

Chapter 1

Variational methods in Image Processing

A track should be constructed to connect some point in space A to a lower altitude point B . Which form the track should take if we wish that a ball released at A reaches B in the shortest time? The curve known as brachistochrone or tautochrone is the answer of this puzzle solved by Jean Bernoulli and a classical problem of the *calculus of variations*.

a function whose variable is itself a function

The main object of calculus of variations are called *functionals* or *energies*, and a simple way to describe it is as a function which the variable is itself function. Minimizing functionals is a more intricate problem than minimizing a normal function, as the variable in a functional has infinite dimension. Nonetheless, by means of the so called *variations*, one can model infinitely small variations in the functional and do a rigorous analysis of its extremum, which the main tool is the *Euler-Lagrange* equation.

usual

the main tool of which

The calculus of variations founded in image processing a fertile field of applications, as image themselves can be seen as a function, and image processing tasks can be modeled as being the results of a functional minimization. In this chapter we present some popular variational techniques to approach image processing tasks, with a particular focus to image segmentation.

some
on

1.1 Inverse problems in imaging

An archaeological museum decided to digitize some of its collections and make them available for digital visits over the internet. The chosen method of digitization consists into take a set of pictures for each object, in different camera positions, execute a

whose

Inverse problem	Forward problem
Projection: Compute vector $v \in \mathbb{R}^3$ which projection is $P(v) \in \mathbb{R}^2$	Compute the projection $P(v) \in \mathbb{R}^2$ of vector $v \in \mathbb{R}^3$
Parameters inference: Given a set of observations Γ , infer the parameters (μ, σ) of the Gaussian distribution that describes Γ	Given a random variable X following a Gaussian distribution with parameters $(\mu = 0, \sigma = 1)$, compute the probability $P(X \leq 0.42)$
Image denoising: Given noisy image $\tilde{\mathbf{I}}$, compute the original image \mathbf{I} , i.e., the image without noise	Add some random noise to a given image \mathbf{I} to produce noisy image $\tilde{\mathbf{I}}$
Image inpainting: Given image $\tilde{\mathbf{I}}$ with a missing patch, reconstruct the removed patch	Remove a patch from image \mathbf{I}
Image segmentation: Given image I , find the labeled partition \mathcal{I}	Given a labeled partition \mathcal{I} of some image I , assemble the pieces to create image I

Table 1.1: Examples of inverse problems and its direct versions. Inverse problems are characterized by uncertainty and parameter inference.

stereo algorithm to estimate point depths and *reconstruct* finally reconstruct the 3D object. The stereo and reconstruction are examples of *inverse problems* in imaging.

Usually, inverse problems are characterized by a degree of *uncertainty* or *lack of information*. The 2D pictures in the problem above miss depth information, that should be *inferred* by the stereo algorithm. On the other hand, if the shape geometry was known, e.g., the values of mean curvature were known for every infinitesimal point of the shape, then constructing a digital 3D representation would be a *forward problem*.

We can find examples of inverse problems in several branches of mathematics [Kir96], geophysics [Zhd15], natural language processing [SY05], astronomy [Luc94] and the list goes on. The image processing field itself is plenty of them [BB98]. In fact, a great part of real world applications consists into *infer* parameters of some model, i.e., an inverse problem. In Table 1.1 we list some examples of inverse problems and its corresponding forward version.

Another characteristic of inverse problems is that they are usually *ill-posed*. A problem is said to be ill-posed if at least one of the properties below are not respected

is
infering

1. A solution exists and it is unique;
2. The solution changes continuously with its parameters

In order to solve ill-posed problems one should include additional information, i.e., create assumptions over the properties of the **sought** solution. In the museum problem, for example, one may assume that missing patches of the reconstructed surface should be filled by patches of minimal area. The process of including additional information in ill-posed problems is called *regularization* and its goal is to better condition an ill-posed problem and in the best scenario, transform it into a well-posed one.

Next, we describe the image model used in this thesis and give a precise definition of the main image problems discussed further on. **Examples are given in ??**

Examples of what ?

Image model

For matters of simplicity, we limit our discussion to grayscale images, the concepts being **extendable** to multichannel images. It is convenient to have in mind two different representations of an image.

$$\begin{aligned} \text{Discrete: } & \mathbf{I} \in \mathbb{F}^{m \times n} \\ \text{Continuous: } & f_{\mathbf{I}} : \Omega \subset \mathbb{R}^2 \rightarrow [0, 1], \end{aligned}$$

where \mathbb{F} is a finite set. In this thesis, we define such set as

$$\mathbb{F} = \left\{ \frac{i}{255} \mid i \in \mathbb{N}, i \leq 255 \right\}. \quad (1.1)$$

The discrete representation is interpreted as a sampling of $m \times n$ elements (pixels) of the continuous representation $f_{\mathbf{I}}$.

Image denoising

Given an image $f_{\tilde{\mathbf{I}}}$ corrupted with some noise from an external source, *image denoising* consists in to find an estimation $f_{\hat{\mathbf{I}}}$ of the original image that respects some quality criteria, usually encoded by the minimum of a functional E .

