Isophotes Selection and Reaction-Diffusion Model for Object Boundaries Estimation

Charles Kervrann, Mark Hoebeke & Alain Trubuil
INRA - Biométrie, Domaine de Vilvert, 78352 Jouy-en-Josas, France
{ck,mh,at}@jouy.inra.fr

Abstract. This paper investigates generic region-based segmentation schemes using area-minimization constraint and background modeling, and develops a computationally efficient framework based on level lines selection coupled with biased anisotropic diffusion. A common approach to image segmentation is to construct a cost function whose minima yield the segmented image. This is generally achieved by competition of two terms in the cost function, one that punishes deviations from the original image and another that acts as a regularization term. We propose a variational framework for characterizing global minimizers of a particular segmentation energy that can generates irregular object boundaries in image segmentation. Our motivation comes from the observation that energy functionals are traditionally complex, for which it is usually difficult to precise global minimizers corresponding to "best" segmentations. In this paper, we prove that the set of curves that minimizes the basic energy model under concern is a subset of level lines or isophotes, i.e. the boundaries of image level sets. The connections of our approach with region-growing techniques, snakes and geodesic active contours are also discussed. Moreover, it is absolutely necessary to regularize isophotes delimiting object boundaries and to determine piecewise smooth or constant approximations of the image data inside the objects boundaries for vizualization and pattern recognition purposes. Thus, we have constructed a reaction-diffusion process based on the Perona-Malik anisotropic diffusion equation. In particular, a reaction term has been added to force the solution to remain close to the data inside object boundaries and to be constant in non-informative regions, that is the background region. In the overall approach, diffusion requires the design of the background and foreground regions obtained by segmentation, and segmentation of the adaptively smoothed image is performed after each iteration of the diffusion process. From an application point of view, the sound initialization-free algorithm is shown to perform well in a variety of imaging contexts with variable texture, noise and lighting conditions, including optical imaging, medical imaging and meteorological imaging. Depending on the context, it yields either a reliable segmentation or a good pre-segmentation that can be used as initialization for more sophisticated, application-dependent segmentation models.

Keywords: grouping and segmentation, energy minimization, level sets, level lines, isophotes, connected components, anisotropic diffusion.

1 Introduction

One of the primary goals of early vision is to segment the domain of an image into regions ideally corresponding to distinct physical objects in the scene. While it has been clear that image segmentation is a critical problem in a wide range of computer vision applications, it has proven difficult to express segmentation criteria that capture non-local properties of an image and to develop efficient algorithms for computing segmentations. This is especially true

if the images were obtained under different conditions and have different content. This work is concerned with the segmentation of salient structures from natural images by integrating two approaches based on the minimization of a segmentation functional and the technique of anisotropic diffusion, to form an overall approach that is robust to noise and poor initialization.

In general, the goal of image segmentation is to find groups which are both homogeneous in the same group and well separated (Deriche (1987); Pavlidis and Liow (1990); Chu and Agarwal (1993); Geman and Geman (1984); Blake and Zisserman (1987); Mumford and Shah (1989); Leclerc (1989)). Many of segmentation techniques rely on the design and minimization of an energy functional which captures the interaction between models and image data. We note that one of the most widely studied mathematical models in image processing is the active contour model (or "*snakes*") to extract the boundaries of homogeneous regions within an image (Kass *et al.* (1987)). In practical imaging, the interactive initialization of the energy-based active contour near the desired boundary significantly reduces the difficulty of segmentation. On the other hand, because they rely on the image gradients, these models can fail in presence of strong noise and may be sensitive to the starting position. In addition, most algorithms optimizing the energy function associated to active contours, find only local minima, and thus have no measure of the significance of the extracted boundary for the image as a whole. Several improvements have been proposed to support several levels of user initialization/interaction: a "balloon force" which controls the interior area of the curve has been introduced in (Cohen (1991)) to pass over weak edges and extract salient edges ; topologically adaptive active contour methods involve solving the energy-based active contours minimization by the computation of geodesics (Kichenesamy *et al.* (1996); Caselles *et al.* (1997)). In this approach, a curve is embedded as a zero level set of higher dimensional surface (Osher and Sethian (1988); Sethian (1996)). The entire surface is evolved to minimize an intrinsic Riemannian metric depending on image gradient. These models are more adaptable to detect an arbitrary number of objects and less sensitive to initialization but use unreliable local information to make a hard premature decision.

In other respects, several works were conducted in the field of regions-based approaches which are more suitable for unsupervised image segmentation. Segmentation consists in finding partitions of the image pixels into zones the most homogeneous possible corresponding to coherent image properties such as brightness, color and texture. Homogeneity may be measured by a given global objective function and hard decisions are made only when information from the whole image is examined at the same time. Thus, past approaches have centered on formulating the problem as the minimization of an energy functional involving the image intensity and edge functions. Some energy models are based on a discrete model of the image, such as Markov random fields (Geman and Geman (1984); Blake and Zisserman (1987); Leclerc (1989); Wang (1998)) or markovian deformable templates (Grenander and Miller (1994)) whereas variational models are based on a continuous model of the image (Mumford and Shah (1989); Geiger and Yuille (1991); Mumford (1994); Morel and Solimini (1994); Schnörr (1998); Jermyn and Ishikawa (1999)). Finally, Blake and Zisserman (1987) and Mumford and Shah (1989) have written about

most aspects of this approach to segmentation and have proposed various complex functionals whose minima correspond to segmented images. In a recent review, Morel and Solimini (1994) have, indeed, shown that most approaches aim at optimizing a cost functional which is the combination of three terms: one which ensures that the smoothed image approximates the observed one, another which states that the gradient of the smoothed image should be small, except on a discontinuity set, and a last one which ensures that the discontinuity set has a small length. Such underlying *a priori* constraints encourage the emergence of few regions whose boundaries are regular (Mumford and Shah (1989); Leclerc (1989)). The integration of boundary and region information sources is also currently achieved through a single-objective function (Zhu and Yuille (1996); Paragios and Deriche (2000)). These functionals include both local and global information, adding robustness to noise and weak boundaries.

In other respects, while these different approaches offer powerful theoretical frameworks and minimizers exist (Morel and Solimini (1994); Schnörr (1998)), it is often computationally difficult to optimize the associated functionals. Depending on the formulation of the cost function with either continuous or discrete variables, different minimization techniques can be applied. For cost functions of continuous variables, it is in many cases possible to compute a set of partial differential equations, the Euler Lagrange equations, that are solved by the minima. For cost functions of discrete variables one must generally rely on direct minimization techniques. Often graduate-non-convexity (Blake and Zisserman (1987)) or Monte Carlo Markov Chain algorithms such as the well known simulated annealing (Geman and Geman (1984)) are the methods of choice. Typically, they estimate the curves that maximally separate unknown statistics inside and outside the curves and avoid bad local minima of cost functionals (Blake and Zisserman (1987); Leclerc (1989); Grenander and Miller (1994)). Unlike previous segmentation techniques, other methods use specific *a priori* knowledge to ease the segmentation task: they may assume that the number of objects (Chesnaud *et al.* (2000); Figueiredo *et al.* (2000)), layers (Darrell and Pentland (1995)), classes (Paragios and Deriche (2000); Yezzi *et al.* (1999)), or the statistics inside region boundaries (Grenander and Miller (1994); Samson *et al.* (1999); Chan and Vese (1999); Amadieu *et al.* (1999)) are known or estimated using Expectation-Maximization procedures or ad-hoc methods (Darrell and Pentland (1995); Paragios and Deriche (2000); Yezzi *et al.* (1999)). The region boundaries propagation can be then implemented using the level set theory (Osher and Sethian (1988); Sethian (1996); Kichenesamy *et al.* (1996)) but these supervised segmentation methods may be sensitive to the initial curve conditions (Amadieu *et al.* (1999); Samson *et al.* (1999); Chan and Vese (1999); Paragios and Deriche (2000)).

In this paper, we present a general approach for image segmentation that incorporates edge-preserving smoothing and energy-based regions extraction. Not much research has been done which analyzes the combination of non-linear diffusion methods with segmentation techniques. However, extracting significant connected regions from a piecewise smooth image estimated using non-linear diffusion methods (Perona and Malik (1990); Black and Rangarajan (1996)) still requires considerable effort. In order to create a true segmentation, the regularized image must be processed by a *ad-hoc* method leading to a partitioning of the entire image into a finite

number of regions. This topic has been already addressed in (Weickert (1998b)) by combining partial differential equations (PDE)-based regularization methods with a watershed algorithm (Vincent and Soille (1991)): the algorithm estimates a piecewise almost constant image, and performs the partitioning of the entire image afterwards. In our work, we shall see that regions can be extracted at each discretized time of the diffusion equation as connected components of level sets of the PDE-based regularized image. From a methodology point of view, the main contributions of the paper are the following:

- We aim at characterizing the minimizers of a basic energy functional combining area-minimization constraint with Gaussian priors for luminance distributions over object and background regions, as a subset image level sets. The energy model under concern, introduced in a discrete setting by Beaulieu and Goldberg (1989)) and reviewed by Morel and Solimini (1994), allows to partition the image into regions, though in a more restrictive manner than widely studied approaches (Mumford and Shah (1989); Blake and Zisserman (1987); Leclerc (1989); Geiger and Yuille (1991); Zhu and Yuille (1996)) or more recent approaches (Chan and Vese (1999); Yezzi *et al.* (1999); Samson *et al.* (1999); Paragios and Deriche (2000)) since it can generate irregular object boundaries. The studied energy model is composed of two terms, one that punishes deviations from the image and another that encourages the emergence of a small number of regions with irregular boundaries. Nevertheless, unlike related classification works (Chan and Vese (1999); Yezzi *et al.* (1999); Samson *et al.* (1999); Paragios and Deriche (2000)), the statistics or the number of classes are not estimated before performing the segmentation: the works proposed by Chan and Vese (1999) and Yezzi *et al.* (1999) with an implementation based on the level set formulation concerns respectively the bi-modal and tri-modal cases. More generally, a characterization of minimizers of energy functionals involving a constant-piecewise model to describe images are not always available ; a fairly complete analysis is available only for a simplified version of the Mumford and Shah model that approximates a given image with piecewise constant functions (Koepfler *et al.* (1994); Morel and Solimini (1994)). The present investigation is based on a variational approach. We prove hereafter that the set of curves that minimizes the energy model under concern is a subset of level lines or *isophotes* defined from level sets of the image. In other words, minimizing the energy is equivalent to select a subset of connected components corresponding to image regions that optimally partition the image domain.

- An additional difference with previous related works lies in the algorithmic implementation. Instead of a commonly used level set formulation (Chan and Vese (1999); Yezzi *et al.* (1999); Paragios and Deriche (2000)), we derive an efficient and initialization-free algorithm, based on incremental level lines selection over uniformly quantized version of original data, which yields a sub-optimal solution to the global minimizer.

- In some respects, it is absolutely necessary to regularize object boundaries and to determine piecewise smooth or constant approximations of the image data inside the boundaries of

objects for vizualization and pattern recognition purposes. Hence, we have coupled the level line selection procedure with a robust anisotropic diffusion scheme with separate object/background processing and robust critical scale estimation to provide a regularized segmentation. The spatially varying "edge-stopping" parameter is herein influenced by the segmentation of the foreground (composed of objects) and background regions, at each discretized time of the diffusion process.

The rest of the paper is organized as follows: in Section 2, we describe the energy-based model for image segmentation and we outline its general features and minimizers. In this section, we describe a numerical implementation of the initialization-free segmentation algorithm. In Section 3, we present experimental results obtained by applying the isophotes selection algorithm to synthetic as well as real images. In Section 4, we propose a robust statistical measure of the gradient variation and use this in an adaptive reaction-diffusion framework to regularize level lines delimiting object boundaries. The qualitative and quantitative assessment of the accuracy of the combined approach is provided in Section 5. Conclusions are presented in Section 6.

2 Isophotes selection for image segmentation

Our model is the minimizer of a segmentation energy. Most of the time, the energy is designed as a combination of several terms, each of them corresponding to a precise property which much be satisfied by the optimal solution. These energy models usually have two parts: a *data model* and a *prior model* also called "regularizer" since it was initially conceived to make the problem of minimizing the data model well-posed (in the sense of Hadamard).

