Geometric Active Contours for Image Segmentation

Vicent Caselles, Ron Kimmel, and Guillermo Sapiro
August 9, 2004

Abstract

In this chapter we review the geometric active contour and related models for image segmentation. We present the underlying geometric energy formulations, their corresponding level-sets based gradient descent flows, and efficient numerical schemes.

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 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 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 partial differential equations (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 partial differential equations, differential geometry, etc. The computer image processing algorithms are actually the numerical implementation of the resulting equations. Related examples of this general approach can be found in chapters 3.3, 3.6, 4.15-4.17, while 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, 11, 10, 12, 13]. It was motivated by the snake model in [37], and the geometric-level sets curve evolution models in [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 ITK project for medical image analysis (www.itk.org). A large number of academic reports and innovative ideas flourished from the philosophy that weighted arclength parametrization is an appropriate measure for boundary detection in images. The work in [9] was of course not the first to make this observation, and among the first to stress the role
of geometry and Euclidean arclength in image segmentation were probably Mumford and Shah in [51]. Yet, the geodesic active contour model provides one of the first clear formulations and efficient numerical implementations for such geometric models.

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

In this chapter we present our, obviously biased yet hopefully balanced, view on the area of PDEs based segmentation methods, and conclude with recent results we have found to be relevant. The presentation partially follows that of [39]. Section 2 introduces some of the mathematical 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 [28], and present an efficient numerical scheme that couples an alternating direction implicit multiplicative scheme, with narrow band [1, 17, 61], and re-distancing 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 look into the books [30, 40, 53, 65, 68] for additional related contributions in image processing and analysis.

2 Mathematical 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, see for example [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 arclength parameter, and whose normal is defined by \( \vec{n}(s) = \{-y_b(s), y_a(s)\} \). 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 (non-geometric) 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 arclength \( s \) 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 [25], \( \kappa \) to be the curvature of the curve \( C \), and the curvature vector \( \kappa \mathbf{n} = C_{ss} \). Let us recall that a Jordan curve is a plane curve which 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 \( \Omega_C \) to be the domain inside the curve \( C \), see Figure 1. We also assume w.l.o.g. that the curves are counterclockwise oriented, so that \( \mathbf{n} \) is the inner unit normal vector.

\[\text{Figure 1: A closed curve } C, \text{ with } C_s \text{ the unit tangent, } \kappa \mathbf{n} = C_{ss} \text{ the curvature vector, and } \Omega_C \text{ the region inside the curve.}\]

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 \( g(\cdot) \) 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 arclength parameterization), 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,

\[
E(C) = \iint_{\Omega_C} f(x, y) dx dy,
\]

where as before, \( \Omega_C \) is the region inside the Jordan curve \( C \). Under general assumptions, this is a measure of the weighted area of \( \Omega_C \), 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 in [41] and proved to be optimal under general conditions in [24].

The importance of orientation information in a variational setting for delicate segmentation tasks was recently also considered in [71], where the authors proposed alignment with a general vector field as a segmentation criterion of complicated closed thin structures in 3D medical images. In [41] it was shown 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 variational principle that combines the alignment term with a topological 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 classical snakes [37], followed by non-variational 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 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 [10] model 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 \( g(\cdot) \), 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 classical snakes ([37]) using least action principles in physics (see also [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 arclength of the curve.
3.2 Alignment Term

As pointed out in [41], \( g(\cdot) \) can be a function not only of the image gradient, but also of its direction. Consider the geometric functional

\[
E_A(C) = -\int_0^L \langle \nabla I(x(s), y(s)), \bar{n}(s) \rangle ds,
\]
or in its more ‘robust’ form

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

where the absolute value of the inner product between the image gradient and the curve normal is our alignment measure, see Figure 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 maximize the alignment. This new energy can be combined with the one in Section 3.1 as well as the ones presented below. 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, e.g., [6, 28, 56].

