned to play the dominant role in the segmentation. The right-hand frame shows the final contour in black painted on the original image in which the dynamic range is mapped into a subinterval of the original one.

In the next example, shown in Fig. 5, high noise and uniform illumination calls for minimal variance coupled with regularization of the contour. The alignment term does not contribute much in such a high noise.

Lastly, Fig. 6 shows 3D examples for brain aneurysms from computed tomography angiography images, [33, 34]. For these difficult data, some of the above-mentioned energies have been extended to 3D and combined with additional shape information.

8 Additional Comments on Related Developments

So far, we have mentioned some of the fundamental contributions in the area of geometric active contours. Let us conclude with some additional comments on related work.

We shall follow the notation conventions of the previous sections. A snake is an active contour defined by $C(q)$, where $q \in [0, 1]$ is a general parameter of C. We also assume that $C(q)$ is a Jordan curve enclosing a domain. As introduced by Kass-Witkin-Terzopoulos [37] the energy of a snake is

$$E(C) = \alpha \int_0^1 |C'(q)|^2 dq + \beta \int_0^1 |C''(q)|^2 dq - \lambda \int_0^1 |\nabla I(C(q))| dq,$$

which is minimized by steepest descent. As argued above, a better choice would be to take the parameter q to be the arc-length of the curve. Trying to integrate the photometric

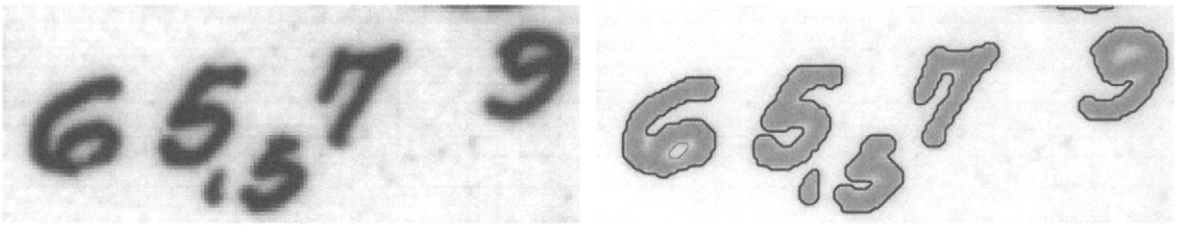

FIGURE 4 The simplest case in which alignment and minimal variance played the dominant factors in finding the exact location of the edges. Image courtesy Kimmel and Brockstein [41].

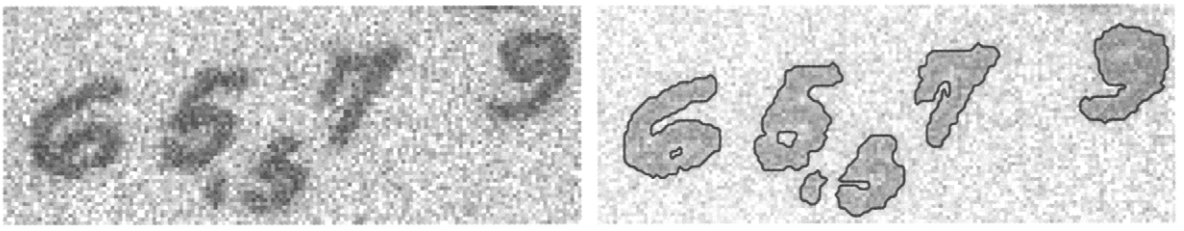

FIGURE 5 For noisy images the alignment term is turned off, while the minimal variance and regularization by the geodesic active contour are the important factors. Image courtesy of Kimmel and Bruckstein [41].

and geometric constraints, Fua and Leclerc [27] proposed to minimize the functional

$$E_P(C) = -\frac{1}{L(C)}\int_0^{L(C)} |\nabla I(C(s))|\, ds \tag{6}$$

where s is the arc-length parameter and $L(C)$ denotes the length of C. The energy is an average along the curve of the magnitude of the gradient, hence, the edge information is integrated along the length of the curve producing a meaningful functional for open. In Kimmel and Bruckstein [41], minimizing the "normalized alignment"

$$E(C) = -\frac{1}{L(C)}\int_0^{L(C)} |\langle \nabla I(C(s)), \vec{n}\rangle|\, ds,$$

was considered. For better regularization, Fua and Leclerc [27] also proposed to couple their measure (6), with

$$E_D(C) = \int_0^{L(C)} |C_{ss}|^2\, ds.$$

We refer to [24] for a detailed discussion on these models.