2.1 Description of the model

Our theoretical setting is the following. Let S be an open subset of $\mathbb{R}^2$ and f a grey-scale image treated as a function defined on S. In practical imaging S is a collection of pixels within a discretized rectangle, and possible values of f are given by integers $[0, 256[\cap\mathbb{N}$, so both the domain and the range of f are discrete sets. Below we will work in the continuous setup, where S is a subset of a Euclidian space and $f : S \to \mathbb{R}$ represents the observed data function. The continuous setup allows us to refer to analytic tools, while leaving always a possibility to "discretize" the problem. We use the notation f to denote both the function and the corresponding grey-scale image. We use the terminology "site" or "pixel" to denote a point of the image, even in the continuous case. Each point $x \in S$ is assigned a grey value $f(x)$. Without further notice we assume that f is upper semi-continuous. This ensures the measurability of f. According to Matheron (Matheron (1975)), we interpret the image f as a nested family of sets defined by $L_\gamma(f) = \{x \in S : f(x) \geq \gamma\}, \gamma \in \mathbb{R}$. Each level set $L_\gamma(f)$ is assumed to be of finite perimeter (Caselles *et al.* (1999)).

We define the solution to the segmentation problem as the global minimum of a regularized criterion over all regions. Let $\Omega_i \subset S$, $i = 1, \ldots, P$, be a non-empty image domain or object

and $\partial\Omega_i$ its boundary. A partitioning of S consists in finding a set $\{\Omega_i\}_{i=1}^P$ and a background $\overline{\Omega}$ defined as the complementary subset of the union of objects (see Fig. 1):

$$S \setminus \overline{\Omega} \;=\; \bigcup_{i=1}^{P} \Omega_i, \quad \Omega_i \bigcap_{i\neq j} \Omega_j \;=\; \emptyset \;\text{ and }\; \Omega_i \bigcap \overline{\Omega} \;=\; \emptyset. \tag{1}$$

We seek a strong segmentation, that is, a partition of the rectangle S into a finite set of patches, each of which corresponding to a part of the image where f approximated by a constant value. Moreover, we just wish to control the number of regions in the image. The regularization theory leads to associate with the unknown domains Ω_i and $\overline{\Omega}$ the following regularized objective function, inspired from (Beaulieu and Goldberg (1989); Istas (1997)):

$$E_\lambda(f, \Omega_1, \ldots, \Omega_P, \overline{\Omega}) \;=\; \sum_{i=1}^{P} \int_{\Omega_i} (f(x) - \overline{f}_{\Omega_i})^2 \, dx + \int_{\overline{\Omega}} (f(x) - \overline{f}_{\overline{\Omega}})^2 \, dx + \lambda \sum_{i=1}^{P} |\Omega_i| \tag{2}$$

where $\overline{f}_{\Omega_i}$ and $\overline{f}_{\overline{\Omega}}$ denote respectively the unknown average values of f over Ω_i and $\overline{\Omega} = S \setminus \cup_{i=1}^P \Omega_i$, $|\Omega_i|$ is the two-dimensional Lebesgue measure which correspond to the area of Ω_i since S is defined as a subset of the Euclidian space, and $\lambda \geq 0$ is the regularization parameter. We point out that the two first terms in equation (2) are coming from the Mumford-Shah energy (Mumford and Shah (1989)) and the last term has been already used in active contours (Zhu and Yuille (1996)), balloons models (Cohen (1991); Zhu and Yuille (1996); Rougon and Preteux (1998)) and also by Nitzberg and Mumford (1990). We shall see in Section 2.5 that objects can be extracted as connected components of image levels sets.

Instead of fixing *a priori* the cardinality of the segmentation, which is a highly arbitrary choice, it seems more natural to control the emergence of regions by an object area-based penalty weighted by a parameter λ. The regularization term $E_p(\Omega_1, \cdots, \Omega_P) = \sum_{i=1}^P |\Omega_i|$ is herein introduced to encourage the emergence of a large background $\overline{\Omega}$ and gives no control on the local smoothness of boundaries. The constant λ can be interpreted as a scale parameter of the functional that only tunes the number of regions (Beaulieu and Goldberg (1989); Morel and Solimini (1994)). If λ is large, the background is encouraged to be a large part of the image. As λ decreases, a lot of regions are allowed and the segmentation is fine. If $\lambda = 0$, each point is potentially a region and $\overline{\Omega} = \emptyset$; the global minimum value coincides with zero and this segmentation is called the "*trivial segmentation*" (Koepfler *et al.* (1994); Morel and Solimini (1994)).

In this modeling, implicitly, a Gaussian distribution for the noise is assumed. This means that one observes a corrupted function $f = f_{\text{true}} + \varepsilon$, where ε is a zero-mean Gaussian white noise. The true image f_{true} is supposed piecewise constant:

$$f_{\text{true}}(x) = \sum_{i=1}^{P} \overline{f}_{\Omega_i} \, 1_{x \in \Omega_i} \;+\; \overline{f}_{\overline{\Omega}} \, 1_{x \in \overline{\Omega}}, \tag{3}$$

where, for a set $E \subset S$, the set indicator function $1_{x \in E} = 1$ if $x \in E$ and $1_{x \in E} = 0$ otherwise. The standard deviation is assumed to be constant over the entire image (Zhu and Yuille (1996)).

Our aim is now to define objects in f. Therefore, we define the following collection $\mathcal{C}_P$ of $P \geq 0$ admissible objects

$$\mathcal{C}_P \;=\; \{\{\Omega_1, \ldots, \Omega_P\} \subset S \;\text{ are connected }\; ;\; S \setminus \overline{\Omega} = \bigcup_{i=1}^{P} \Omega_i \;; \Omega_i \bigcap_{1 \leq i \neq j \leq P} \Omega_j = \emptyset \;\} \tag{4}$$

where the subsets $\{\Omega_1, \ldots, \Omega_P\}$ are the objects of the image and $\overline{\Omega}$ is the background. When $P = 0$, there is no object in the image. An optimal segmentation of image f is by definition a global minimum of the energy (when exists)

$$(\Omega_1^\star, \ldots, \Omega_{P\star}^\star, \overline{\Omega}^\star) \;=\; \text{argmin}_{0 \leq P \leq T} \;\; \text{argmin}_{\{\Omega_1, \ldots, \Omega_P\} \in \mathcal{C}_P} \;\; E_\lambda(f, \Omega_1, \ldots, \Omega_P, \overline{\Omega}) \;. \tag{5}$$

where the collection $\mathcal{C}_T$ denotes the bank of all admissible objects, i.e. $\mathcal{C}_P \subseteq \mathcal{C}_T, \forall P \leq T$, and T is the maximum number of registered objects in the image. In Fig. 1, $T = 4$ and a minimal energy-based partition is given by the set of $(P = 2)$ objects $\{\Omega_2, \Omega_3\}$ and a background $\overline{\Omega} = \Omega_1 \cup \Omega_4$.

FIGURE 1 TO BE PLACED HERE

A joint minimization with respect to all unknown domains Ω_i and parameters $\overline{f}_{\Omega_i}$ is an intricate problem (Blake and Zisserman (1987); Mumford and Shah (1989); Morel and Solimini (1994); Zhu and Yuille (1996); Cohen (1996)), even if T is low, because of the large number of possibilities of placing objects inside S and P is unknown. In addition, the set of unknown variables (sets and functions) are, by definition, clearly not independent. In Section 2.4, we prove that the object boundaries that minimize (2) are level lines of function f, which makes the problem tractable.

2.2 Related works

The Mumford-Shah functional (Mumford and Shah (1989)) is known as the most synthetic criterion for the segmentation process (Morel and Solimini (1994)), depending on two variables, the unknown image function $\hat{f}$ and the set of unknown boundaries denoted D. This functional is the combination of three terms: the first term ensures that the estimated function $\hat{f}$ approximates the observed function f ; the second term states that the gradient $\nabla \hat{f}$ is small, except on a discontinuity set D ; the last term leads to a discontinuity set D having a small length. The simplest energy functional associated with this generic description of Mumford and Shah is the restriction to piecewise-constant functions (Koepfler *et al.* (1994); Morel and Solimini (1994); Zhu and Yuille (1996)):

$$E_{MS}(\hat{f}, D) \;=\; \int_{S \setminus D} (f(x) - \hat{f}(x))^2 \, dx \;+\; \nu \, \ell(D), \tag{6}$$

where D is a union of boundaries in S with length $\ell(D)$ and $\hat{f}$ is piecewise constant on $S \setminus D$. The constant ν represents the scale parameter of the functional and measures the amount of boundaries: if ν is low a lot of boundaries are allowed ; as ν increases, the segmentation gets coarser. This functional represents the simplest compromise between accuracy of the regions and parcimony of the boundaries. If we fix the boundaries D, then the corresponding minimal

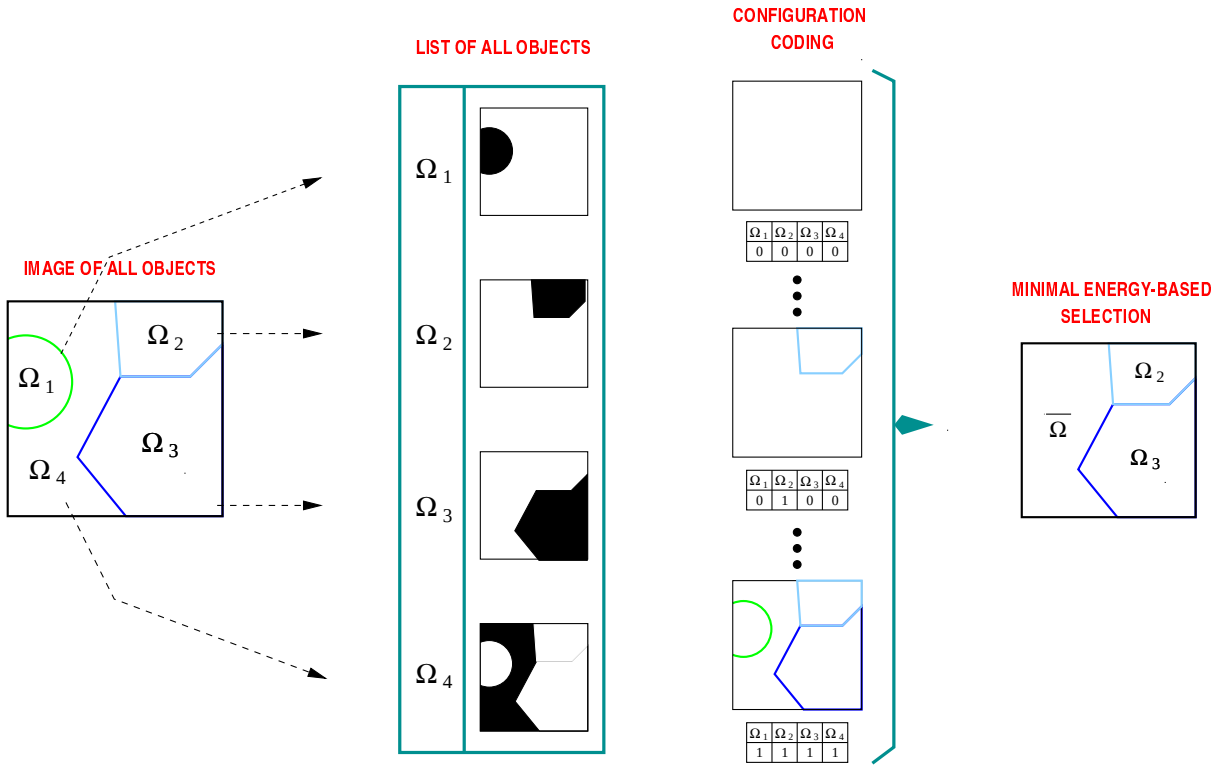


Figure 1: A minimal energy-based partition is given by the set of ($P = 2$) objects $\{\Omega_2, \Omega_3\}$ and a background $\overline{\Omega} = \Omega_1 \bigcup \Omega_4$.