Given $f_{\tilde{\mathbf{I}}}$, find estimation $f_{\hat{\mathbf{I}}}$ such that

$$f_{\hat{\mathbf{I}}} = \arg \min_f E(f, f_{\tilde{\mathbf{I}}})$$

Applications: **Restauration** of old pictures; enhancement of satellite images.

restoration

Image segmentation

in finding a

Given an image f_I , the *image segmentation* problem consists in to find partition \mathcal{I} of f_I such that each element of \mathcal{I} is identified with some desired property, usually encoded by the minimum of some functional E

Given $f_I : \Omega \rightarrow [0, 1]$ and a positive integer n , find partition $\mathcal{I}^* = \{\Omega_i \subset \Omega \mid i \leq n\}$ such that

$$\mathcal{I}^* = \arg \min_{\mathcal{I}} E(\mathcal{I}, f_I) \quad \text{subject to} \quad \begin{array}{l} \forall i \neq j : \Omega_i \cap \Omega_j = \emptyset \\ \bigcup_i^n \Omega_i = \Omega \end{array}$$

Applications: enhance blood vessels in angiograms; track roads in satellite images; identify objects in a scene.

Image inpainting

Given an image $f_{\tilde{I}}$ with a collection of missing patches \mathcal{P} , *image inpainting* consists in to create an image f_I with the reconstructed missing patches such that a quality criteria, encoded as the minimum of some functional, is respected

Given $f_{\tilde{I}}$ and missing patches \mathcal{P} , find image f_I such that

$$f_I = \arg \min_f E(f, f_{\tilde{I}}) \quad \text{subject to } f(\Omega \setminus \mathcal{P}) = f_{\tilde{I}}(\Omega \setminus \mathcal{P}).$$

Applications: removal of undesired objects in a scene; restauration of old pictures.

I would provide figures/images to illustrate the different applications (and examples of processing methods that provide reasonable results)

1.2 Bayesian rationale and total variation

As remarked in the previous section, inverse problems involve some level of uncertainty about the solution. In order to solve an ill-posed problem we need to regularize it by including additional information, or make assumptions.

otherwise said

The maximum a posteriori method was first introduced in the image processing community in the work of [GG84] and we are going to reproduce here the rational for image denoising. We make two assumptions

1. The noisy image \tilde{I} was obtained by addition of a normal Gaussian noise ($\mu = 0, \sigma = 1$) to the original image, i.e.,

$$\tilde{I} = I + N, \tag{A.1}$$



(a) Original image



(b) 10-partition as in [CCP08]



(c) Noisy image



(d) Image denoised with FISTA [BT09]

(e) Inpainting mask



(f) Inpainted image with [FFA15]

Figure 1.1: Imaging problems examples

Good examples : you must reference the figure in the text when you speak of applications

where \mathbf{N} is a $(m \times n)$ matrix of random variables $N_{i,j}$ and $Pr(N_{i,j} = n) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{n^2}{2}\right)$.

2. Given some function ρ , a candidate image estimation \mathbf{C} has probability

$$Pr(\mathbf{C}) = \exp(-\rho(\mathbf{C})). \quad \text{\textcolor{blue}{\exp}} \quad (\text{A.2})$$

We estimate the unknown original image \mathbf{I} as the image $\hat{\mathbf{I}}$ that is more likely to occur. Applying Bayes' theorem we obtain

$$\hat{\mathbf{I}} = \arg \max_{\mathbf{C}} Pr(\mathbf{C} | \tilde{\mathbf{I}}) = \arg \max_{\mathbf{C}} \frac{Pr(\tilde{\mathbf{I}} | \mathbf{C}) Pr(\mathbf{C})}{Pr(\tilde{\mathbf{I}})}. \quad (\text{1.4})$$

We have already all the elements to expand [Equation \(1.4\)](#). The probability of having the corrupted image $\tilde{\mathbf{I}}$ given a candidate image \mathbf{C} is derived from [Equation \(A.1\)](#), i.e.,

$$Pr(\tilde{\mathbf{I}} | \mathbf{C}) = Pr(\mathbf{N} = \tilde{\mathbf{I}} - \mathbf{C}) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\|\tilde{\mathbf{I}} - \mathbf{C}\|^2}{2}\right). \quad (\text{1.5})$$

The denominator term is computed as the joint probability

$$Pr(\tilde{\mathbf{I}}) = \sum_{\mathbf{J} \in \mathbb{F}^{m \times n}} Pr(\tilde{\mathbf{I}} | \mathbf{J}) Pr(\mathbf{J}) = \frac{1}{\sqrt{2\pi}} \sum_{\mathbf{J} \in \mathbb{F}^{m \times n}} \exp\left(-\frac{1}{2} \|\tilde{\mathbf{I}} - \mathbf{J}\|^2 - \rho(\mathbf{J})\right). \quad (\text{1.6})$$

Substituting [Equations \(1.5\)](#) and [\(1.6\)](#) in [Equation \(1.4\)](#) we obtain

$$\hat{\mathbf{I}} = \arg \max_{\mathbf{C}} \frac{1}{\sqrt{2\pi}} \frac{\exp\left(-\frac{1}{2} \|\tilde{\mathbf{I}} - \mathbf{C}\|^2 - \rho(\mathbf{C})\right)}{\sum_{\mathbf{J} \in \mathbb{F}^{m \times n}} \exp\left(-\frac{1}{2} \|\tilde{\mathbf{I}} - \mathbf{J}\|^2 - \rho(\mathbf{J})\right)} \quad (\text{1.7})$$

Finnaly, solving [Equation \(1.7\)](#) is equivalent to solve

$$\widehat{\mathbf{I}} = \arg \min_{\mathbf{C}} \frac{1}{2} \|\widetilde{\mathbf{I}} + \mathbf{C}\|^2 + \rho(\mathbf{C}). \quad (1.8)$$

The first term appears so often in imaging problems that it has a special name: *data fidelity*. In the denoising problem, the data fidelity term appeared as a consequence of the Gaussian noise model assumption. The second term is also a regularization term and it favors images that respect some desirable property for the problem to be solved. Since natural images have a higher spatial dependency, a reasonable guess for ρ would be a function that has lower value for piecewise smooth data, i.e., images composed by closed regions with smooth variations in its interior but possibly strong discontinuities in their boundaries.