Zhu, Lee, and Yuille [75, 76] proposed a statistical framework for image segmentation they named *region competition*. It combines active contours with region growing techniques using statistical criteria. Let O_i, $i = 1, \ldots, M$, be the regions of the segmentation. The regions O_i are considered to be homogeneous and the intensity values are generated by a prespecified probability distribution which we assume to be

$$P(I(x)|(\mu_i, \sigma_i)) = \frac{1}{\sqrt{2\pi}\sigma_i}\exp(-\frac{(I-\mu_i)^2}{2\sigma_i^2},$$

where (μ_i, σ_i) denote the mean and variance of the image on the region O_i. Then, the authors [75, 76] proposed to minimize the MDL (minimum description length criterion)

$$E(\{O_i\}_{i=1}^M, \{(\mu_i, \sigma_i\}_{i=1}^M)$$
$$= \sum_{i=1}^{M}\left(\frac{\mu}{2}L(\partial O_i) - \int_{O_i} log\, P(I(x)|(\mu_i, \sigma_i))\, dx + \lambda\right)$$

where $\lambda, \mu > 0$. Here μ represents the code length for unit arc-length, and λ is the code length needed to describe the distribution and code system for region O_i, which is assumed to be similar for all regions.

In a series of papers [56–60], Paragios and Deriche proposed an extension of the work of Zhu and Yuille using active contours (as mentioned before, Chan and Vese [14] and Yezzi et al. [74] worked on related region models). They integrate boundary and region-based terms to create an energy for curves whose minima determines a partition of the image

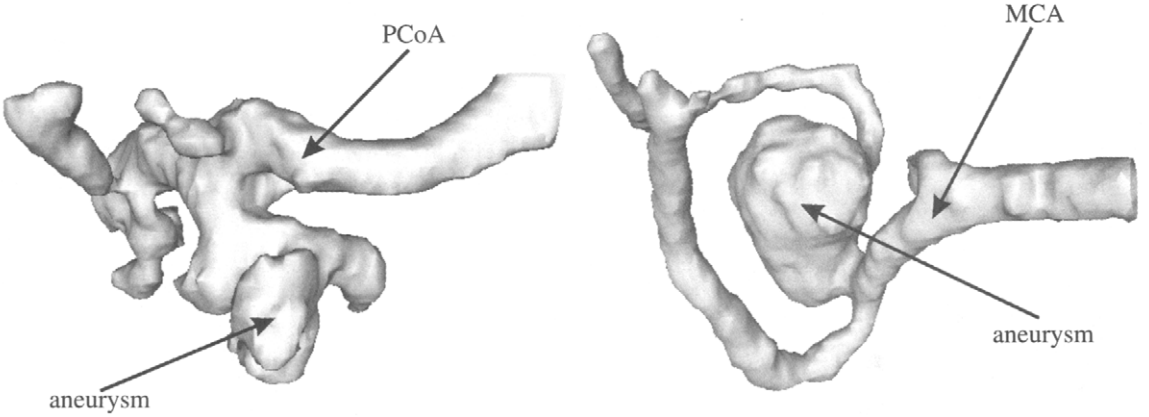

FIGURE 6 Representative examples of the three-dimensional models from computed tomography angiography data, obtained by the segmentation algorithm in Hernandez et al. [34]. PCoA, posterior communicating artery aneurysm; MCA, middle cerebral artery aneurysm.

which is optimal according to these criteria. Using different region descriptors they were able to introduce various applications: image segmentation [60], texture segmentation [58], and detection and tracking of moving objects [56, 57, 59]. A related level-set variational framework was used [63] for (supervised) image classification (with a predetermined number of regions). General functionals formed by addition of boundary and region based terms were considered [3, 36] and its Euler-Lagrange equation was computed using shape-derivative techniques.