$\hat{f}$ is completely defined by the fact that its value on each region Ω_i of $S \setminus D$ is completely defined by the fact that its value on each region Ω_i of $S \setminus D$ is equal to the average value of $\overline{f}_{\Omega_i}$ of f over Ω_i. This corresponds to the "*cartoon*" model (Zhu and Yuille (1996))

$$E_{\text{cartoon}}(f, \Omega_1, \ldots, \Omega_P) \;=\; \sum_{i=1}^{P} \int_{\Omega_i} (f(x) - \overline{f}_{\Omega_i})^2 \, dx \;+\; \nu \; \ell(\bigcup_{i=1}^{P} \partial\Omega_i). \tag{7}$$

which also includes the definition of a connected "background" region since, by definition, the union of $\Omega_i = S$. Unlike this model, the "background" region defined in our segmentation energy (2) is intentionally a non-connected region. In addition, from a numerical point of view, it is not easy to compute a minimizer of (7) since the set of boundaries is unknown. We mention the level set method (Osher and Sethian (1988); Sethian (1996)) methods and the merging regions method (Koepfler *et al.* (1994); Morel and Solimini (1994)) to solve the so-called "*minimal partition problem*".

We note also that our energy functional (2) is also related to *balloons* models (Cohen (1991); Zhu and Yuille (1996); Rougon and Preteux (1998)). Snakes and balloons have been typically used with creative external forces to segment various anatomical structures. An active contour is a closed contour of a region $\partial\Omega$, defined by $\partial\Omega(s)$ where s can be the arc length of the contour. The two-dimensional balloon model extends the snake energy to include a force $\lambda\vec{N}(s)$ which pushes the contour out (or in) along its normal $\vec{N}(s)$. Changing the sign of λ makes this normal force inflating or deflating. Contrasting the balloon and snake models, we note that incorporation of the normal force in the balloon model allows the initial position of the contour to be further from the intended final position, while still enabling convergence. As well, in the balloon model, the initial position can lie either inside or outside the intended contour, while the snake model requires the initial position to surround the intended contour (Cohen (1991); Rougon and Preteux (1998)). The additional inflation or deflation force can be derived from an energy functional. It corresponds to the following energy term $|\Omega| = \lambda \int_\Omega dx$ which is a reduced form of our prior model $E_p(\Omega_1, \cdots, \Omega_P) = \lambda \sum_{i=1}^{P} |\Omega_i|$ where $P = 1$. Thus, the balloon model maximizes its area while maximizing the intensity gradient along the contour. Because of the isoperimetric inequality, this results almost everywhere into a smoothing behavior.

We want also to point out that the approach presented in this paper shares common aspects with the *region competition* approach of Zhu and Yuille (Zhu and Yuille (1996)). The Zhu-Yuille's algorithm is derived by minimizing a generalized Bayes/MDL criterion using the variational principle. In this part of the section, we propose to determine the motion for a point x_b at the common boundary $\partial\Omega_i$ (parametrized by $s \in [0,1]$) of a region Ω_i and the background $\overline{\Omega}$ by computing the gradient descent on (2). The time dependent position of the boundary $\partial\Omega_i$ can be expressed parametrically by $x_b(s,t)$. The motion of the boundary $\partial\Omega_i$ is governed by the Euler-Lagrange differential equation. For any point $x_b(s,t)$ on the boundary we obtain

$$\begin{aligned} \frac{dx_b(s,t)}{dt} &= -\frac{\delta E_\lambda(f, \Omega_i, \overline{\Omega})}{\delta x_b(s)} \\ &= \lambda\vec{N}(x_b(s,t)) \;+\; \left[(f(x) - \overline{f}_{\Omega_i})^2 - (f(x) - \overline{f}_{\overline{\Omega}})^2\right] \vec{N}(x_b(s,t)) \end{aligned} \tag{8}$$

where $\vec{N}(x_b(s,t))$ is the unit normal to $\partial\Omega_i$ at point $x_b(s,t)$. There are two forces acting on the contour, both pointing along the normal: the first term is analogous to a pressure term ; the second term is the *statistics force*. The better the point $x_b(s,t)$ satisfies the homogeneity requirement, and the weaker is the statistics force. This equation can be seen as a degenerate case of the *region competition* algorithm (Zhu and Yuille (1996)). The Euler-Lagrange equations solving for each region can be complex and the *region competition* algorithm finds a local minima. Notice the equation (7) does not hold at T-junctions between domains boundaries $\partial\Omega_i$ where the normal vector is probably ill-defined. The level set framework has been recently investigated for treating this issue (Merriman *et al.* (1994)). In particular, using the level-set formulation (Osher and Sethian (1988); Chan and Vese (1999); Samson *et al.* (1999); Paragios and Deriche (2000); Yezzi *et al.* (1999)), suitable numerical schemes have been derived for solving propagating equations. However, in both cases, seed regions must be provided by the user or randomly put across the image, and average values $\overline{f}_{\Omega_i}$ and $\overline{f}_{\overline{\Omega}}$ are updated at each step of the iterative algorithm. In this paper, we shall see that the steady solution of a set of P equations of the form (8) can be characterized (P regions superimposed on a background $\overline{\Omega}$).

2.3 Upper bound of the objects number

As pointed out by Zhu and Yuille (1996), regularization parameters calibrating energy terms in segmentation models can be interpreted in the sense of statistics [1]. It appears, most of the time, that variations in the values of these parameters λ have significant effects on the qualitative properties of the minimizer (Younes (2000)). In this section, we show that the maximum number of objects is explicitly influenced by λ.

LEMMA 1. *If there exists an optimal segmentation defined by (2) and* $|\Omega_i| \geq |\Omega_{min}|, i = 1, \cdots, P$, *then the optimal number* $P^\star$ *of objects is upper bounded by*

$$P_{\max} \;=\; (\lambda\ |\Omega_{\min}|)^{-1} \int_S (f(x) - \overline{f_S})^2 dx \quad \text{if} \quad E_{\mathrm{p}}(\Omega_1, \ldots, \Omega_P) \;=\; \sum_{i=1}^{P} |\Omega_i|.$$

Proof (of Lemma 1). Suppose that the segmentation energy is maximum when there is no object in the image. We have

$$\lambda \sum_{i=1}^{P^\star} |\Omega_i^\star| \;\leq\; E_\lambda(f, \Omega_1^\star, \ldots, \Omega_{P^\star}^\star, \overline{\Omega}^\star) \;\leq\; E_\lambda(f, S) = \int_S (f(x) - \overline{f_S})^2 dx.$$

[1] We can reformulate the minimization of $\sum_{i=1}^{P} \int_{\Omega_i} (f(x) - \overline{f}_{\Omega_i})^2\, dx + \int_{\overline{\Omega}} (f(x) - \overline{f}_{\overline{\Omega}})^2\, dx + \lambda \underbrace{\sum_{i=1}^{P} |\Omega_i|}_{|S|\setminus|\overline{\Omega}|}$ as the minimization of $\sum_{i=1}^{P} \int_{\Omega_i} (f(x) - \overline{f}_{\Omega_i})^2\, dx + \int_{\overline{\Omega}} \left[(f(x) - \overline{f}_{\overline{\Omega}})^2 - \lambda\right]\, dx$. λ may be then interpreted as a classical Gaussian threshold chosen according to the normal law table for a given percentage.

If $|\Omega_i| \geq |\Omega_{\min}|$, we have $P^\star \, |\Omega_{\min}| \;\leq\; \sum_{i=1}^{P^\star} |\Omega_i^\star| \;\leq\; \lambda^{-1} \int_S (f(x) - \overline{f_S})^2 \, dx$. Finally, we obtain

$$P^\star \;\leq\; (\lambda \;\; |\Omega_{\min}|)^{-1} \int_S (f(x) - \overline{f_S})^2 \, dx. \qquad \square$$

2.4 Minimizer description and isophotes

In this paper, our aim is not to investigate conditions for having global minima of the energy under concern and discuss their existence, which is quite difficult (Morel and Solimini (1994)). In what follows, we just assume existence of minimizers of the energy $E_\lambda(f, \Omega_1, \cdots, \Omega_P, \overline{\Omega})$ among functions of sets of finite perimeter (or of bounded variation (noted BV)). Our estimator is defined by (when exists)

$$(\widehat{\Omega}_1, \ldots, \widehat{\Omega}_{\widehat{P}}, \widehat{\overline{\Omega}}) \;=\; \text{argmin}_{0 \leq P \leq T} \, \text{argmin}_{\{\Omega_1, \ldots, \Omega_P\} \in \mathcal{C}_P} \, E_\lambda(f, \Omega_1, \ldots, \Omega_P, \overline{\Omega}) \tag{9}$$

Hence, we propose the following lemma:

LEMMA 2. *If there exist minimizers and that infinitesimal variations in the neighborhood of the minimal solution do not introduce topological changes, then the set of curves that globally minimizes the energy is a subset of level lines of f:*

$$f_{|\partial \widehat{\Omega}_i} \;\equiv\; \mu_i, \qquad i = 1, \ldots, \widehat{P} \; .$$

i.e. the boundary $\partial \widehat{\Omega}_i$ of each $\widehat{\Omega}_i$ is a boundary of a connected component of a level set of f.

We defer the proof of Lemma 2 to the appendix and study a situation where the perturbation is connected. Finally, Equation (35) (see Appendix) states a *necessary condition* which is essential to prove that a subset of level lines is a minimizer of the energy. In this approach, the family of possible partitions, based on level lines which do not cross each other's, includes local and global minimizers. Indeed, placing a level curve inside S corresponds to select a local or a global minimum of the energy. A global minimizer (not unique) can be computed by supervising all combinations of boundaries of connected components. The rest of the paper is devoted to the computation a subset of level curves that minimizes the segmentation energy. In particular, the computing of energies for all possible partitions may produce an expensive computational cost. Accordingly, we propose a fast stepwise greedy algorithm (see Section 2.5) to select a suboptimal configuration of objects corresponding to a local minima of the energy. We basically consider further that a connected component is an object Ω_i: each bounded connected component has a topological border that is composed of edgels called *lines lines* ; a level line separates the plane into two disjoint connected parts, it bounded interior and its unbounded exterior ; Ω_i is comprised in the interior of one of the level lines.

2.5 A stepwise algorithm for image segmentation

This section describes our algorithmic procedure for object boundaries estimation using the result described above. We discuss issues that have arisen in converting the theory to practice for our applications. We present a formal description of the method starting with a description of the input parameters of the algorithm. Our recommendations for the concrete choice of these parameters as well as the default choices used in our simulations and applications are collected in this section. The algorithm we propose is automatic and does require neither the number of regions nor any initial average values for regions and background.

Level sets and object boundaries The key ingredient of the procedure is the construction of objects whose boundaries are isophotes in the image (Caselles *et al.* (1999)). In practical imaging, both the domain S and the range of f are discrete sets. We recall that a discrete gray-scale image can be interpreted as a nested family of level sets (Matheron (1975)), computed by simple thresholding. In the today's technology, we can traditionally associate with an image 256 level sets $\{L_\gamma(f)\}$, $\gamma \in \{0, 1, 2, \cdots, 255\}$. Let γ be a fixed level of the image, $0 \leq \gamma \leq 255$, and let u_γ be the binary image at level γ of the discretized original image f, defined by $u_\gamma(x) = 1$ if $f(x) \geq \gamma$ and $u_\gamma(x) = 0$ otherwise. A crude way to build pixels sets corresponding to objects would be to proceed to a connected components labeling (Monasse and Guichard (2000)) of binary images $\{u_\gamma\}$, $0 \leq \gamma \leq 255$, and to associate each label with an object Ω_i. Its boundary $\partial\Omega_i$ would be the border of the connected component within the image level sets (Caselles *et al.* (1999)).