Tikhonov regularization

The classical way to optimize Equation (1.8) is to shift it to a continuous setting, analytically derive some optimization properties and then use these properties to solve the problem in a discrete setting. The continuous reformulation of Equation (1.8) consists in to optimize the energy functional below

optimizing

$$f_{\tilde{\mathbf{I}}} = \arg \min_f F(f) = \frac{1}{2} \int_{\Omega} \|f_{\tilde{\mathbf{I}}} - f\|^2 dx + R(f), \quad (1.9)$$

where R is a functional derived from the choice of ρ . A popular choice for R is to define it as the $L2$ norm of ∇f , also called the *Tikhonov* regularization term. Equation (1.9) is rewritten as

$$f_{\tilde{\mathbf{I}}} = \arg \min_f F(f) = \frac{1}{2} \int_{\Omega} \|f_{\tilde{\mathbf{I}}} - f\|^2 dx + \int_{\Omega} \|\nabla f\|^2 dx. \quad (1.10)$$

Euler-Lagrange equation

We can establish some necessary optimization conditions for Equation (1.10) by deriving its *Euler-Lagrange* equation. Assume that function g minimizes functional F , i.e.,

$$g = \arg \min_f F(f).$$

Further, assume that there exists a function w that agrees with g at the boundary of f' domain, i.e., $w(x) = 0, \forall x \in \partial\Omega$. Define the function h as

$$h(\epsilon) = F(g + \epsilon w)$$

Therefore, h has a minimum at $\epsilon = 0$. Thus,

$$\begin{aligned} 0 &= \frac{dh}{d\epsilon}_{|\epsilon=0} = \frac{d}{d\epsilon}_{|\epsilon=0} \frac{1}{2} \int_{\Omega} \|f_{\tilde{\mathbf{I}}} - g - \epsilon w\|^2 + \|\nabla(g + \epsilon w)\|^2 \\ &=_{|\epsilon=0} \frac{1}{2} \int_{\Omega} 2\|f_{\tilde{\mathbf{I}}} - g - \epsilon w\| \frac{(f_{\tilde{\mathbf{I}}} - g - \epsilon w)}{\|f_{\tilde{\mathbf{I}}} - g - \epsilon w\|} w + 2\|\nabla(g + \epsilon w)\| \frac{(\nabla g + \epsilon w)}{\|\nabla(g + \epsilon w)\|} \nabla w \\ &= \int_{\Omega} (f_{\tilde{\mathbf{I}}} - g)w + (\nabla g)\nabla w. \end{aligned}$$

Applying integration by parts and using the fact that $w(x) = 0, \forall x \in \partial\Omega$.

$$0 = \int_{\Omega} (f_{\tilde{\mathbf{I}}} - g - \Delta g)w$$

Since w could be any function, we can write

$$f_{\tilde{\mathbf{I}}} - g - \Delta g = 0 \tag{1.11}$$

Therefore, if g is a minimum of [Equation \(1.9\)](#), then it respects the convex [Equation \(1.11\)](#). Hence, given an initial solution f , one can execute a descent method (gradient descent, for example) to find its minimum. In practice, [Equation \(1.9\)](#) is discretized using the samplings $\hat{\mathbf{I}}, \tilde{\mathbf{I}}$ of $f_{\hat{\mathbf{I}}}, f_{\tilde{\mathbf{I}}}$ and a finite differences scheme is defined to estimate the Laplacian Δ .

The Tikhonov term favors images with smooth variations in color, but the smoothness is not restricted to the interior of regions. Thus, Tikhonov tends to obfuscate the discontinuities that will likely be present in the [edges](#) and we have the impression that the image is blurred (see [Figure 1.2](#)). Nonetheless, Tikhonov term is attractive due to its optimization properties.

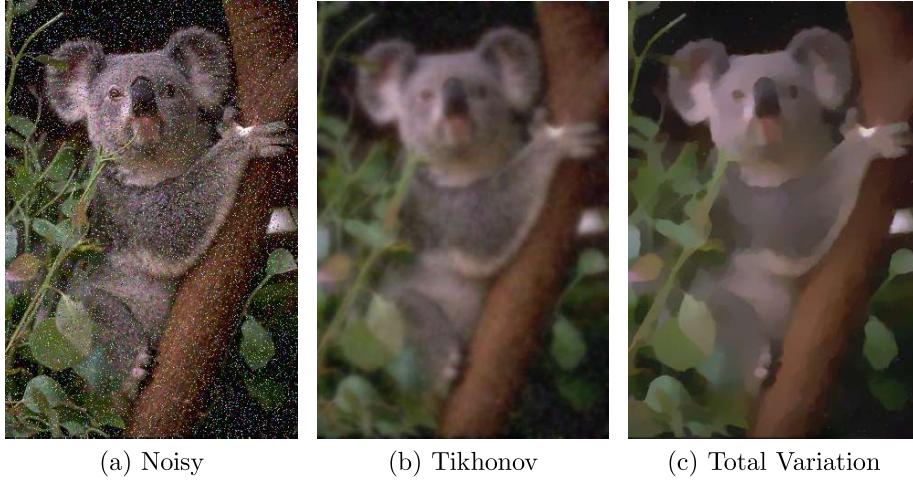


Figure 1.2: Denoising algorithms results for the Tikhonov and Total variation regularization terms.