9 Summary

In this chapter we reviewed some of the basic concepts of image segmentation via partial differential equations, and in particular, those related to the geodesic active contours. We are still far from the end of the road of deriving efficient segmentation techniques and low level vision operators. Often the segmentation tasks are difficult so that user interface and human interaction are still required [4, 19]. Good numeric schemes for so-called "solved" problems would probably change the way we process and analyze images in the future. Simple operations we take for granted, like edge detection and shape reconstruction, should be revisited and refined for the best possible solution. The exploration of the basic image analysis tools would improve our understanding of these processes and enable faster progress of feature extraction, learning, and classification procedures. Our philosophy of geometric variational interpretation for fundamental low-level image analysis operators seem to be one promising direction for improving existing tools and designing new ones.

Acknowledgments

The work here reported is partially supported by Departament d'Universitats, Recerca i Societat de la Informació de la Generalitat de Catalunya, Proyecto Nacional de Promocion General del Conocimiento project, reference BFM2003-02125, Office of Naval Research, National Science Foundation, Israeli Science Foundation, and National Institutes of Health. We thank all our collaborators in the theories of geodesic active contours, edge detection, and partial differential equations–based image processing.

References

[1] D. Adalsteinsson and J. A. Sethian. "A fast level set method for propagating interfaces," *J. Comp. Phys.* **118**, 269–277, (1995).

[2] G. Aubert and L. Blanc-Feraud, "Some remarks on the equivalence between 2D and 3D classical snakes and geodesic active contours," *Int. J. Comp. Vision* **34**, 19–28 (1999).

[3] G. Aubert, M. Barlaud, O. Faugeras, and S. Jehan-Besson, "Image segmentation using active contours: Calculus of variations or shape gradients?," *SIAM J. Appl. Math.* 63, 2128–2154 (2003).

[4] A. Bartesaghi, G. Sapiro, S. Lee, J. Lefman, and S. Subramaniam, "A new approach for 3D segmentation of cellular tomograms obtained using three-dimensional electron microscopy," *Proc. IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, Arlington, April 2004.

[5] M. Bertalmio, G. Sapiro, and G. Randall, "Region tracking on level-sets methods," *IEEE Trans. Medical Imaging* **18**, 448–451 (1999).

[6] M. Bertalmio, G. Sapiro, and G. Randall, "Morphing active contours," *IEEE Trans. Pattern Anal. Mach. Intell.* **22**, 733–738 (2000).

[7] J. Canny. "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.* **8**, 679–698 (1986).

[8] V. Caselles, F. Catte, T. Coll, and F. Dibos. "A geometric model for active contours." *Numerische Mathematik*, **66**, 1–31 (1993).

[9] V. Caselles, R. Kimmel, and G. Sapiro. "Geodesic active contours," in *Proceedings Int. Conf. on Computer Vision* (Boston, Massachusetts, June 1995), pp. 694–699.

[10] V. Caselles, R. Kimmel, and G. Sapiro. "Geodesic active contours," *Int. J. Comput. Vis.* **22**, 61–79 (1997).

[11] V. Caselles, R. Kimmel, G. Sapiro, and C. Sbert. "Minimal surfaces: A three dimensional segmentation approach," EE PUB No. 973, Technion–Israel Institute of Technology, Israel, June 1995.

[12] V. Caselles, R. Kimmel, G. Sapiro, and C. Sbert. "Minimal surfaces: A geometric three dimensional segmentation approach," *Numerische Mathematik* **77**, 423–425 (1997).

[13] V. Caselles, R. Kimmel, G. Sapiro, and C. Sbert. "Minimal surfaces based object segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.* **19**, 394–398 (1997).