Instead of computing all the 256 level sets, we restrict this computation to a small number of $L(< 256)$ level sets and uniformly quantize the image histogram. Now, we consider the scenario where a point x belongs to one single connected component at once within the image level sets. We take into account this fact and define the bilevel sets of f with levels v and w, $0 \leq v \leq w$, as the set of pixels $x \in S$ such as $v \leq f(x) \leq w$. For $l \in \mathbb{N}$ varying from 1 to L, let b_l be the binary image with $b_l(x) = 1$ if $f(x) \in [t_{l-1}, t_l[$ and $b_l(x) = 0$ otherwise, where t_l is a threshold ($t_l \in [f_{\min}, f_{\max}]$). We call those images L-bilevel sets of f (Alvarez *et al.* (1999)). Each bilevel image represents a quantization level of the original image. In general, each bilevel set is made up of $n(t_l)$ connected components, where $n(t_l)$ is a function of the threshold t_l. Notice that the connected components $\{\Omega_{t_l,1}, \Omega_{t_l,2}, \cdots, \Omega_{t_l,n(t_l)}\}$, $1 \leq l \leq L$ are disjoint and their union is the image domain S:

$$\bigcup_{l=1}^{l=L} [\Omega_{t_l,1} \bigcup \Omega_{t_l,2} \bigcup \cdots \bigcup \Omega_{t_l,n(t_l)}] = S. \tag{10}$$

The connected components $\{\Omega_i\}$ of level sets can be characterized by their surrounding curves $\{\partial\Omega_i\}$, that is the level lines (Caselles *et al.* (1999); Alvarez *et al.* (1999)). If we map these level lines for a given set of L levels, we get a segmentation of the image also called *topographic map* (Caselles *et al.* (1999); Froment (2000).) Recently, Monasse and Guichard proposed a fast discrete algorithm to compute a *topographic map* using a contrast invariant tree representation

of connected components (Monasse and Guichard (2000)). As made clear in (Caselles *et al.* (1999)), the topographic map is the basic structure of the image. More generally, one can consider a segmentation achieved using only some connected components of level sets, which is the philosophy of our approach. The most perceptible level lines can be determined by an isoperimetric criterion (Froment (2000)) or the detection of T-junctions of level lines (Caselles *et al.* (1999)). Both criteria are strong indicators of region boundaries.

Instead, we define herein perceptually significant level lines as the level sets boundaries of an uniformly quantized image by using L quantizers. As a consequence, the detection of meaningful level lines will depend on the quantization parameter L. Unlike previous criteria (Caselles *et al.* (1999); Froment (2000)), this quantization operation is not invariant to contrast changes. Nevertheless, we shall see that, in practice, $L = \{4, \ldots, 8\}$ seems sufficient to detect physically meaningful objects with large areas in the image.

The procedure The proposed algorithm is not a region growing algorithm as described in (Koepfler *et al.* (1994); Morel and Solimini (1994); Beaulieu and Goldberg (1989); Pavlidis and Liow (1990)) since all objects are built once and for all. It differs from the wathershed approach since regions that emerge from the watershed segmentation are not necessarily connected components within the image level sets (Vincent and Soille (1991)). In addition, the watershed approach may include an over-segmentation and computational expense. Nevertheless, over-segmentation can be avoided by using non-linear diffusion as a preprocessing step (Weickert (1998b)) and parallel real-time implementations of the watershed algorithm exit. In our approach, we prevent over-segmentation by choosing a small number L of quantization levels and post-processing the connected components to remove any components whose surface area $|\Omega_i|$ is less than some threshold $|\Omega_{\min}|$ (a parameter of the method). This parameter is commonly used in image segmentation (Salembier and Serra (1995); Acton and Mukherjee (2000)) to eliminate regions corresponding to noise and artifacts in the original image. To implement our level set image segmentation based on energy minimization, a four stages method is used. Let L, λ, $|\Omega_{\min}|$ be the input parameters set by the user.

1. BILEVEL SET CONSTRUCTION The first step completes a crude mapping of each image pixel on a given bilevel set. At present, we uniformly quantize the function f in $L = 4$, 8, 16 or 32 equal-sized and non-overlapping intervals $[t_{l-1}, t_l[$, $l = \{1, \cdots, L\}$. Given this set of intervals, let b_l be the bilevel set image with $b_l(x) = 1$ if $f(x) \in [t_{l-1}, t_l[$ and $b_l(x) = 0$ otherwise.

2. OBJECT EXTRACTION A crude way to build pixels sets corresponding to objects is to proceed to a connected components labeling of bilevels image b_l and to associate each label with an object Ω_i. Though this process may work in the noise-free case, in general we would also need some smoothing effect of the connected components labeling. So we consider a size-oriented morphological operator acting on sets that consists in keeping all connected components of the output of area larger than a limit $|\Omega_{\min}|$. This area operator does not

introduce new features or edges and boundaries of connected components are preserved (Salembier and Serra (1995); Acton and Mukherjee (2000)). The list of remained connected components then forms the bank $\mathcal{C}_T$ of admissible T objects $\{\Omega_1, \ldots \Omega_T\}$ $(P \leq T)$ such as $|\Omega_i| \geq |\Omega_{\min}|$. The connected components of area lower than $|\Omega_{\min}|$ are a part of the background $\overline{\Omega}$.

3. CONFIGURATION DETERMINATION The connected components are then combined during the third step to form object configurations. For instance, these configurations can be built by enumeration of all possible object combinations, i.e. 2^T configurations. Each configuration is made of a subset of objects taken in the bank $\{\Omega_1, \ldots \Omega_T\}$ The background $\overline{\Omega}$ corresponds to the complementary set of objects selected for each configuration. Each possible configuration can then be indexed by a binary number $\mathtt{b}_i$ which is the binary expansion of i $(0 \leq i \leq 2^T)$. The binary value of each bit in $\mathtt{b}_i$ determines the presence or absence of a given object in the configuration (see Fig. 1).

4. ENERGY COMPUTATION AND OBJECT CONFIGURATION SELECTION Energy calculations take the image intensities of the original (not quantized) image to establish piecewise-constant approximation errors. Energies of the form $\{\int_{\Omega_i}(f(x) - \overline{f}_{\Omega_i})^2 \, dx\}$ are computed once and stored in RAM memory. The energy term $\int_{\overline{\Omega}}(f(x) - \overline{f}_{\overline{\Omega}})^2 \, dx$ is efficiently updated for each configuration since $\overline{\Omega}$ is the complementary subset of the union of objects $\{\Omega_i\}_{i=1}^{P}$. The configuration that globally minimizes the energy functional corresponds to the optimal segmentation. The time necessary to perform image segmentation essentially depends on the size of the object bank $\mathcal{C}_T$, i.e. the number T of registered connected components. Nevertheless, all configurations are independent and could be potentially evaluated on suitable parallel architectures.

Computational issues Now we discuss how some parameters of the procedure can be selected and indicate one possible choice used in our experimental results.

Connectivity We define object boundaries as connected components borders within the image level sets. Two pixels are said to be connected when they are neighbors according to some neighborhood relationship. On the discrete domain S (with rectangular tessellation) the neighborhoods of a pixel x are typically defined via 4-connectivity or 8-connectivity.

Number of bilevel sets The value of L is mainly determined by the number of meaningful objects that one wishes to extract and the computational effort one is able to spend. Decreases L allows to reduce the number of connected components and process a small bank of T objects. We propose herein a method for mapping a set of pixels to a small set of levels such that each connected component forms a relatively large and meaningful region. We select the optimal configuration of objects by supervising a small set of levels (Ishikawa and Geiger (1998)). In practice, our approach successfully segmented various images into only 4 or 8 levels. Presenting

typically level lines with levels multiple of a fixed amount say $\frac{f_{\max}-f_{\min}}{L}$ will preserve all edges of an uniformly quantized image with L quantizers.

Minimal size of objects This area operator affects the image by remaining connected components within the image level sets that do not satisfy the minimum area criterion (Salembier and Serra (1995); Acton and Mukherjee (2000)). Boundaries of connect components are not distorted by this operator since it does not incorporate any shape of structuring element on the processed image. We post-process the connected components to eliminate patches corresponding to noise in the original date. Our default choice is $|\Omega_{\min}| \in [0.0001 - 0.015] \times |S|$.

Hyperparameter λ The choice of this parameter determines mostly the properties of the segmentation result. Increasing this parameter reduces the final number of objects to be extracted. If f is a function from S to $[0, 255]$, a default choice for the hyperparameter is $\lambda \in [0.001 - 1.] \times 255^2$. Of course larger values of λ lead to extraction of only one object. In practice, $\lambda = 0.01 \times 255^2$ provides a reasonable compromise for most cases. However, we keep a possibility to tuning this parameter in some specific situations depending on what is important in each particular case.

Energy minimization For a fixed bank $\mathcal{C}_T = \{\Omega_1, \cdots, \Omega_T\}$ of T objects, one way to choose the optimal set of objects $\{\widehat{\Omega}_1, \cdots, \widehat{\Omega}_{\widehat{P}}\}$, $\widehat{P} \leq T$ defined by (9), is to search for all possible combinations of P objects and compute the corresponding energy $E_\lambda(f, \Omega_1, \cdots, \Omega_P, \overline{\Omega})$. Then, by comparing the energies, we can see which collection of objects is the best. Enumerating all possible sets of objects in the object bank and comparing their energies is computationally too expensive if T is large (typically, it is infeasible if $T > 32$). Instead of a such a brute force search, only used in experiments when $T \leq 20$, we propose the following stepwise greedy algorithm for minimizing $E_\lambda(f, \Omega_1, \cdots, \Omega_P, \overline{\Omega})$.

We start from $P = 0$ and introduce one object Ω_j at a time. Energies of all objects are assumed to be already stored in a RAM memory. At the first step, we compute the T energies with one single object Ω_j at once against the complementary subset $\overline{\Omega} = S \setminus \cup_{j \neq i=1}^{T} \Omega_i$. Let $\widehat{\Omega}_1$ the estimated object that best lowers E_λ. This object is stored on a RAM memory as an object of the optimal configuration. It is removed from the initial bank of objects $\mathcal{C}_T$. At any steps of the algorithm, a new estimated object is chosen to maximally decrease the energy E_λ.

Suppose that at the P-th step, $\widehat{P}$ and $\widehat{\overline{\Omega}}$ are not known but we have estimated P objects $\{\widehat{\Omega}_1, \cdots, \widehat{\Omega}_P\}$ and a current background $\overline{\Omega} = S \setminus \{\widehat{\Omega}_1, \cdots, \widehat{\Omega}_P\}$. Let $E_\lambda(f, \widehat{\Omega}_1, \cdots, \widehat{\Omega}_P, \overline{\Omega})$ the current computed energy. Then at the $(P+1)$-th step, for each object $\Omega_j \in \mathcal{C}_T \setminus \{\widehat{\Omega}_1, \cdots, \widehat{\Omega}_P\}$, we calculate the following energy difference:

$$\Delta E_\lambda(f, \Omega_j, \overline{\Omega}, \overset{\circ}{\overline{\Omega}}) \;=\; E_\lambda(f, \widehat{\Omega}_1, \cdots, \widehat{\Omega}_P, \overline{\Omega}) - E_\lambda(f, \widehat{\Omega}_1, \cdots, \widehat{\Omega}_P, \Omega_j, \overset{\circ}{\overline{\Omega}}) \tag{11}$$

FIGURE 2 TO BE PLACED HERE

where $\overline{\Omega} = \overset{\circ}{\overline{\Omega}} \cup \Omega_j$. Intuitively, we choose the object which has the maximal difference, i.e.,

$$\widehat{\Omega}_{P+1} = \arg \max_{\Omega_j \in \mathcal{C}_T \setminus \{\widehat{\Omega}_1, \cdots, \widehat{\Omega}_P\}} \Delta E_\lambda(f, \Omega_j, \overline{\Omega}, \overset{\circ}{\overline{\Omega}}). \quad (12)$$

The algorithm stops at P-th step when the adding of any object does not decrease E_λ. This means that the optimal number of objects is $\widehat{P} = P$ and the remained objects of the bank are a part of the estimated background, i.e. $\widehat{\overline{\Omega}} = S \setminus \{\widehat{\Omega}_1, \cdots, \widehat{\Omega}_{\widehat{P}}\}$. In summary, we propose the algorithm of object selection described in Fig. 2.

This algorithm selects a suboptimal configuration of objects corresponding to a local minima of the energy functional. Using this algorithm, $\frac{T \times (T+1)}{2}$ object configurations are examined at the most, whereas the supervision of all the configurations correspond to 2^T iterations. The CPU-time taken by our implementation using the parameter settings is reported in Section 3.

3 Experimental results in image segmentation

This section presents experiments on synthetic images as well as real-world images. We are interested in the use of the technique in the context of optical, medical and meteorological imagery. Our system successfully segmented various images into a few regions. For the bulk of the experiments, the algorithm parameters were set in these experiments as follows: $L = 4, 8, 16$ or 32, and regions which areas $|\Omega_i| < |\Omega_{\min}|$ are discarded. The cleaning threshold $|\Omega_{\min}| \in [0.0005 - 0.015] \times |S|$ indicates the minimum area of connected components to be retained in the segmentation. Recall that obtaining the most significant objects is the goal of this work. For this reason, L was set fairly low in the experiments to obtain large regions and to improve robustness to noise and artifacts in the image. For our method, $\lambda \in [0.001 - 1.] \times 255^2$ varies across the images depending on the image content. It is set empirically and values that gave visually better results were chosen. In our experiments, only a few values (3 or 4 values) with a crude precision were tried for this parameter. The choice of λ is crucial and a special attention must be granted by the user. Nevertheless, a small perturbation of the chosen parameter does not drastically alter the segmentation results from experiments. Most segmentations took approximately about 5-30 seconds on a 296MHz workstation. Simulations were conducted on synthetic as well as real-world images to evaluate the performance of the algorithm. We start by examining the influence of the penalization parameter λ on results.