![Figure 2: The curve C, its normal \( \bar{n} \) at a specific point, and the image gradient \( \nabla I \) at that point. The alignment term integrates the projection of \( \nabla I \) on the normal along the curve.](image)

3.3 Weighted Region

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

\[
E_W(C) = \int \int_{\Omega_C} f(x, y) dxdy,
\]
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 $\Omega_C$ that we also denote by $|\Omega_C|$. Other approaches try to find uniform regions inside $C$, see for example [56, 57, 58, 59, 69, 75] and the description in next subsection.

### 3.4 Minimal Variance

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

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

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) 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, as well as for 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.

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

One could consider more generic region based measures like

$$E(C) = \int_{\Omega_C} \|T(I(x,y)) - \vec{c}_1\| dx dy + \int_{\Omega \setminus \Omega_C} \|T(I(x,y)) - \vec{c}_2\| dx dy,$$

where $T$ is a general transformation of the image, the norm $\| \cdot \|$ can be chosen according to the problem in hand, and $\vec{c}_1$ and $\vec{c}_2$ are vectors of possible parameters. One such example is the robust measure

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

### 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 [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

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

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 $\Omega_C$ that we also denote by $|\Omega_C|$. Other approaches try to find uniform regions inside $C$, see for example [56, 57, 58, 59, 69, 75] and the description in next subsection.

### 3.4 Minimal Variance

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

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

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) 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, as well as for 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.

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

One could consider more generic region based measures like

$$E(C) = \int_{\Omega_C} \|T(I(x,y)) - \vec{c}_1\| dx dy + \int_{\Omega \setminus \Omega_C} \|T(I(x,y)) - \vec{c}_2\| dx dy,$$

where $T$ is a general transformation of the image, the norm $\| \cdot \|$ can be chosen according to the problem in hand, and $\vec{c}_1$ and $\vec{c}_2$ are vectors of possible parameters. One such example is the robust measure

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

### 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 [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

$$E(C) = \int C L(C_p,C) dp,$$
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 $\frac{\delta E(C)}{\delta C} = 0$. A dynamic process known as gradient descent, that takes an arbitrary curve towards a minimum of $E(C)$, is given by the curve evolution equation

\[
\frac{\partial C}{\partial t} = -\frac{\delta E(C)}{\delta C},
\]

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 in [5, 27, 41]. We now present a few examples of this gradient descent flow.

Unless stated otherwise, the arclength 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

\[
E_{GAC}(C) = \int_C g(C(s)) ds,
\]

for which the first variation is given by

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

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

\[
\frac{d}{d\epsilon}|C_s + \epsilon \eta_s|_{\epsilon=0} = \langle C_s, \eta_s\rangle,
\]

we compute

\[
\frac{d}{d\epsilon} E_{GAC}(C + \epsilon \eta)|_{\epsilon=0} = \frac{d}{d\epsilon}|_{\epsilon=0} \int g(C(s) + \epsilon \eta(s))|C_s + \epsilon \eta_s| ds
\]

\[
= \int \langle \nabla g(C(s)), \eta(s)\rangle ds + \int g(C(s)) \langle C_s, \eta_s\rangle ds
\]

\[
= \int \langle \nabla g(C(s)), \eta(s)\rangle ds - \int \langle \nabla g(C(s)), C_s\rangle \langle C_s, \eta\rangle ds
\]

\[
+ \int g(C(s)) \langle C_{ss}, \eta\rangle ds
\]

\[
= \int \langle \nabla g(C(s)), C_s^\perp\rangle \langle C_s^\perp, \eta\rangle ds - \int g(C(s)) \langle C_{ss}, \eta\rangle ds
\]

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

$$\frac{\delta E(C)}{\delta C} = F'(\langle \vec{V}(C), \vec{n} \rangle)(\text{div}V)(C)\vec{n} - (F'(\langle \vec{V}(C), \vec{n} \rangle))sR^tV(C)$$

$$-\{G(\langle \vec{V}(C), \vec{n} \rangle))sC_s + G(\langle \vec{V}(C), \vec{n} \rangle)k\vec{n}\}. \quad (3)$$