Total variation regularization

An alternative to Tikhonov regularization is to use the so called *total variation* of the image function. For a smooth function f , its total variation is computed as

$$TV(f) = \int_{\Omega} \|\nabla f\|.$$

For a more general (possibly not differentiable) locally integrable function $f : \Omega \rightarrow \mathbb{R}^n$, its total variation is defined as

$$\begin{aligned} TV(f) &= \sup \left\{ \int_{\Omega} \nabla u \cdot \phi \mid \phi \in C_c^1(\Omega, \mathbb{R}^n) \text{ and } \|\phi\|_{\infty} \leq 1 \right\} \\ &= \sup \left\{ - \int_{\Omega} u \nabla \cdot \phi \mid \phi \in C_c^1(\Omega, \mathbb{R}^n) \text{ and } \|\phi\|_{\infty} \leq 1 \right\}, \end{aligned}$$

where ϕ is vector-valued infinitely differentiable function with compact support. The image denoising total variation model is written as

$$f_{\hat{I}} = \arg \min_f \frac{\lambda}{2} \int_{\Omega} \|f_{\tilde{I}} - f\|^2 dx + TV(f). \quad (1.12)$$

The ROF model [ROF92] assumes the image representation is smooth, and from the Euler-Lagrange equation of [Equation \(1.12\)](#) the following gradient flow is derived:

$$\frac{df}{dt} = \nabla \cdot \left(\frac{\nabla f}{\|\nabla f\|} \right) - \lambda(f_{\tilde{I}} - f) \quad \text{\textbackslash} \text{frac}\{\partial f\}\text{\textbackslash}\partial t \}$$

The left term is not differentiable, and a small $\epsilon > 0$ is added to the denominator in order to avoid numerical instability. However, the calibration of ϵ might be delicate, since a small ϵ might not be sufficient to avoid instability and a larger ϵ may disfigure the model. In [Cha04] the total variation definition is exploited to create a convergent algorithm that works by successive projections and that solves model [Equation \(1.12\)](#). In [BT09] a modified version of the previous algorithm has proven to have faster convergence.

The total variation term is characterized by its smooth properties while partially preserving some discontinuities in the edges. In this sense, total variation models produce results with sharper edges than those produced by the Tikhonov term (see [Figure 1.2](#)).

across

1.3 Standard techniques

In this section we give an overview of the key techniques in variational models for problems in image processing. We start by describing the most influential model in this category.

Mumford-Shah

The L_2 -norm regularization has nice optimization properties, but it does not preserve discontinuities along the object boundaries. This effect is attenuated using a L_1 -norm, but it is not sufficient to avoid blurred edges. The *Mumford-Shah* functional [MS89] handles this issue by incorporating the edges in its formulation in the form of a set of discontinuities \mathcal{K} and limiting the L_2 -norm regularization to points in the interior of objects, i.e., $\Omega \setminus \mathcal{K}$. Moreover, the set \mathcal{K} itself is compelled to be of small length. The Mumford-Shah model consists in to minimize the following functional minimizing

$$(f_{\hat{I}}, \hat{\mathcal{K}}) = \arg \min_{f, \mathcal{K}} \alpha \int_{\Omega} \|f_{\tilde{I}} - f\| dx + \beta \int_{\Omega \setminus \mathcal{K}} \|\nabla f\| dx + \lambda Per(\mathcal{K}) \quad (1.13)$$

The functional can be seen as a model for both denoising and segmentation problems. The function $f_{\hat{I}}$ being the denoising solution and $\hat{\mathcal{K}}$ the segmentation

solution. [Equation \(1.13\)](#) is proven to have a minimizer [DCL89], and in the case \mathcal{K} is fixed, the minimizer is unique (see chapter 25 of [Bar+11]). However, to find a minimizer of [Equation \(1.13\)](#) is a challenging task due to its non-convexity.

Nonetheless, there exist several approximations to the Mumford-Shah functional. We refer to the phase-field model of [AT90]; the finite-differences scheme of [Cha99]; the level-set method of [VC02]; the convex relaxations of [Poc+09; SC14]; and the discrete calculus approach of [FLT17].

1.3.1 Curve evolution

Active contours

Active contours or snakes is a supervised method for doing image segmentation. In the original work [KWT88], an initial parametric curve $C_0(q) \rightarrow (x(q), y(q))$ is evolved towards the local minimum of the snakes energy

$$\begin{aligned} F(C) &= \alpha \text{Length}(C) + \beta \text{Smoothness}(C) + \gamma \text{Edge}(C) \\ F(C) &= \alpha \int_0^1 \left\| \frac{dC}{dq} \right\|^2 dq + \beta \int_0^1 \left\| \frac{d^2C}{dq^2} \right\|^2 dq - \gamma \int_0^1 \|\nabla f_I(C(q))\|^2 dq \end{aligned} \quad (1.14)$$

The length and smoothness regularization term favors curves of smooth variations and small length while the edge term compels the curve to stop at regions of high variation of color intensity.