Influence of the penalization parameter λ Figure 3a shows an artificially computed 256×256 image representing the superposition of two bidimensional Gaussian functions located respectively at $x_0 = (64, 128)$ and $x_1 = (160, 128)$ with variance of $\sigma_0 = 792$ and $\sigma_1 = 1024$.

ALGORITHM 1: <<OBJECT SELECTION>>

Let $\mathcal{C}_T$ be a bank of T objects and f the observed image.
Compute the T energies of all individual objects.

Initialize $P = 0$, $E_\lambda(f, \Omega_1, \cdots, \Omega_P, \overline{\Omega}) \leftarrow E_\lambda(f, \Omega_S)$.

Repeat

- For each object $\Omega_j \in \mathcal{C}_T \setminus \{\widehat{\Omega}_1, \cdots, \widehat{\Omega}_P\}$ compute $\Delta E_\lambda(f, \Omega_j, \overline{\Omega}, \overset{\circ}{\overline{\Omega}})$.
- Choose $\widehat{\Omega}_{P+1}$ so that $\widehat{\Omega}_{P+1} = \arg \max_{\Omega_j \in \mathcal{C}_T \setminus \{\widehat{\Omega}_1, \cdots, \widehat{\Omega}_P\}} \Delta E_\lambda(f, \Omega_j, \overline{\Omega}, \overset{\circ}{\overline{\Omega}})$.
- $\{\widehat{\Omega}_1, \cdots, \widehat{\Omega}_P\} \leftarrow \{\widehat{\Omega}_1, \cdots, \widehat{\Omega}_P\} \bigcup \widehat{\Omega}_{P+1}$, $P \leftarrow P + 1$.

Until $\Delta E_\lambda(f, \Omega_j, \overline{\Omega}, \overset{\circ}{\overline{\Omega}}) < 0$, $\forall \Omega_j \in \mathcal{C}_T \setminus \{\widehat{\Omega}_1, \cdots, \widehat{\Omega}_P\}$.

Set $\widehat{P} = P$ and $\widehat{\overline{\Omega}} = S \setminus \{\widehat{\Omega}_1, \cdots, \widehat{\Omega}_{\widehat{P}}\}$.

Figure 2: Object selection algorithm.

FIGURE 3 TO BE PLACED HERE

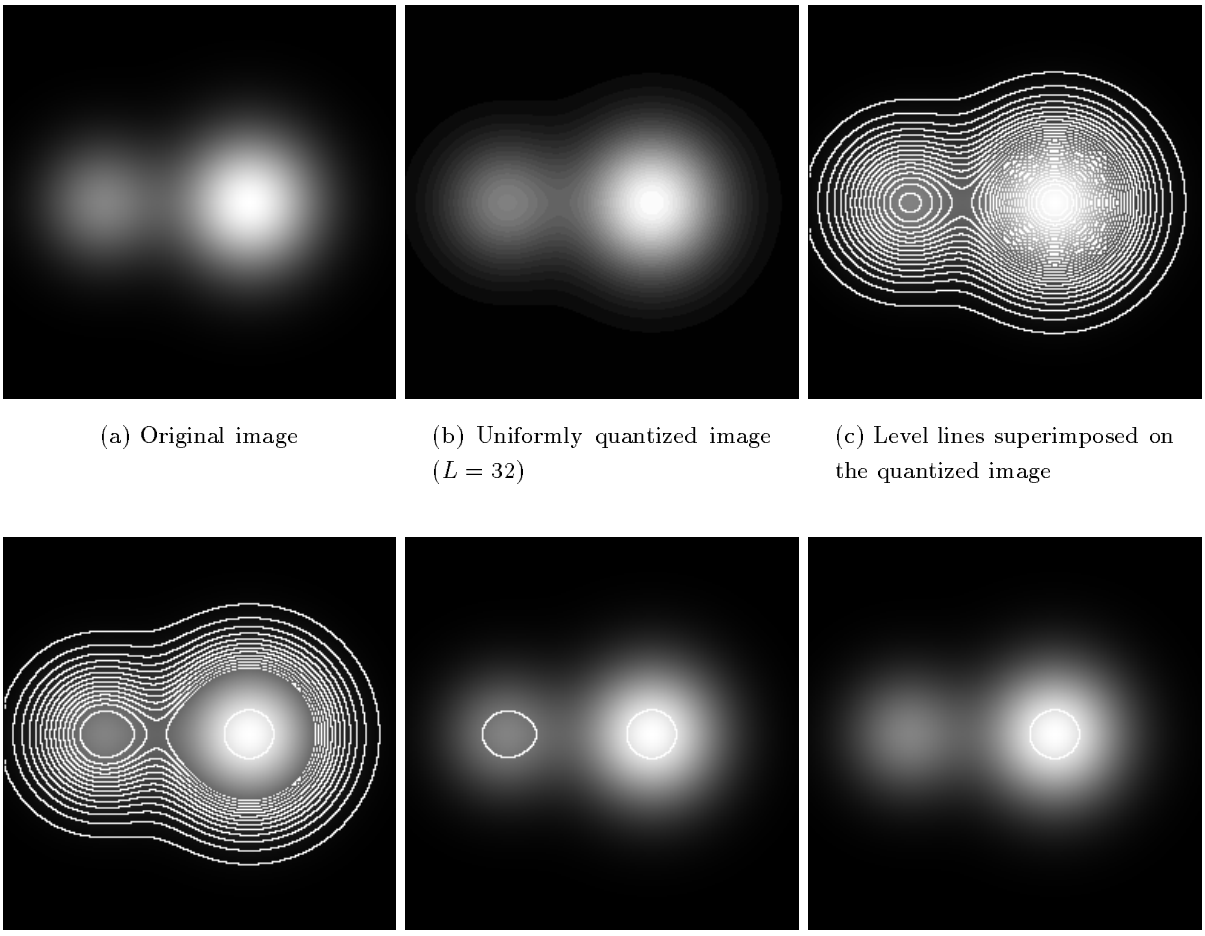
Figure 3b shows the result of the uniform quantization operation applied on Fig. 3a ($L = 32$). The levels lines associated with the quantized image are displayed on Fig. 3c. Note that level sets of area too small are suppressed. Figures 3d-f show how the penalization parameter influences the segmentation results. The white borders denote the object boundaries resulting from the segmentation.

FIGURE 4 TO BE PLACED HERE

Figure 4 illustrates how our method selects the number of segments in the "*house*" image (256×256). The first row shows, from left to right, the original image, the topographic map when $L = 4$ and $|\Omega_{\min}| = 0.0005 \times |S|$ and the image histogram. The four equal-sized intervals are displayed in Fig. 4c. In this experiment, the maximum number of admissible objects is $T = 49$. The second row shows the piecewise-constant approximation of the original image when $\lambda = 0.001 \times 255^2$, $\lambda = 0.005 \times 255^2$ and $\lambda = 0.05 \times 255^2$. The third row shows the corresponding boundary sets. Notice, even if the number of objects is low ($T = 49$), the obtained boundaries in Fig. 4 are relatively numerous since the background region is a non-connected set. A coarser segmentation ($P = 17$) is obtained in Figs. 4f and 4i by choosing a higher value for λ. It takes 12 seconds (1176 iterations) of computing time for building the object bank and selecting the best configuration using the stepwise greedy algorithm. Enumerating all the configurations is infeasible since $2^T = 5,6310^{14}$ iterations !

FIGURE 5 TO BE PLACED HERE

Dynamic segmentation process An example of cloud detection is provided in Figs. 5 and 6. For this set of parameters $L = 4$, $\lambda = 0.1 \times 255^2$ and $|\Omega_{\min}| = 0.0001 \times |S|$, the algorithm labeled seas, continents and small clouds as "*background*". The significant clouds are crudely extracted from the 383×260 image and labeled as "*objects*" (first row in Fig. 6). Figure 5 shows, from left to right, the original image, the topographic map and the piecewise constant approximation of the image if all the $T = 207$ regions are used. Figure 6 depicts four samples of the dynamic segmentation process: the piecewise-constant approximation and the boundary sets superimposed on the original image are shown at significant iterations of the stepwise greedy algorithm. Figure 6 shows the results at the 214th ($P = 2$), 2053th ($P = 11$) and the last 9781th ($\widehat{P} = 55$) iteration. The algorithm finally selects $\widehat{P} = 55$ regions and stops at the 9781th iteration (=21s of CPU time), i.e. before the maximal iteration $\frac{T \times (T+1)}{2} = 21528$. For this experiment, numerical results of the stepwise greedy algorithm are shown in Table 1. Note that the larger significant regions are first extracted before examination of small clouds.



(a) Original image

(b) Uniformly quantized image ($L = 32$)

(c) Level lines superimposed on the quantized image

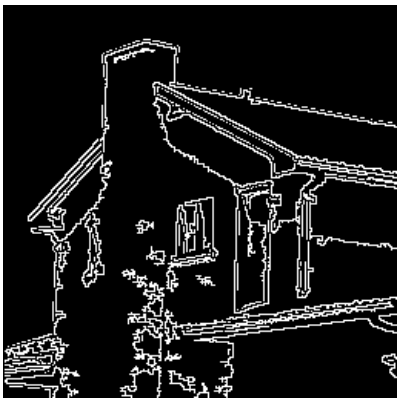
(d) Segmentation with $\lambda = 0.01 \times 255^2$

(e) Segmentation with $\lambda = 0.1 \times 255^2$ (two objects)

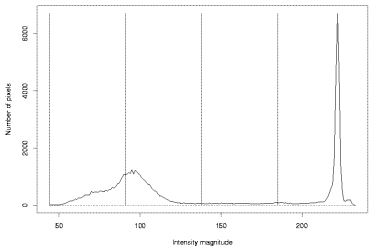
(f) Segmentation with $\lambda = 1. \times 255^2$ (one object)

Figure 3: Segmentation of a synthetic image ($L = 32$).

(a) Original image



(b) Topographic map

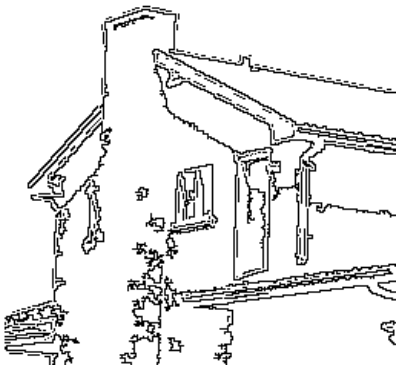


(c) Image histogram

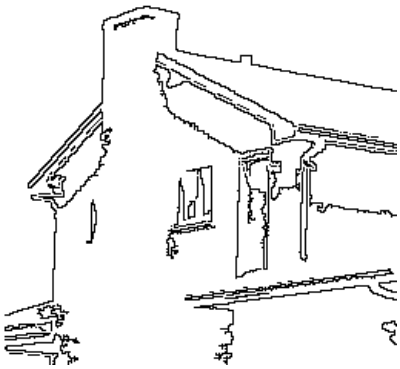
(d) Piecewise-constant approximation using $\lambda = 0.001 \times 255^2$ ($P = 49$)

(e) Piecewise-constant approximation using $\lambda = 0.005 \times 255^2$ ($P = 29$)

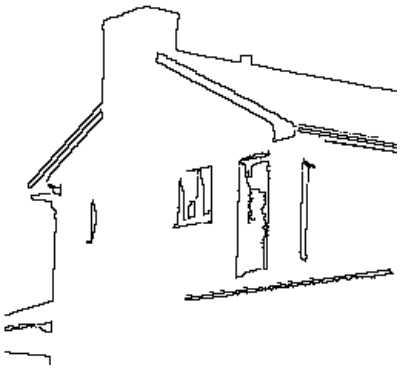
(f) Piecewise-constant approximation using $\lambda = 0.05 \times 255^2$ ($P = 17$)



(g) Boundary set using $\lambda = 0.001 \times 255^2$



(h) Boundary set using $\lambda = 0.005 \times 255^2$



(i) Boundary set using $\lambda = 0.05 \times 255^2$

Figure 4: Segmentation of the "*house*" image ($L = 4, |\Omega_{\min}| = 0.0005 \times |S|$ pixels, $T = 49$).