Typical examples of $F$ are $F(z) = -|z|^p$, $p \geq 1$. Observe that when $1 \leq p < 2$, then $F$ is of class $C^2$ on $\mathbb{R} \setminus \{0\}$, but with proper adaptation 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 [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 $\epsilon$ and $\eta$ are not zero, the curve $C + \epsilon \eta$ is not parameterized by arc-length anymore. Using (2) and

$$\frac{d}{d\epsilon} \frac{C_s}{C_s + \epsilon \eta} \big|_{\epsilon = 0} = \eta_s = \langle C_s + \eta_s \rangle$$

we compute

$$\frac{d}{d\epsilon} E(C + \epsilon \eta) \big|_{\epsilon = 0} = \frac{d}{d\epsilon} \left. \oint_C F(\langle \vec{V}(C + \epsilon \eta), \frac{C_s + \epsilon \eta}{C_s + \epsilon \eta} \rangle) ds + \epsilon \eta \right|_{\epsilon = 0}$$

$$= \oint_C F'(\langle \vec{V}(C) + \epsilon \eta \rangle)(\text{div}\vec{V}(C))\eta + F'(\langle \vec{V}(C) + \epsilon \eta \rangle)\langle \vec{V}(C) + \epsilon \eta, \eta \rangle ds$$

$$- \oint_C F'(\langle \vec{V}(C) + \epsilon \eta \rangle)\langle \vec{V}(C) + \epsilon \eta, \eta \rangle ds + F'(\langle \vec{V}(C), \eta \rangle)\langle \vec{V}(C), \eta \rangle ds.$$

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

$$\frac{d}{d\epsilon} E(C + \epsilon \eta) \big|_{\epsilon = 0} = \oint_C F'(\langle \vec{V}(C), \eta \rangle)\langle \vec{V}(C) + \epsilon \eta, \eta \rangle ds$$

$$- \oint_C F'(\langle \vec{V}(C), \eta \rangle)\langle R^tV(C), \eta \rangle + F'(\langle \vec{V}(C), \eta \rangle)\langle R^t \text{div} \vec{V}(C) \rangle ds$$

$$- \oint_C G(\langle \vec{V}(C), \eta \rangle)s\langle \eta, C_s \rangle + G(\langle \vec{V}(C), \eta \rangle)s \langle \eta, C_s \rangle ds$$

$$= \oint_C F'(\langle \vec{V}(C), \eta \rangle)\langle \text{div} \vec{V}(C), \eta \rangle - \langle R^t \text{div} \vec{V}(C) \rangle ds$$

$$- \oint_F F'(\langle \vec{V}(C), \eta \rangle)\langle R^tV(C), \eta \rangle$$

$$- \oint_C G(\langle \vec{V}(C), \eta \rangle)s \langle C_s, \eta \rangle + G(\langle \vec{V}(C), \eta \rangle)s \langle \eta, \eta \rangle ds.$$
Thus, the the first variation of $E_A(C)$ is
\[
\frac{\delta E(C)}{\delta C} = F'((\hat{V}(C), \vec{n})) [D\hat{V}(C)\hat{q}(\vec{n}) - R^tD\hat{V}(C)(C_s)] \\
- (F'((\hat{V}(C), \vec{n})), R^tV(C) \\
- [(G((\hat{V}(C), \vec{n})� C_s + G((\hat{V}(C), \vec{n}))k\vec{n}].
\]

Now, observe that $D\hat{V}(C)\hat{q}(\vec{n}) - R^tD\hat{V}(C)(C_s) = D\hat{V}(C)^trC_s - R^tD\hat{V}(C)(C_s) = (\text{div}V)(C)\vec{n}$, which gives (3).