The snakes method was devised having an interactive framework in mind. First of all, the user must set the initial curve close to the object to be segmented, and besides that, a set of additional tools as anchor points, repulsion and spring forces are available for online modification of the problem. The user can make use of these tools to conveniently perturb the current solution and force the curve to evolve to the expected local optimum.

The active contours is an influential paradigm for image segmentation and it was particularly popular for segmenting medical images [TD96]. Variations of the original model include extension to 3D-segmentation [MT99] and topologically adaptable snakes [MT95].

drawbacks ? Some minus in the active contours formulation includes its non-intrinsic definition, i.e., the curve is not defined in terms of its geometric properties and its representation depends on the chosen parametrization; and, partially as consequence of the latter, its inability to change the initial curve topology. One needs to initialize several snakes in order to correctly segment a picture with several holes, for example.

Geometric active contours

Parametric models as snakes are often criticized because of their non-intrinsic definition, i.e., the energy is not defined in terms of the geometric properties of the contours. That makes the theoretical analysis of the snakes model harder, as the evolution of the contour itself depends of the chosen parametrization. In [Cas+93] the authors propose a model based on the mean curvature motion of the level-sets of a C^2 function u .

Let $u : \Omega \subset \mathbb{R}^2 \rightarrow [0, 1]$ be a C^2 function. The curvature at its k -th level set is given by

$$\kappa(x, y) = \nabla \cdot \left(\frac{\nabla u}{\|\nabla u\|} \right), \quad \forall (x, y) \in \{ (x, y) \mid u(x, y) = k \}$$

The *geometric active contour* model consists into **evolve** an extension of u , by including an artificial time parameter t , and compute the steady solution of the flow

$$\begin{aligned} u(0, x, y) &= u(x, y) \\ \frac{du}{dt} &= g(\|\nabla f_I\|) \|\nabla u\| \nabla \cdot \left(\frac{\nabla u}{\|\nabla u\|} + v \right), \end{aligned} \tag{1.15}$$

where g is a non-increasing function that plays the role of an edge-detector, e.g., $g(x) = 1/(1+x)^2$. The function u can be initially defined as a smoothed version of $1 - \chi_C$, where χ_C is the characteristic function of some set $C \in \Omega$ that contains the objects to be segmented.

Following [Equation \(1.15\)](#), the gray level at some point (x, y) changes proportionally to the curvature of its belonging level set. The constant v forces the change in u to be always positive, i.e., pixels gets lighter, never darker. The term $\|\nabla u\|$ allows u to evolve only at some neighborhood of the 0-level-set boundary and the term $g(\|\nabla f_I\|)$ makes the evolution to stop if an edge is reached. At the steady solution of [Equation \(1.15\)](#) the segmented objects of I corresponds to the 0-level set of u .

Differently from the snakes, the geometric active contours handle changes in topology of the initial curve. In figure [Figure 1.3](#), the initial 0-level set of u splits in three disjoint sets at the steady solution of [Equation \(1.15\)](#). However, the geometric active contour models cannot segment objects with holes without including a region-based term [Che+06].

evolving

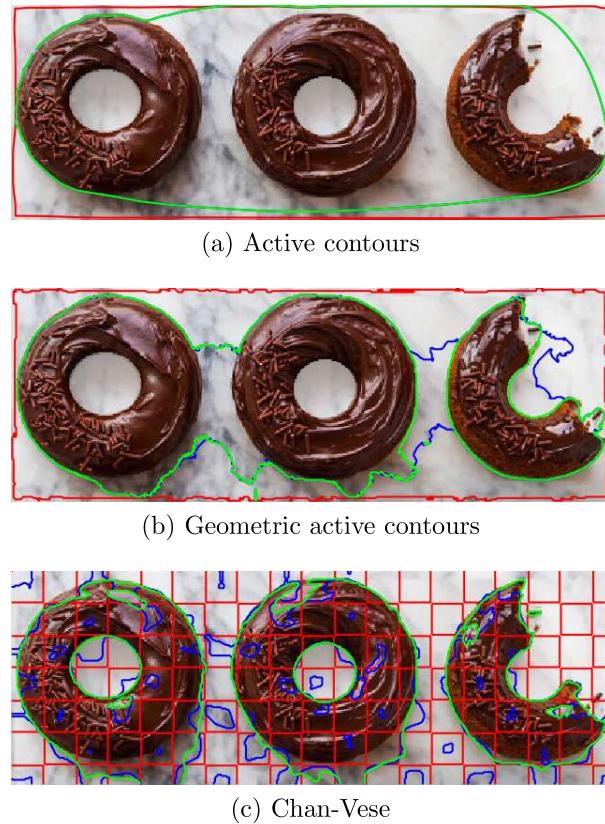


Figure 1.3: Illustration of evolution curve models. The initial curve (0-level set) is colored in red and the final one is colored in green.

1.3.2 Level set

The active contour and its geometric version are both edge-based methods, a natural strategy for image segmentation but with limitations, e.g., the models may encounter some difficulties to segment objects with holes. The *Chan-Vese* method proposes the inclusion of a region-based term and it generalizes the level-set approach already presented in the **geomtric** active contour model.