(a) Original image

(b) Topographic map ($L = 4, |\Omega_{\min}| = 0.0001 \times |S|$)

(c) Piecewise-constant approximation ($T = 207$)

Figure 5: Level lines of a meteorological image.

TABLE 1 TO BE PLACED HERE

FIGURE 6 TO BE PLACED HERE

Exhaustive search vs. deterministic minimization To go further in the comparisons, we can focus on the results we obtain by minimization of (2) using the stepwise greedy algorithm (Fig. 2) to those obtained by examining all the object configurations. The performance of the minimization procedure is demonstrated for a MR image (256×228) where the flat black background has been previously eliminated during a pre-processing step. The first row in Fig. 7 shows, from left to right, the original image, the image histogram and the topographic map. In order to run the algorithms, we experimentally adjust the parameters: $L = 8$ and $|\Omega_{\min}| = 0.0025 \times |S|$ pixels. For this set of parameters, the total number of objects is $T = 25$. The aim is to segment both the corpus callosum and brain by selecting $\lambda = 0.5 \times 255^2$. In Fig. 7, the second row shows the convergence to the global minima of the energy function using both optimization algorithms. In this experiment, the suboptimal stepwise greedy algorithm converges to a local minima corresponding to $P = 5$ objects after $276 < \frac{(T+1)\times T}{2}$ iterations and $6.4s$ of CPU time on a 296 Mhz workstation. We need $395s$ to evaluate all the $2^T = 33554432$ configurations and obtain the same segmentation result shown in Fig. 7. In addition, the time necessary for building the bilevel sets and extracting the connecting components is about $5s$ in most experiments. We are not claiming that this algorithm always converges to the global minima since many toy-examples can be designed to contradict this fact. However, this heuristic approach seems adequate in most cases and then allow us to manipulate many objects in natural images.

FIGURE 7 TO BE PLACED HERE

Mumford and shah functional In these experiments, we compare our segmentation results to those provided by a region growing method using the simplified version of the Mumford and Shah model which aims at approximating a given image with piecewise constant functions (Koepfler *et al.* (1994); Morel and Solimini (1994)). The prior model ensures the discontinuity set has a small length (Mumford and Shah (1989)). A satisfying reconstruction using one hundred regions is shown in Figs. 8c and 8d. Our approach produces comparative visual results using 29 regions and a non-connected background (see Fig. 4e). In comparison, we run the region-growing algorithm to extract only 24 regions. Figs. 8 and 8b displays the crudely piecewise-constant approximation result.

FIGURE 8 TO BE PLACED HERE

P	Iteration	E_λ	$\sum_{i=1}^{P} \lvert\Omega_i\rvert$	P	Iteration	E_λ	$\sum_{i=1}^{P} \lvert\Omega_i\rvert$
1	5	8804	15864	29	5420	7808	32378
2	214	8360	22185	30	5610	7807	32402
3	420	8089	27670	31	5785	7806	32424
4	662	7983	29051	32	5960	7805	32442
5	832	7928	30082	33	6140	7804	32456
6	1028	7902	30418	34	6313	7803	32473
7	1229	7887	30676	35	6481	7802	32491
8	1449	7877	30838	36	6661	7801	32506
9	1630	7868	31026	37	6834	7800	32520
10	1844	7862	31137	38	7003	7799	32531
11	2053	7855	31230	39	7170	7798	32545
12	2258	7850	31293	40	7341	7798	32556
13	2431	7846	31362	41	7508	7797	32567
14	2620	7843	31415	42	7674	7796	32579
15	2817	7839	31502	43	7834	7796	32591
16	3019	7836	31569	44	8007	7795	32616
17	3203	7833	31629	45	8167	7794	32628
18	3401	7830	31671	46	8352	7794	32663
19	3602	7827	31713	47	8490	7793	32674
20	3783	7825	31746	48	8651	7793	32685
21	3977	7822	31788	49	8824	7792	32720
22	4147	7820	31833	50	8969	7792	32731
23	4352	7818	32194	51	9142	7791	32747
24	4530	7816	32222	52	9286	7791	32765
25	4694	7814	32260	53	9457	7791	32782
26	4881	7812	32293	54	9614	7790	32793
27	5069	7811	32323	55	9781	7790	32809
28	5248	7809	32352				

Table 1: Object selection at each step of the stepwise greedy algorithm.

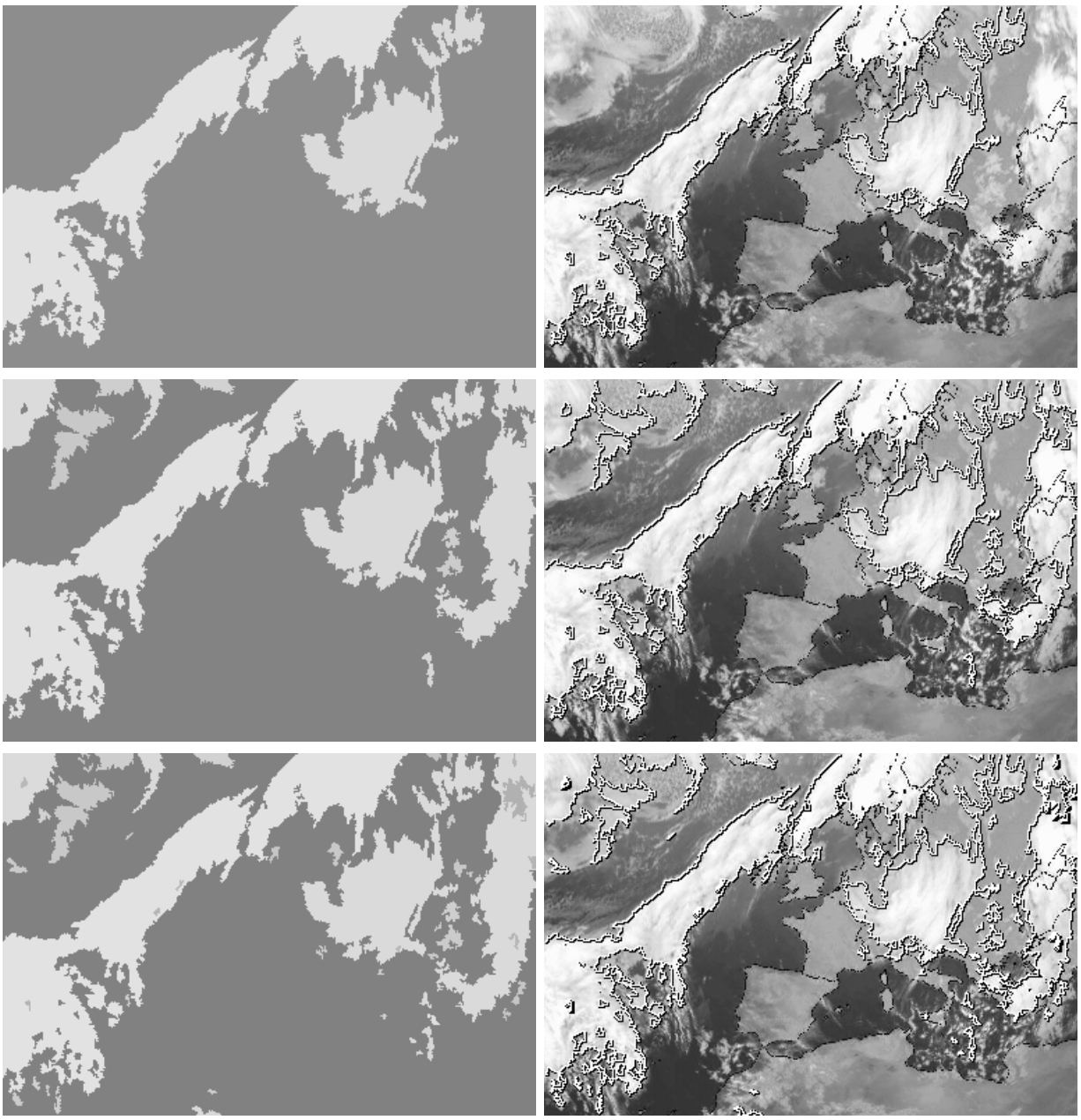
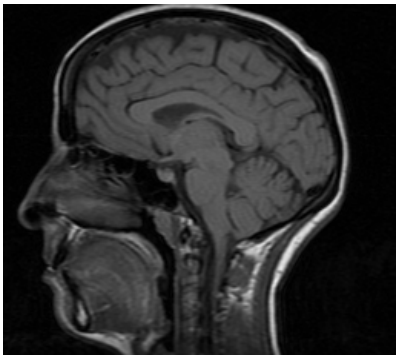
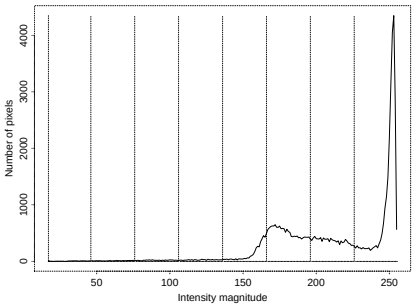


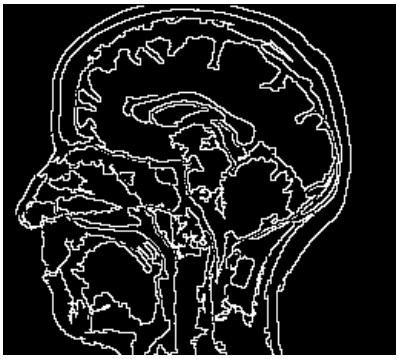
Figure 6: Segmentation of a meteorological image ($\lambda = 0.1 \times 255^2$). Left column (top to bottom): piecewise-constant approximation at iteration 214 ($P = 2$), 2053 ($P = 11$) and 9781 ($P = 55$). Right column (top to bottom): boundary sets superimposed on the original image at iteration 1214, 2053 and 9781.



(a) Original image

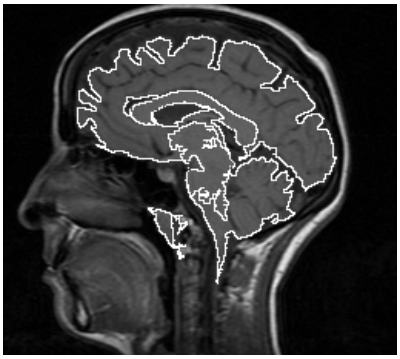


(b) Image histogram



(c) Topographic map ($L = 8$)

(d) Segmentation map



(e) Boundary set superimposed on the image

Figure 7: Segmentation results of a MR image with $\lambda = 0.5 \times 255^2$.

(a) Piecewise-constant approximation using 29 regions

(b) Boundary set of 29 regions

(c) Piecewise-constant approximation using 100 regions

(d) Boundary set of 100 regions

Figure 8: Segmentation of the "*house*" image using the simplified Mumford and Shah model (Koepfler *et al.* (1994); Morel and Solimini (1994)).

4 Reaction-diffusion model

The results of Section 2 and Section 3 suggest that the forms of connected components in noisy images are not generally smooth. This section deals with object boundaries-preserving isophotes regularization. Numerous partial differential equations-based algorithms have been proposed recently to tackle problems of noise removal, image enhancement and image restoration (Perona and Malik (1990); Nordström (1990); Alvarez *et al.* (1992); Catte *et al.* (1992); Alvarez *et al.* (1993); Alvarez and Mazorra (1994); Ter Haar Romeny (1994); Rougon and Preteux (1995); Charbonnier *et al.* (1997); Malladi and Sethian (1996); Kornprobst *et al.* (1997); Weickert (1998a); Caselles *et al.* (1998)). They have been recently injected into the segmentation problem since they can be cast as the minimization of an energy functional (Morel and Solimini (1994); You *et al.* (1996)). With this variational approach, edges are traditionally unknown and must be detected at the same time as the object is reconstructed and regularized.