Notice that the second term in the right hand side of (3) can be further expanded as ([41])
\[
(F'((\hat{V}(C), \vec{n})), R^tV(C) = (F'((\hat{V}(C), \vec{n})), [(V(C), R\vec{n})\vec{n} + (V(C), R^tC_s)C_s] \\
= F''(\hat{V}(C), \vec{n}))[(V(C), \vec{n})(V(C), C_s) - (V(C), C_s)^2]|\vec{n} + \text{tangential components}.
\]

An important example of (3) corresponds to the choice $\hat{V} = \nabla I$, for which the formula for the first variation is
\[
\frac{\delta E(C)}{\delta C} = F'(\nabla I(C), \vec{n}))[\nabla I(C)\vec{n}] - (F'(\nabla I(C), \vec{n}))R^t\nabla I(C) \\
- [(G(\nabla I(C), \vec{n}))C_s + G(\nabla I(C), \vec{n}))k\vec{n}]
\]

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

The robust alignment term
\[E_{AR}(C) = -\int_C |(\hat{V}, \vec{n})| ds,\]
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
\[
\frac{\delta E_{AR}(C)}{\delta C} = -\text{sign}(\nabla I(C), \vec{n}(s))\Delta I(C)\vec{n} - (\text{sign}(\nabla I(C), \vec{n}(s))C_s + (\nabla I(C), \vec{n})k\vec{n}]
\]

If $\langle \nabla I(C), \vec{n}(s) \rangle$ has constant sign, then the second term vanishes and we obtain
\[
\frac{\delta E_{AR}(C)}{\delta C} = -\text{sign}(\nabla I(C), \vec{n}(s))\Delta I(C)\vec{n}. \tag{4}
\]

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

$$E(C) = \iint_{\Omega_C} f(x,y) dx dy = \iint_{\Omega_C} (Q_x - P_y) dx dy$$

$$= \int_C (P dx + Q dy) = \int_C (P x + Q y) ds$$

$$= \int_C \langle -Q, P \rangle \vec{n} ds,$$

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

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

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 [21, 15] 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

$$\frac{\delta E_{MV}}{\delta C} = -(c_2 - c_1) \left( I - \frac{c_1 + c_2}{2} \right) \vec{n}$$

$$\frac{\delta E_{MV}}{\delta c_1} = \iint_{\Omega_C} (c_1 - I) dx dy$$

$$\frac{\delta E_{MV}}{\delta c_2} = \iint_{\Omega \setminus \Omega_C} (c_2 - I) dx dy.$$

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

$$\frac{\delta E_{MV}}{\delta C} = -\frac{1}{2} \left( (I - c_1)^2 - (I - c_2)^2 \right) \vec{n}$$

$$= -\frac{1}{2} \left( I^2 - 2 I c_1 + c_1^2 - I^2 + 2 I c_2 - c_2^2 \right) \vec{n}$$

$$= - \left( (c_2 - c_1) I - \frac{(c_1 + c_2)(c_2 - c_1)}{2} \right) \vec{n}$$

$$= -(c_2 - c_1) \left( I - \frac{c_1 + c_2}{2} \right) \vec{n}.$$

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

\[ E_{RMV}(C,c_1,c_2) = \int \int_{\Omega_C} |I - c_1| dx dy + \int \int_{\Omega \setminus \Omega_C} |I - c_2| dx dy, \]

for which the first variation is

\[ \frac{\delta E_{RMV}}{\delta C} = - (|I - c_1| - |I - c_2|) \hat{n}, \]
\[ \frac{\delta E_{RMV}}{\delta c_1} = \int \int_{\Omega_C} \text{sign}(c_1 - I) dx dy, \]
\[ \frac{\delta E_{RMV}}{\delta c_2} = \int \int_{\Omega \setminus \Omega_C} \text{sign}(c_2 - I) dx dy, \]

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 \( \Omega_C \) that splits its area into two equal parts. From obvious reasons we refer to this value as the median of \( I \) in \( \Omega_C \), 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 over-segment such events or require large regularization that could over-smooth 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 [35].