Let $f_I : \Omega \subset \mathbb{R}^2 \rightarrow \mathbb{U}^2$ a grayscale image and $F \subset \Omega$ an open set such that the pair $(F, \Omega \setminus F)$ is the searched binary partition. Further, assume that there exists a function $\phi : \Omega \rightarrow \mathbb{R}$ with bounded first derivative. The image partitions are identified in the following fashion

$$\begin{aligned}\phi(x) &< 0, \quad \forall x \in F \\ \phi(x) &= 0, \quad \forall x \in \partial F \\ \phi(x) &> 0, \quad \forall x \in \Omega \setminus \overline{F}\end{aligned}$$

In possession of the partition descriptor ϕ , the following energy is proposed

$$\begin{aligned}F(\phi, x) &= \mu Length(\phi, x) + \nu Area(\phi, x) + \lambda_1 Foreground(\phi, x) + \lambda_2 Background(\phi, x) \\ &= \mu \int_{\Omega} \delta_0(\phi(x)) \|\nabla \phi(x)\| dx + \nu \int_{\Omega} H(\phi(x)) dx \\ &\quad + \lambda_1 \int_{\Omega} (1 - H(\phi(x))) \|f_I(x) - c_F\|^2 dx + \lambda_2 \int_{\Omega} H(\phi(x)) \|f_I(x) - c_B\|^2 dx,\end{aligned}\tag{1.16}$$

where $H(x)$ is the Heaviside function and δ_0 the standard Dirac delta function, i.e.,

$$H(x) = \begin{cases} 1, & x \geq 0 \\ 0, & \text{otherwise,} \end{cases} \quad \delta_0(x) = \begin{cases} +\infty, & x = 0 \\ 0, & \text{otherwise.} \end{cases} \quad \text{and } \int_{-\infty}^{+\infty} \delta_0(x) dx = 1.$$

The parameters c_f, c_b are defined as the average color intensity in the interior of the foreground and background regions, respectively

$$c_F = \frac{\int_{\Omega} (1 - H(\phi(x))) f_I(x) dx}{\int_{\Omega} (1 - H(\phi(x))) dx}, \quad c_B = \frac{\int_{\Omega} H(\phi(x)) f_I(x) dx}{\int_{\Omega} H(\phi(x)) dx}.$$

Next, the Euler-Lagrange equation of [Equation \(1.16\)](#) is calculated and used to define a gradient flow to minimize [Equation \(1.16\)](#), in a similar fashion as done in [Section 1.2](#). In order to be numerically tractable, the Heaviside and Dirac delta function are regularized as

$$H_\epsilon(x) = \frac{1}{2} \left(1 + \frac{2}{\pi} \arctan\left(\frac{x}{\epsilon}\right) \right), \quad \delta_\epsilon(x) = \frac{\epsilon}{\pi(\epsilon^2 + x^2)}.$$

The initial level set function can be set as any function with bounded first derivative, but it is recommended to use the checkerboard function $\phi = \sin(\pi/5x_1)\sin(\pi/5x_2)$ which is reported [[Get12](#)] to have fast convergence. In [[VC02](#)], the Chan-Vese authors extended their method to contemplate colored images and multisegmentation. An illustration is presented in [Figure 1.3](#)

1.3.3 Minimum path

In a sequel work, the authors of geometric active contours established a link between their geometric model in [[Cas+93](#)] and the computation of geodesics in a regular surface [[CKS97](#)].

The length of a parametric curve $C(q)$ according to [with](#) an isotropic metric of potential $W(C)$ is calculated as

$$L(C) = \int W(C) \|C_q\| dq. \quad (1.17)$$

[Equation \(1.17\)](#) is used to compute shortest paths between two points according to the given metric. By properly setting the potential W , we can make object boundaries in an image f_I to match the curves of shortest length, for example, letting $W = g(\|\nabla f_I\|)$ as in [Equation \(1.15\)](#) we obtain

$$L(C) = \int g(\|\nabla f_I(C(q))\|) \|C_q\| dq. \quad (1.18)$$

In [[CKS97](#)] the authors show that the snakes model without the smoothness term ($\beta = 0$) is equivalent to geodesics computations, the metric changing accordingly with the models parameters. The isotropic metric above is equivalent to the case in which the length and image terms of the snakes model are equal.

Given an initial curve $C_0(q)$, a local minimizer for [Equation \(1.18\)](#) can be computed by finding the steady solution of the following flow derived from its Euler-Lagrange equation

$$C(0, q) = C_0(q) \quad (1.19)$$

$$\frac{dC}{dt} = (g\kappa - \nabla g \cdot \mathbf{n})\mathbf{n}, \quad (1.20)$$

where \mathbf{n} is the normal vector to the curve C at (t, q) . One can show that, given an initial function $u \in \mathcal{C}^2$ such that u is negative (positive) in the interior (exterior) of its 0-level set, the solution of [Equation \(1.20\)](#) equals the steady solution of

$$u(0, x, y) = u_0(x, y) \quad (1.21)$$

$$\frac{du}{dt} = g\|\nabla u\|\kappa + \nabla g \cdot \nabla u \quad (1.22)$$

Comparing [Equation \(1.22\)](#) with [Equation \(1.15\)](#) we notice that the $\nabla g \cdot y$ term was included while the v parameter was removed. The geometric active contour stops as soon as an ideal edge is found (a threshold should be set), which is particularly bad for real images segmentation, as it is likely that the flow will stop at the first variation of color intensity. The new term allows the flow to evolve even in those cases. Nonetheless, one can include again the v parameter, as it permits to increase the convergence in some cases.