[14] T. Chan and L. Vese. "An active contour model without edges," in *Scale-Space Theories in Computer Vision* Springer Verlag, Berlin, Heidelberg, Germany, 1999, pp. 141–151.

[15] C. S. Chiang, C. M. Hoffmann, and R. E. Lync. "How to compute offsets without self-intersection," in *Proc. SPIE* **1620**, 76 (1992).

[16] D. L. Chopp. Computing minimal surfaces via level set curvature flow. Ph.D Thesis, Lawrence Berkeley Lab. and Dep. of Math. LBL-30685, Uni. of CA. Berkeley, May 1991.

[17] D. L. Chopp. "Computing minimal surfaces via level set curvature flow," *J. Comput. Physics* **106**, 77–91 (1993).

[18] L. D. Cohen. "On active contour models and balloons," *CVGIP: Image Understanding* **53**, 211–218 (1991).

[19] L. D. Cohen, and R. Kimmel, "Global minimum for active contours models: A minimal path approach," *Int. J. Comput. Vis.* **24**, 57–78 (1997).

[20] D. Cremers, F. Tischhäuser, J. Weickert, C. Schnörr, "Diffusion snakes: introducing statistical shape knowledge into the Mumford–Shah functional," *Int. J. Comput. Vis.* **50**, 295–313 (2002).

[21] P. Danielsson. "Euclidean distance mapping," *Comput. Graph. Image Process.* **14**, 227–248 (1980).

[22] A. Dervieux and F. Thomasset, "A finite element method for the simulation of Rayleigh-Taylor instability," *Lecture Notes in Mathematics* **771**, 145–158 (1979).

[23] A. Dervieux and F. Thomasset, "Multifluid incompressible flows by a finite element method," *Lecture Notes in Physics* **11**, 158–163 (1981).

[24] A. Desolneux, L. Moisan, and J.M. Morel, "Variational snake theory," in *Level Set Methods and Their Applications in Computer Vision* (Springer Verlag, New York, 2003).

[25] M. P. Do Carmo, *Differential Geometry of Curves and Surfaces* (Prentice-Hall, New Jersey, 1976).

[26] O. D. Faugeras and R. Keriven, "Variational principles, surface evolution, PDE's, level-set methods, and the stereo problem," *IEEE Trans. Image Process.* **7**, 336–344 (1998).

[27] P. Fua and Y. G. Leclerc. "Model driven edge detection," *Mach. Vis. Appl.* **3**, 45–56 (1990).

[28] R. Goldenberg, R. Kimmel, E. Rivlin, and M. Rudzsky, "Fast geodesic active contours," *IEEE Tran. Image Process.* **10**, 1467–1475 (2001).

[29] R. Goldenberg, R. K. E. Rivlin, and M. Rudzsky. "Cortex segmentation — a fast variational geometric approach," in *IEEE Workshop on Variational and Level Set Methods in Computer Vision* (Vancouver, Canada, July 2001).

[30] F. Guichard and J. M. Morel, *Introduction to Partial Differential Equations in Image Processing, Tutorial Notes (IEEE Int. Conf. Image Proc.*, Washington, DC, October 1995).

[31] R. Haralick. "Digital step edges from zero crossing of second directional derivatives," *IEEE Trans. Pattern Anal. Mach. Intell.* **6**, 58–68 (1984).

[32] J. Helmsen, E. G. Puckett, P. Collela, and M. Dorr, "Two new methods for simulating photolithography development in 3D," *Proc. SPIE Microlithography* **IX**, 253 (1996).

[33] M. Hernáandez Giménez, R. Barrena, G. Hernandez, G. Sapiro, and A. Frangi Caregnato, "Pre-clinical evaluation of implicit deformable models for three-dimensional segmentation of brain aneurysms from CTA images," *Medical Imaging 2003: Image Processing Proc. of SPIE* **5032**, 1264–1127 (2003).

[34] M. Hernández Giménez, A. Frangi Caregnato, and G. Sapiro, "3D segmentation of brain aneurysms in CTA using non-parametric region-based information and implicit deformable models: Method and evaluation," *Medical Image Computing and Computer-Assisted Intervention (MICCAI 2003)* 594–602 (2003).