4.1 Another Intermezzo

We reviewed a number of possible curve flows to deform a given initial curve towards 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 employed 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 $C_t = \gamma \bar{n}$, 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 $\phi(t, x, y)$, 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 $\bar{n} = -\nabla \phi / |\nabla \phi|$, hence

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

More on the rigorous equivalence between both flows and the particular relevance for the geodesic active contours can be found in [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 classical 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, while 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

$$\arg \min_{C,\ell_1,\ell_2} E(C, \ell_1, \ell_2),$$

for the functional

$$E(C, \ell_1, \ell_2) = E_{AR}(C, \ell_1, \ell_2) + \alpha E_{GAC}(C) + \beta E_{MV}(C)$$

$$= - \int_C |\langle \nabla I, \vec{n} \rangle| ds + \alpha \int_C g(C(s)) ds$$

$$+ \frac{\beta}{2} \left( \int_{\Omega \setminus \Omega_C} (I - \ell_1)^2 dx dy + \int_{\Omega \cap \Omega_C} (I - \ell_2)^2 dx dy \right),$$

where $\alpha$ and $\beta$ are positive constants, and, in practice, $\alpha$ 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

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

$$c_1 = \frac{1}{|\Omega_C|} \int_{\Omega_C} I(x, y) dx dy,$$

$$c_2 = \frac{1}{|\Omega \setminus \Omega_C|} \int_{\Omega \setminus \Omega_C} I(x, y) dx dy,$$

where $|\Omega_C|$ denotes the area of the regions $\Omega_C$. 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

$$\phi_t = \left[ -\text{sign}(\langle \nabla \phi, \nabla I \rangle) \Delta I + \text{div} \left( g(x, y) \frac{\nabla \phi}{|\nabla \phi|} \right) \right] + \beta (c_2 - c_1) \left( I - \frac{c_1 + c_2}{2} \right) |\nabla \phi|. $$
Efficient solutions for this equation can use either AOS [46, 47, 72], ADI, LOD methods [39], coupled with a narrow band approach [1, 16, 61], see [28] for the geodesic active contour 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.

<table>
<thead>
<tr>
<th>Measure</th>
<th>$E(C)$</th>
<th>$\delta E/\delta C$</th>
<th>$\delta E/\delta C$ in level set form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weighted Area</td>
<td>$\iint \alpha f(x, y) dx dy \sum_{\alpha} f(x, y) \hat{n}$</td>
<td>$f(x, y) \nabla \phi \cdot \hat{n}$</td>
<td>$(c_2 - c_1) \nabla \phi</td>
</tr>
<tr>
<td>Minimal Variance</td>
<td>$\iint (I - c_1)^2 f(x, y) dx dy + \sum_{\alpha} (I - c_2)^2 f(x, y) \hat{n}$</td>
<td>$(c_2 - c_1) \nabla \phi</td>
<td>\hat{n}$</td>
</tr>
<tr>
<td>GAC</td>
<td>$g(C(s)) ds \sum_{\alpha} (c_2 - c_1) \nabla \phi</td>
<td>\hat{n}$</td>
<td>$\text{div} \left( \frac{\nabla g}{</td>
</tr>
<tr>
<td>Alignment</td>
<td>$- \int (\nabla I, \hat{n}) ds \sum_{\alpha} \frac{\text{sign}(\nabla I, \hat{n}) \Delta I \hat{n}}{-\text{sign}(\nabla I, \hat{n}) \Delta I \hat{n}}$</td>
<td>$\text{sign}(\nabla I, \hat{n}) \Delta I \hat{n}$</td>
<td>$-\text{sign}(\nabla I, \hat{n}) \Delta I \hat{n}$</td>
</tr>
</tbody>
</table>

6 Efficient Numerical Schemes

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

Let us analyze the following level set formulation,

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

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

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 $\phi$ is thereby

$$
\phi_t = \text{adiv} \left( g(x, y) \frac{\nabla \phi}{|\nabla \phi|} \right) + \eta(\phi, \nabla I) + \alpha \frac{\partial}{\partial x} \left( g(x, y) \frac{\partial \phi}{\partial x} \right) + \alpha \frac{\partial}{\partial y} \left( g(x, y) \frac{\partial \phi}{\partial y} \right) + \eta(\phi, \nabla I).
$$

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 in order 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.