The method we present consists in smoothing the input image to impose some length on level curves of the image. We also intend to keep the object boundaries, encourage intra-region diffusion and find an oversmoothed representation of the background in the input image. It is clear that such requirement will be fulfilled when we use an a priori knowledge about objects to generate a geometry-driven reaction-diffusion scheme that reflects the characteristics of a given spatial position in the image. For that reason we have chosen the anisotropic diffusion of Perona and Malik (Perona and Malik (1990)) which will not blur or move the edges, where an additional reaction term forces the solution to remain close to the data in objects and to be a constant in the background. Due to this modeling, the coefficient of spatial diffusion will be designed to smooth adaptively the foreground, background and edges. Moreover, the method which consists in moving the iso-intensity contours in an image under curvature dependent speed laws is not appropriate: the mean curvature operator (Alvarez *et al.* (1992)) is free parameter and therefore enables to easily encourage diffusion inside regions and discriminate background and foreground regions over time. The background region cannot be labeled after each iteration of this diffusion process since all the iso-intensity contours are similarly smoothed according to curvature. Nevertheless, the classification between foreground and background can be performed independently of the diffusion process in a post-processing step. In our approach, the luminance statistics inside level lines are examined after each iteration to estimate the coefficients of spatial diffusion inside and outside regions. When the original image is used as the initial condition, this process produces over time, a set of smooth regions, a background with a nearly constant luminance separated by well located and sharp edges.

4.1 Related works

Perona-Malik equation *Anisotropic diffusion* as proposed by Perona and Malik (Perona and Malik (1990)) is an alternative to the linear scale-space described by the heat equation. Intuitively, the idea of non-homogeneous diffusion is to limit the flow of intensity according to local gradient information. This would result in the blurring of details within an object while

keeping the edges sharp, producing a *cartoon* of the image. Basically, the idea is to evolve from an original image $f_0(x)$, a family of increasingly smooth images $f(x,t)$ derived from the solution of the following partial differential equation (Perona and Malik (1990)):

$$\begin{cases} \dfrac{\partial f(x,t)}{\partial t} & = \text{ div } (g(|\nabla f(x,t)|)\, \nabla f(x,t)), \\ \left.\dfrac{\partial f(x,t)}{\partial n}\right|_{\partial S} & = 0, \\ f(x,0) & = f_0(x) \end{cases} \tag{13}$$

where n denotes the normal to the image boundary ∂S, t is the time or evolution parameter, ∇ the gradient operator, div the divergence operator and $g(\cdot)$ the conductance modulating term which is a bounded, positive, decreasing and tending to zero at infinity as a function of $|\nabla f|$. For this reason $g(\cdot)$ is also called *edge-stopping* function. Perona and Malik suggested an *edge-stopping* function $g(|\nabla f(x,t)|) = \left(1 + \frac{|\nabla f(x,t)|^2}{2\sigma^2}\right)^{-1}$ which depends explicitly on the gradient magnitude of the function itself: equation (13) has the effect of limiting blurring near edges and also increasing the steepness or gradient of edges that are sufficiently steep in the original image f_0. In practice, areas in which the gradient magnitude is lower than the scale parameter σ will be blurred more strongly than areas with a higher gradient magnitude. This tends to smooth noise and unwanted texture regions, while preserving the edges between different regions. To remove undesirable "staircase" artifacts, we can iterate these equations but, although the noise is eradicated, the edges are softened.

Let us now consider the simplest discrete approximation of this process. We consider discrete times $t \in \mathbb{N}$ and a time step $\triangle t$. The standard anisotropic diffusion equation is known to present potential numerical instability. However this numerical instability is not observed by adopting the discrete anisotropic diffusion update proposed in (Perona and Malik (1990)):

$$f(x,t+1) \;=\; f(x,t) \;+\; \frac{\triangle t}{|\mathcal{N}_x|} \sum_{x^- \,\in\, \mathcal{N}_x} g(\delta(f,x,x^-,t),\sigma)\; \delta(f,x,x^-,t) \tag{14}$$

where $\mathcal{N}_x$ represents the first-order spatial neighborhood of x, $|\mathcal{N}_x|$ is the number of directions in which diffusion is computed (typically $|\mathcal{N}_x| = 4$) and σ is a scale parameter. The directional derivative $\delta(f,x,x^-,t)$ (not to be confused with gradient operator $\nabla f(x,t)$) located at x can be defined as a simple difference on the East, North, West and South directions:

$$\delta(f,x,x^-,t) \;=\; f(x^-,t) - f(x,t), \quad x^- \in \mathcal{N}_x. \tag{15}$$

This discretization, also used in our own framework, has been widely referred to as anisotropic diffusion and we will refer hereafter to this implementation.

Robust anisotropic diffusion It has been shown that anisotropic diffusion models derive from functionals of the form (Geiger and Yuille (1991); You *et al.* (1996))

$$E(f) \;=\; \int_S \rho(|\nabla f(x)|,\sigma)\, dx. \tag{16}$$

$E(f)$ specifies a *generalized membrane* defined by an adaptive potential $\rho(\cdot)$ related $g(\cdot)$ by

$$g(|\nabla f(x)|, \sigma) \;=\; \frac{\rho'(|\nabla f(x)|, \sigma)}{|\nabla f(x)|}. \tag{17}$$

This form of energy can be interpreted in the sense of robust statistics where the diffusion process aims at estimating a piecewise smooth image from noisy input image (Black *et al.* (1998)). In this robust statistical framework, edges are interpreted as outliers and σ as a scale parameter that will adaptively reduce unwanted noise when the local statistics properties of the image are Gaussian and will preserve edges when the gradient is high.

In (Black *et al.* (1998)), it is established that the function $g(\cdot)$ is proportional to the Lorentzian error norm used in robust statistics. Nevertheless, experimental results showed that the *Tukey's biweight* ρ-error function produces sharper discontinuities than the original *Lorentzian* function (Black and Rangarajan (1996)). We have also used in our anisotropic smoothing context the *edge-stopping* function derived from *Tukey's biweight* ρ-error

$$g(|\nabla f(x)|, \sigma) \;=\; \begin{cases} \left(1 - \dfrac{|\nabla f(x)|^2}{\sigma^2}\right)^2 & \text{if } \; |\nabla f(x)| \leq \sigma, \\ 0 & \text{otherwise.} \end{cases} \tag{18}$$

The scale parameter σ can be directly estimated in a robust way (see Rousseeuw and Leroy (1987)) from the image gradients. Intuitively σ should characterize the variance of the majority of the data within a region. In (Black *et al.* (1998)), Black *et al.* took the region for computing $\sigma = \sqrt{5}\,\tilde{\sigma}$ to be the entire image where

$$\tilde{\sigma} \;=\; 1.4826\, \text{median}_f \big| |\nabla f(x)| - \text{median}_f(|\nabla f(x)|) \big|. \tag{19}$$

This follows from the fact that the median value of the absolute values of a large enough sample of unit-variance normal distributed one-dimensional values is $0.6745 = 1/1.4826$. This approach works well when edges are distributed homogeneously across the image but this is rarely the case. The authors recently explored the computation of this measure in image patches to prevent the amplification of undesirable texture gradients and noise in homogeneous image regions (Black and Sapiro (1999)). However, information about the image structure and the data generation process can be valuable to improve tasks such as noise reduction and image segmentation.

Variational methods for denoising images Coming from the optimization approach for image denoising, it is also interesting to add a reaction term which asserts that we must remain close to the original image (Nordström (1990); Morel and Solimini (1994); Catte *et al.* (1992); Charbonnier *et al.* (1997); Kornprobst *et al.* (1997); Caselles *et al.* (1998)). Nordström has proposed to modify purely diffusive models by introducing a bias term constraining the solution to remain close to the initial condition f_0. This reaction term avoids the selection of a stopping time and termination of the diffusion at a trivial solution, such as a constant image (Nordström

(1990)).The biased anisotropic diffusion is defined by

$$\begin{cases} \dfrac{\partial f(x,t)}{\partial t} &= [f_0(x) - f(x,t)] \ + \ \alpha \ \mathrm{div} \ (g(|\nabla f(x,t)|, \sigma) \ \nabla f(x,t)), \\ \left.\dfrac{\partial f(x,t)}{\partial n}\right|_{\partial S} &= 0, \\ f(x,0) &= f_0(x). \end{cases} \tag{20}$$

Obviously, the left-hand side $[f_0(x) - f_t(x)]$ enforces an additional constraint that penalizes deviation form the input image. The intermediate solutions of (20) may be regarded as a descent on this following energy functional (Morel and Solimini (1994)):

$$E(f) \ = \ \int_S (f - f_0)^2 \, dx \ + \ \frac{\alpha}{2} \ \int_S \rho(|\nabla f(x)|, \sigma) \, dx. \tag{21}$$

The first summand encourages similarity between the restored image and the original one, while the second summand rewards smoothness. The smoothness weight $\alpha > 0$ is called regularization parameter. Choosing a non-convex potential does not guarantee well-posedness and stable algorithms, but (20) however converges globally to the steady-state of the regularization process. Notice that computing $g(\cdot)$ using some regularized first order derivatives (e.g. Gaussian derivatives) leads to a well-posed PDE, even for non-convex ρ-function (Catte *et al.* (1992)).

4.2 Reaction-diffusion model for background simplification

In addition to the classical anisotropic diffusion given in Section 4.1, other diffusion mechanisms may be employed to adaptively filter an image for image segmentation purposes. For instance, the idea of Alvarez and Esclarin (1997) is to propose a reaction-diffusion partial differential equation which attracts the image over time to a finite number of specified grey levels. This approach aims at adaptively quantizing the image using a variational method. The philosophy of our approach is similar to (Alvarez and Esclarin (1997)) in a certain point of view but the methodologies are quite separate. We follow hereafter the philosophy of Nordström (Nordström (1990)) by incorporating a reaction term to penalize adaptively the background and foreground regions. The additional reaction term forces the solution to remain close to the data in objects and to be a constant in the background. Hnece, one can attempt to solve the following set of more sophisticated partial differential equations:

$$\begin{cases} \dfrac{\partial f(x,t)}{\partial t} &= [f_0(x) - f(x,t)] \ + \ \alpha_\Omega \ \mathrm{div} \ (g(|\nabla f(x,t)|, \sigma_\Omega(t)) \ \nabla f(x,t)), \quad \text{if } \ x \in S/\overline{\Omega}, \\ \dfrac{\partial f(x,t)}{\partial t} &= [\overline{f}_{\overline{\Omega}}(t) - f(x,t)] \ + \ \alpha_{\overline{\Omega}} \ \mathrm{div} \ (g(|\nabla f(x,t)|, \sigma_{\overline{\Omega}}(t)) \ \nabla f(x,t)), \quad \text{if } \ x \in \overline{\Omega}, \\ \left.\dfrac{\partial f(x,t)}{\partial n}\right|_{\partial S} &= 0, \\ \left.\dfrac{\partial f(x,t)}{\partial n}\right|_{\partial \Omega_i} &= 0, \\ f(x,0) &= f_0(x) \end{cases} \tag{22}$$

where $\overline{\Omega}$ is the background, $\Omega = S \setminus \overline{\Omega}$ is the union of objects, that is the foreground, $\overline{f}_{\overline{\Omega}}(t)$ is the average value of f over $\overline{\Omega}$ at iteration t, α_Ω and $\alpha_{\overline{\Omega}}$ are two distinct regularization

parameters which control the competition between the two energy terms and determine the smoothing, and $\sigma_\Omega(t)$ and $\sigma_{\overline{\Omega}}(t)$ are respectively the scale parameters related to the foreground and background at time t. We note n the normal to either the image boundary ∂S or the object boundary $\partial\Omega_i$. Notice the background region is a non-connected region since it is defined as the complementary subset of the union of objects : $\bar{\Omega} = S \setminus \cup_{i=1}^{P}\Omega_i$. It is made up of the union of several connected regions of any size and $f(x) \in [f_{\min}, f_{\max}]$ if $x \in \overline{\Omega}$. Accordingly, the variance $\sigma_{\overline{\Omega}}$ region may be large. Unlike the background, the objects Ω_i are connected regions and $l\frac{f_{\max}-f_{\min}}{L} \leq f(x) < (l+1)\frac{f_{\max}-f_{\min}}{L}$ if $x \in \Omega_i$, $i = 1, \cdots, P$ and $l = 1, \cdots, L$. These two foreground/background regions must be then treated and smoothed differently. Finally, each object could be also separately processed. However, the computational load can be prohibitive if the number P of objects is large due to the robust but computationally expensive estimation of variances (see (19) and (23)) used in our framework.