[35] M. Holtzman-Gazit, D. Goldsher, and R. Kimmel, "Hierarchical segmentation of thin structures in volumetric medical images," *Medical Image Computing and Computer-Assisted Intervention (MICCAI 2003)* 562–569 (2003).

[36] S. Jehan-Besson, "Modèles de contours actifs basés régions pour la segmentation d'images et de vidéos," Ph.D. Thesis, Université de Nice - Sophia Antipolis, January 2003.

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

[38] J. Kim, J. W. Fisher III, M. Cetin, A. Yezzi, and A. S. Willsky, "Incorporating complex statistical information in active contour-based image segmentation," *Proc. IEEE International Conference on Image Processing*, September 2003.

[39] R. Kimmel. "Fast edge integration," in *Level Set Methods and Their Applications in Computer Vision*, Paragios and Osher, eds. (Springer Verlag, NY, 2003), Chap. 3.

[40] R. Kimmel, *Numerical Geometry of Images: Theory, Algorithms, and Applications* (Springer-Verlag, New York, 2004).

[41] R. Kimmel and A. M. Bruckstein. "On regularized Laplacian zero crossings and other optimal edge integrators," *Int. J. Comput. Vis.* **53**, 225–243 (2003).

[42] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi, "Gradient flows and geometric active contour models," *Proc. Int. Conf. Comp. Vision '95*, 810–815, Cambridge, June 1995.

[43] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi, "Conformal curvature flows: from phase transitions to active vision," *Arch. Rat. Mech. Anal.* **134**, 275–301 (1996).

[44] M. Leventon, E. Grimson, and O. Faugeras, "Statistical shape influence in geodesic active contours," *Proc. IEEE CVPR* 316–322 (2002).

[45] L. M. Lorigo, O. Faugeras, W. E. L. Grimson, R. Keriven, and R. Kikinis, "Segmentation of bone in clinical knee MRI using texture-based geodesic active contours," *Proceedings Medical Image Computing and Computer-Assisted Intervention, MICCAI '98* (Cambridge, MA, Springer, 1998), pp. 1195–1204.

[46] T. Lu, P. Neittaanmaki, and X.-C. Tai. "A parallel splitting up method and its application to Navier-Stokes equations," *Appl. Math. Letters* **4**, 25–29 (1991).

[47] T. Lu, P. Neittaanmaki, and X.-C. Tai. "A parallel splitting up method for partial differential equations and its application to Navier-Stokes equations," *RAIRO Math. Model. Numer. Anal.* **26**, 673–708 (1992).

[48] R. Malladi, J. A. Sethian, and B. C. Vemuri. "Shape modeling with front propagation: A level set approach," *IEEE Trans. Pattern Anal. and Mach. Intell.* **17**, 158–175 (1995).

[49] D. Marr. *Vision* (Freeman, San Francisco, 1982).

[50] D. Marr and E. Hildreth. "Theory of edge detection," *Proc. Royal Society London B*, **207**, 187–217 (1980).

[51] D. Mumford and J. Shah. "Boundary detection by minimizing functionals," in *Proceedings of IEEE CVPR, Computer Vision and Pattern Recognition* (San Francisco, 1985), pp. 22–26.

[52] S. J. Osher and R. Fedkiw, *Level Set Methods and Dynamic Implicit Surfaces IEEE*, (Springer-Verlag, New York, 2003).

[53] S. J. Osher and N. Paragios, *Geometric Level Set Methods in Imaging, Vision, and Graphics* (Springer-Verlag, New York, 2003).

[54] S. J. Osher and J. A. Sethian. "Fronts propagating with curvature dependent speed: Algorithms based on Hamilton-Jacobi formulations," *J. Comp. Phys.* **79**, 12–49 (1988).

[55] N. Paragios, "Shape-based segmentation and tracking in cardiac image analysis," *IEEE Trans. Med. Imaging*, to appear (2003).