1As it evolves, $\phi$ immediately departs from being a distance function of its zero set. We can ignore this second order effect, while re-distancing every numerical time step.
Denote the operators

\[ A_1 = \frac{\partial}{\partial x} g(x, y) \frac{\partial}{\partial x} , \quad A_2 = \frac{\partial}{\partial y} g(x, y) \frac{\partial}{\partial y} . \]

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

\[
\begin{align*}
\phi^{n+1} &= \phi^n + \tau (\alpha (A_1 + A_2) \phi^n + \tau \eta(\phi^n, \nabla I)) \\
&= (I + \tau \alpha (A_1 + A_2)) \phi^n + \tau \eta(\phi^n, \nabla I),
\end{align*}
\]

where, after sampling the \( x, y \) plane, \( I \) is the identity matrix and \( I \) is our input image. The operators \( A_l \) become matrix operators, and \( \phi^n \) 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 numerical 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

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

where \( N(i) \) 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

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

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

\[ \phi^{n+1} = \prod_{l=1}^2 \left( I - \tau \alpha A_l \right)^{-1} \left( \phi^n + \tau \eta(\phi^n, \nabla I) \right). \]

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

\[
(I - \tau A_1)^{-1}(I - \tau A_2)^{-1}(\phi + \tau \eta) = \left( I - \tau A_1 - \tau A_2 + O(\tau^2) \right)^{-1}(\phi + \tau \eta) = \phi + \tau (A_1 + A_2) \phi + \tau \eta + O(\tau^2),
\]

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 \( I - \tau A_l \) 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 [40].

7 Examples

We now present a number of examples of the above describe 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 frame shows the final contour in black painted on the original image in which the dynamic range is mapped into a sub-interval of the original one.

![Figure 4](image1.png)

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 of the authors of [41].

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.

![Figure 5](image2.png)

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 the authors of [41].

Lastly, Figure 6 shows 3D examples for brain aneurysms from CTA images, following [33, 34]. For this difficult data, some of the above mentioned energies have been extended to 3D and combined with additional shape information.
Figure 6: Representative examples of the 3D models from CTA data, obtained by the segmentation algorithm in [34]. Posterior Communicating Artery and Middle Cerebral Artery aneurysms, respectively.

8 Additional Comments on Related Developments

So far we 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 arclength of the curve. Trying to integrate the photometric and geometric constraints, Fua and Leclerc [27] proposed to minimize the functional

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

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 [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

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

17
We refer to [24] for a detailed discussion on these models.

In [75, 76], Zhu, Lee, and Yuille 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 pre-specified probability distribution which we assume to be

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

where $(\mu_i, \sigma_i)$ 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)

$$E(\{O_i\}_i^{M}, \{\mu_i, \sigma_i\}_{i=1}^{M}) = \sum_{i=1}^{M} \left(\frac{\mu}{2} L(\partial O_i) - \int_{O_i} \log P(I(x)|(\mu_i, \sigma_i)) \, dx + \lambda\right)$$ (8)

where $\lambda, \mu > 0$. Here $\mu$ represents the code length for unit arclength, and $\lambda$ 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, 57, 58, 59, 60], Paragios and Deriche proposed an extension of the work of Zhu and Yuille using active contours (as mentioned before, Chan and Vese as well as Yezzi *et al.* 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 which is optimal according to these criteria. Using different region descriptors they were able to introduce various applications: image segmentation in [60], texture segmentation in [58], and detection and tracking of moving objects [56, 57, 59]. A related level set variational framework was used in [63] for (supervised) image classification (with a predetermined number of regions). General functionals formed by addition of boundary and region based terms were considered in [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 numerical 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, PNPGC project, reference BFM2003-02125, ONR,
NSF, ISF, and NIH. We thank all our collaborators in the theories of geodesic active contours, edge detection, and PDEs based image processing.

References