The relation between segmentation and geodesic computation inspired several works. Elongated and thin objects as blood vessels or roads in satellite images are the global optimum of a geodesic computation in which the minimum path is constrained to lie between two points [CK97]. Further development of this work reduced the initialization to just a single point [BC09]. Anisotropic metrics aligned to the image edges are reported to return improved solutions for blood vessels segmentation [Jba+08; BC11]. Ideas from Chan-Vese and Geodesic models are put together in [Che+06]. Finally, an elucidating review of geodesic methods in computer vision can be found in [Pey+10].

1.3.4 Convex relaxation

A set Ω is convex if for all $a, b \in \Omega$, every element in the line connecting a, b also lie in Ω . A function $f : \Omega \rightarrow \mathbb{R}$ is convex if its domain Ω is convex and the following inequality holds

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y).$$

Convexity is a desirable property in optimization problems because any local minimizer (maximizer) is also a global one. In other words, methods that optimizes f do not depend on its initialization. Therefore, it is of great interest to find convex formulations for image processing problems.

Ideally, one would look for the so called convex envelope of f , which is the tightest convex function \tilde{f} such that $\tilde{f} \leq f$. In fact, one can compute the envelope of a function f by taking its convex biconjugate.

Definition 1 (Convex conjugate): Let $f : \Omega \rightarrow \mathbb{R} \cup \{+\infty, -\infty\}$. Its convex conjugate is defined as

$$f^*(y) = \sup_x y^T x - f(x)$$

The biconjugate f^{**} is the convex envelope of f . In fact, if f is convex and lower-semicontinuous, $f^{**} = f$ (Frenchel's inequality). Unfortunately, the computation of the biconjugate is known only for a few functions. Nonetheless, the conjugate is key to prove properties on convex optimization algorithms as those based on the proximal operator [Cha04; BT09].

In order to use tools from convex optimization one needs to define a convex energy. Very often in imaging problems the functionals are defined over a non-convex function space, for example, in the binary denoising or the multilabeling problem, in which the optimization function has a discrete range. A straightforward approach is to simply relax the range to a continuous set and execute a standard optimization method. The discrete solution is then obtained by simple rounding. However, fewer are the cases in which the projected solution is optimum or even meaningful to the original problem. An technique that gives guarantees with respect to the quality of the back projected solution is *functional lifting*.

Functional lifting

Consider the binary image denoising problem. Let $f_{\tilde{\mathbf{I}}}$ the observed noisy image and consider the total variation model for binary denoising

$$\min_{f_{\mathbf{I}}: \Omega \rightarrow \{0,1\}} E^{2-den}(f_{\mathbf{I}}) = \min_{f_{\mathbf{I}}} \int_{\Omega} \|\nabla f_{\mathbf{I}}\| + \lambda \int_{\Omega} (f_{\mathbf{I}} - f_{\tilde{\mathbf{I}}})^2, \quad (1.23)$$

This model is not convex, as the optimization variable $f_{\mathbf{I}}$ belongs to the non-convex domain of binary functions. The corresponding level-set formulation of [Equation \(1.23\)](#) (in the same spirit of the Chan-Vese model) is given by

$$\min \int_{\Omega} \|\nabla H(\phi(x))\| + \lambda \int_{\Omega} (H(\phi(x)) - f_{\tilde{I}}(x))^2 \quad (1.24)$$

whose

Which a local minimum is a steady solution of

$$\frac{d\phi}{dt} = H'_\epsilon(\phi) \left(\nabla \cdot \left(\frac{\nabla \phi}{\|\nabla \phi\|} \right) + 2\lambda(f_{\tilde{I}}(x) - H_\epsilon(\phi)) \right) \quad (1.25)$$

We reproduce the example given in [CEN06] to illustrate a situation in which the Chan-Vese method returns a local optimum solution. Assume that the observed image is the characteristic function of a ball of radius R , i.e., $f_{\tilde{I}}(x) = \mathbf{1}_{B_R}(x)$. Moreover, we set the initial guess of the level-set function to be precisely the observed image, i.e., $\phi_0(x) = f_{\tilde{I}}(x)$. The evolution given by Equation (1.25) maintains the radial symmetry of ϕ_0 , which means that $\phi^{(t)}$ will represent the characteristic function of some ball of radius r . In other words, for this particular setting, Equation (1.24) is equivalent to

$$\min_r g(r, \lambda) = \min_r 2\pi r + \lambda\pi|R^2 - r^2|, \quad r \geq 0. \quad (1.26)$$