[56] N. Paragios and R. Deriche, "Geodesic active regions for tracking," *European Symposium on Computer Vision and Mobile Robotics CVMR'98* (Santorini, Greece, 1998).

[57] N. Paragios and R. Deriche, "Geodesic active regions for motion estimation and tracking," *Technical Report - INRIA - Sophia Antipolis* **3631** (1999).

[58] N. Paragios and R. Deriche, "Geodesic active regions for supervised texture segmentation," *Proc. International Conference Computer Vision* (Corfu, Greece, September 1999).

[59] N. Paragios and R. Deriche. "Geodesic active contours and level sets for the detection and tracking of moving objects," *IEEE Trans. Pattern Anal. Mach. Intell.* **22**, 266–280, (2000).

[60] N. Paragios and R. Deriche. "Coupled geodesic active regions for image segmentation: A level set approach," *European Conference in Computer Vision* (Dublin, Ireland, June 2000).

[61] D. Peng, B. Merriman, S. Osher, H. Zhao, and M. Kang, "A PDE-based fast local level-set method," *J. Computat. Phys.* **155**, 410–438 (1999).

[62] M. Rousson and N. Paragios, "Shape priors for level-set representations," *Proc. ECCV* (Denmark, 2002), pp. 78–93.

[63] C. Samson, L. Blanc-Feraud, G. Aubert, and J. Zerubia, "A level set model for image classification," *International Conference on Scale-Space Theories in Computer Vision* (1999), pp. 306–319.

[64] G. Sapiro, "Color snakes," *Comput. Vis. Image Underst.* **68**, 247–253 (1997).

[65] G. Sapiro, *Geometric Partial Differential Equations and Image Processing* (Cambridge University Press, January 2001).

[66] J. Sethian, "Fast marching level set methods for three-dimensional photolithography development," *Proc. SPIE International Symposium on Microlithography* (Santa Clara, California, March, 1996).

[67] J. A. Sethian, "A fast marching level-set method for monotonically advancing fronts," *Proc. Nat. Acad. Sci.* **93**, 1591–1595, (1996).

[68] J. A. Sethian. *Level Set Methods: Evolving Interfaces in Geometry, Fluid Mechanics, Computer Vision and Materials Sciences* (Cambridge University Press, 1996).

[69] A. Tsai, A. Yezzi, and A. S. Willsky, "Curve evolution implementation of the Mumford-Shah functional for image segmentation, denoising, interpolation, and magnification," *IEEE Trans. Image Process.* **10**, 1169–1186 (2001).

[70] J. N. Tsitsiklis, "Efficient algorithms for globally optimal trajectories," *IEEE Trans. Auto. Cont.* **40**, 1528–1538 (1995).

[71] A. Vasilevskiy and K. Siddiqi, "Flux maximizing geometric flows," in *Proceedings International Conference on Computer Vision* (Vancouver, Canada, July 2001), pp. 149–154.

[72] J. Weickert, B. M. ter Haar Romeny, and M. A. Viergever, "Efficient and reliable scheme for nonlinear diffusion filtering," *IEEE Trans. Image Process.* **7**, 398–410 (1998).

[73] R. T. Whitaker, "Algorithms for implicit deformable models," *Proc. ICCV'95* (Cambridge, June 1995), pp. 822–827.

[74] A. Yezzi, S. Kichenassamy, P. Olver, and A. Tannenbaum, "Geometric active contours for segmentation of medical imagery," *IEEE Trans. Med. Imaging* **16**, 199-210 (1997).

[75] S. Zhu, T. Lee, and A. Yuille. "Region competition: Unifying snakes, region growing, energy/bayes/mdl for multi-band image segmentation," in *Proceedings International Conference on Computer Vision* (Cambridge, June 1995), pp. 416–423.

[76] S. Zhu, and A. Yuille, "Region competition: Unifying snakes, region growing, and bayes/mdl for multi-band image segmentation," in *IEEE Trans. Patt. Anal. Mach. Intell.* **18**, 884–900 (1996).