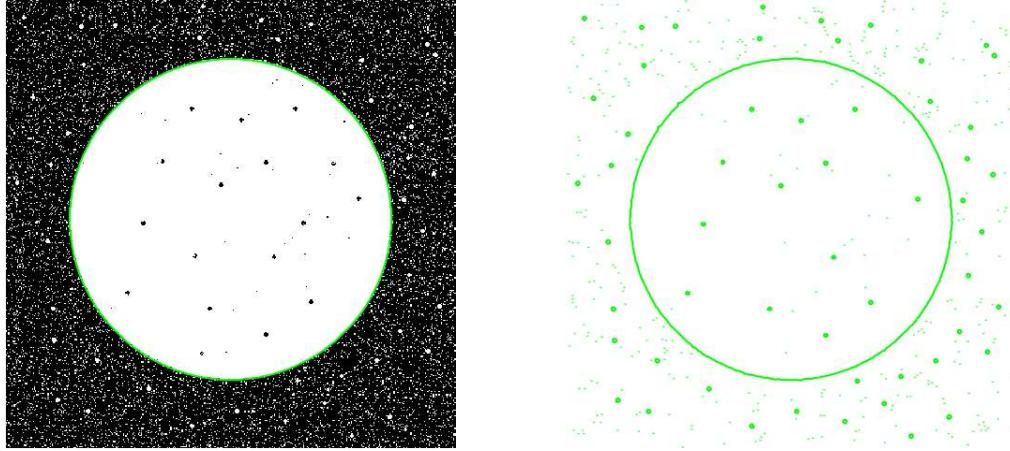
Equation (1.26) has a local minimum at $r = R$ and a local maximum at $r = 1/\lambda$. Depending on the value of λ , its global minimizer will be at $r = 0$ or $r = R$. For a fixed λ , let's consider the case in which $g(0, \lambda) < g(R, \lambda)$.

$$\begin{aligned} g(0, \lambda) &\leq g(R, \lambda) \\ \lambda\pi R^2 &\leq 2\pi R \\ R &\leq \frac{2}{\lambda} \end{aligned}$$

Therefore, for $R \in (1/\lambda, 2/\lambda)$, the flow of Equation (1.25) is going to stop at the local minimum $r = R$, even the global minimum being located at $r = 0$.

The parameter λ is set by the user to calibrate the denoise capabilities of the model. In Figure 1.4 it indicates which ball radius should be considered an object or a noise element. In the illustrated case, with $\lambda = 1$, the optimal solution consists to segment only the ball of larger radius, but Chan-Vese fails to converge to this solution.

je ne comprends pas la phrase



(a) Input image and contour of lowest energy for $\lambda = 1$

(b) Result of level-set method for $\lambda = 1$

Figure 1.4: Level-set result of Equation (1.23) stops at a local minimum in the right, instead of the green contour in the left.

In [CEN06] the authors use an upper-level set representation to derive an equivalent model to Equation (1.23) but with the particular property that global optimum solution can be recovered by simple thresholding. The function u can be rewritten in terms of its upper level representation as

$$f_I(x) = \int_0^1 \varphi(x, \mu) d\mu,$$

where $\varphi(x, \mu)$ is its μ -th upper level set, i.e.,

$$\varphi(x, \mu) = \mathbf{1}_{\{f_I > \mu\}} = \begin{cases} 1, & f_I(x) > \mu \\ 0, & \text{otherwise.} \end{cases}$$

Using the co-area formula we rewrite the total variation term as

$$\int_{\Omega} \|\nabla f_I\| = \int_{\Omega} \int_0^1 \|\nabla \varphi(x, \mu)\| dx d\mu$$

and we can rewrite Equation (1.23) as

$$\min_{\varphi: \Omega \rightarrow \{0,1\}} E^{2-den} = \min_{\varphi} \int_{\Omega} \int_0^1 \|\varphi(x, \mu)\| + (\mu - f_{\tilde{I}}(x))^2 \delta(f_I(x) - \mu) dx d\mu \quad (1.27)$$

$$= \min_{\varphi} \int_{\Sigma} \|\varphi(x, \mu)\| + (\mu - f_{\tilde{I}}(x))^2 |\partial_{\mu} \varphi(x, \mu)| d\Sigma, \quad (1.28)$$

where $\Sigma = [\Omega \times [0, 1]]$. It happens that in the new formulation Equation (1.28), one can recover a binary solution from simple thresholding. From its relaxation

$$\min_{\varphi: \Omega \rightarrow [0,1]} E^{2-den} \quad (1.29)$$

we can once again rewrite E^{2-den} in terms of the upper level set representation of φ

$$\begin{aligned} E^{2-den} &= \int_{\Sigma} \int_0^1 \|\mathbf{1}_{\{\varphi > \gamma\}}\| + (\mu - f_{\tilde{I}}(x))^2 |\partial_{\mu} \mathbf{1}_{\{\varphi > \gamma\}}| d\Sigma d\gamma \\ E^{2-den} &= \int_0^1 E^{2-den}(\mathbf{1}_{\{\varphi > \gamma\}}) d\gamma \end{aligned}$$

Therefore, if φ^* is the solution of the convex relaxed problem Equation (1.29), $\mathbf{1}_{\{\varphi^* > \gamma\}}$ is also an optimal solution of E^{2-den} for almost every choice of γ .

The functional lifting technique creates an equivalent higher dimensional model with the property that binary solutions can be easily recovered from its relaxed solution. In [Poc+08] this strategy is extended for multilabeling problems and in [Poc+09; SCC12] they are used to create a convex relaxation of the Mumford-Shah model.

To optimize the higher dimensional energy one could regularize the indicator and dirac delta functions in the same spirit of the Chan-Vese method, but it is usually preferable to use a convex optimization method that is suitable for non-differentiable functions as the proximal gradient [Cha04], FISTA [BT09] or the primal-dual [CP11] algorithm. The results are very satisfying, but the running times very high.