The idea behind the model given by (22) is to combine diffusion for noise filtering and a reaction term to either remain close to the initial condition f_0 inside objects or remove structures in the background region. A constant value $\overline{f}_{\overline{\Omega}}(t)$ for the background is herein desired since it is regarded as a non-informative region for image interpretation. In our case, we choose a non-convex potential $g(\cdot)$ corresponding to the *Tukey's biweight* norm. Of course, the non-convexity of the Tukey's biweight function will give an ill-posed problem. However, we use this function later in our experiments because non-convex functions leads to edge enhancement and give better results in practice. Notice that finite difference discrete schemes are known to compensate the ill-posedness of the continuous model, i.e. introduce additional regularization. We refer to (Weickert (1998b)) for a thorough discussion concerning the implementation of discrete schemes.

In this modeling, the choice of σ_Ω and $\sigma_{\overline{\Omega}}$ is then an important issue and the main concern of this section. The estimation of these scale parameters differs from previous definitions (Black *et al.* (1998); Black and Sapiro (1999)). We propose values with an explicit dependence on the spatial coordinates and the data function f. First, we interpret the standard deviations $\tilde{\sigma}_\Omega(t)$ and $\tilde{\sigma}_{\overline{\Omega}}(t)$ defined by (Rousseeuw and Leroy (1987); Black and Rangarajan (1996); Black and Sapiro (1999))

$$\begin{cases} \tilde{\sigma}_\Omega(t) &= 1.4826\ \text{median}_f||\nabla f| - \text{median}_f(|\nabla f|)|, \text{ if } \ x \in \Omega = S \setminus \overline{\Omega}, \\ \tilde{\sigma}_{\overline{\Omega}}(t) &= 1.4826\ \text{median}_f||\nabla f| - \text{median}_f(|\nabla f|)|, \text{ if } \ x \in \overline{\Omega}, \end{cases} \tag{23}$$

as the gradient magnitudes at which outliers begin to be downweighted in the background and foreground regions. Hence, we choose a value for the scale parameters so that the Tukey influence function begins rejecting outliers at the values $2\tilde{\sigma}_\Omega(t) = \sigma_\Omega(t)/\sqrt{5}$ and $2\tilde{\sigma}_{\overline{\Omega}}(t) = \sigma_{\overline{\Omega}}(t)/\sqrt{5}$ (Black *et al.* (1998)). Thus, starting form $f(x,0) = f_0(x)$, the first term of (22) diffuses adaptively the image in the background and foreground regions, while the reaction terms form either a constant background or preserve the original image. Since the computation involves determining the background and foreground regions for solving (22), a way for efficiency is to estimate a segmentation map at each iteration t and update the scale parameters $\{\sigma_\Omega(t), \sigma_{\overline{\Omega}}(t)\}$ and the

average intensity $\overline{f}_{\overline{\Omega}}(t)$. The set of parameters $\{\sigma_\Omega(t), \sigma_{\overline{\Omega}}(t), \overline{f}_{\overline{\Omega}}(t)\}$ regulates the filtering from an *a priori* knowledge of the geometry of the image (composed of objects). In the following, we fix regularization parameters α_Ω and $\alpha_{\overline{\Omega}}$. High values for $\alpha_{\overline{\Omega}}$ suggest that a simplified background is desired. This choice of these parameters has of course to depend on the image data and the desired application. This modeling improved the existing PDE-schemes: standard anisotropic diffusion are inefficient for filtering structural distortions and partial non-gaussian clutter (Zhu and Mumford (1997)).

4.3 Discrete formulation and algorithmic scheme

In the previous section, we have introduced a robust method to incorporate "fresh" information from the data structure and thus improving the behavior the diffusion process in presence of meaningful objects. In previous works (Perona and Malik (1990); Black *et al.* (1998); Black and Sapiro (1999); Weickert (1998b)), the regions for computing the scale parameter are either the entire image or rectangular patches. Besides, image smoothing and image segmentation are two independent tasks (Weickert (1998b)). We found that two distinct adaptive scale parameters for the foreground and background are a simple way to control the diffusion rate since we have

$$\begin{cases} l\,\frac{f_{\max}-f_{\min}}{L} \;\leq\; f(x) \;<\; (l+1)\,\frac{f_{\max}-f_{\min}}{L} & \text{if } x \in \Omega_i,\ i=1,\cdots,P, \\ f_{\min} \;\leq\; f(x) \;<\; f_{\max} & \text{if } x \in \overline{\Omega}. \end{cases} \tag{24}$$

The new conductance function depends not only on the local behavior of the data function f which is subject to noise, but also on the particular characteristics of the system at every location x of the image.

In what follows, we consider that, on the outside of Ω_i, f has a constant value corresponding to the value on $\partial\Omega_i$, i.e the value of a level line. Adopting the discrete anisotropic diffusion update in (Perona and Malik (1990)), diffusion may be implemented by

$$\begin{cases} f(x,t+1) = f(x,t) + \dfrac{\triangle t\,\alpha_\Omega}{|\mathcal{N}_x|} \displaystyle\sum_{x^- \in \mathcal{N}_x} g(\delta(f,x,x^-,t),\sigma_\Omega)\,\delta(f,x,x^-,t) + \triangle t[f_0(x) - f(x,t)] & \text{if } x \in S/\overline{\Omega} \\ f(x,t+1) = f(x,t) + \dfrac{\triangle t\,\alpha_{\overline{\Omega}}}{|\mathcal{N}_x|} \displaystyle\sum_{x^- \in \mathcal{N}_x} g(\delta(f,x,x^-,t),\sigma_{\overline{\Omega}})\,\delta(f,x,x^-,t) + \triangle t[\overline{f}_{\overline{\Omega}}(t) - f(x,t)] & \text{if } x \in \overline{\Omega} \end{cases} \tag{25}$$

where $g(\delta(f,x,x^-,t),\sigma) \in [0,1]$ given in (18) indicates the presence ($g(\cdot)$ close to 0) or absence ($g(\cdot)$ close to 1) of discontinuities or outliers inside patches (Black *et al.* (1998)). Traditionally, the number of iterations in the discrete implementation is set interactively by visually inspecting the results. The discrete scheme is a naive scheme and probably a more performant scheme would be suitable. Nevertheless, at $t=0$, a crude segmentation is obtained by selection of a subset of level lines. According to (15) and (25), the directional derivative $\delta(f,x,x^-,t)$ is computed as:

$$\delta(f,x,x^-,t) \;=\; \begin{cases} f(x^-,t) - f(x,t) & \text{if } \ x,\ x^- \in \Omega_i & \text{and } \ x^- \in \mathcal{N}_x, \\ 0 & \text{if } \ x \in \Omega_i,\ x^- \notin \Omega_i & \text{and } \ x^- \in \mathcal{N}_x, \end{cases}$$

where $\mathcal{N}_x$ is is the first order-spatial neighborhood of x. This naive discretization enables to diffuse across object boundaries.

FIGURE 9 TO BE PLACED HERE

The number of iterations of the diffusion process is indirectly controlled by the segmentation map corresponding to a partition of the image into P objects Ω_i: identical level lines/isophotes are invariably selected if the restored image tends to a piecewise-constant image. In our experiments, this can be accomplished with less than $t_{\max} = 50$ iterations in sufficient precision. In contrast to other approaches (Perona and Malik (1990); Black *et al.* (1998); Black and Sapiro (1999); Weickert (1998b)), the diffusion stopping rule is then defined with respect to objects that are being selected and the segmentation map. This relationship is more underlined through a practical segmentation algorithm described in Fig. 9.

5 Experimental results in noise and clutter removal

In this section, we compare results of different diffusion schemes applied to synthetic data, optic and confocal images. All images were produced using the same set of parameters. In our experiments, $\triangle t = 0.025$ and $\alpha_{\bar{\Omega}} = \alpha_{\Omega} = 20$ show better performance in forming piecewise constant regions and simplifying background. These values have been fixed manually from experiments. In addition, α_{Ω} and $\alpha_{\overline{\Omega}}$ are not adjusted to select automatically the stopping time of the diffusion process but to regularize the solution ; in our experiments, the stopping time $t_{\max}$ is heuristically estimated. A subjective examination of examples suggests that the choice of parameters $L, |\Omega_{\min}|$ and λ will depend on the image that is being analyzed.

We point out the simplified Mumford-Shah functional yields to a piecewise-constant approximation/segmentation of the original image. A smoothed and diffused version of the original of the image is not available if we consider only this functional. In our approach, the overall algorithm yields to, on the one hand a piecewise-constant approximation of the image by optimizing (2) along with an identification of the non-connected background, and on the other hand, an object-based restoration of the original image. In fact, further comparison with the complete Mumford-Shah model for segmentation/restoration (Mumford and Shah (1989)) are probably more suitable but the optimization/implementation of this complex model is quite tricky. Finally, we have chosen to reconstruct objects when the preliminary segmentation (i.e. at time $t = 0$) is clearly unsatisfying (see Fig. 10 vs. Fig. 4), which is a more challenging task.

FIGURE 10 TO BE PLACED HERE

FIGURE 11 TO BE PLACED HERE

FIGURE 12 TO BE PLACED HERE

ALGORITHM 2: <<OBJECT SELECTION AND REACTION-DIFFUSION>>

Let $f_0(x)$ be the original image and Let L, λ, $|\Omega_{\min}|$ and $t_{\max}$ be the input parameters set by the user where $t_{\max}$ denotes the maximal number of iterations of the reaction-diffusion process.

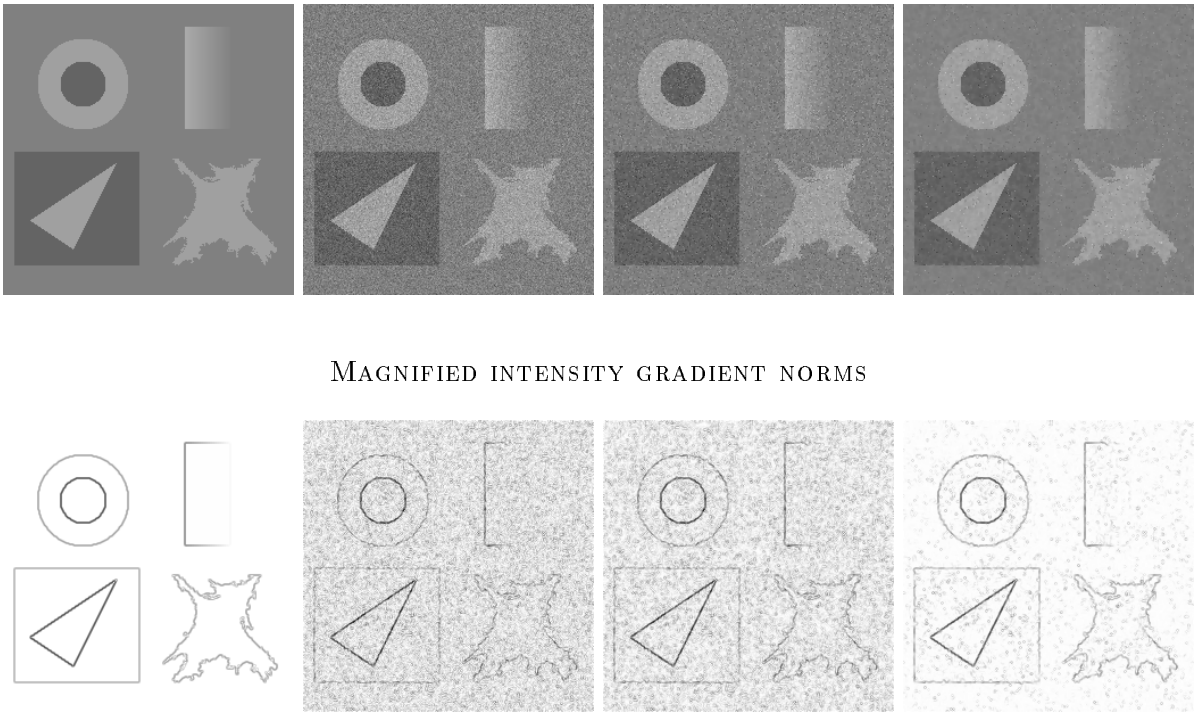
Initialize $t = 0$, $f(x, t = 0) \leftarrow f_0(x)$.

Repeat

- Apply the <<OBJECT SELECTION>> algorithm (Fig. 2)
- Calculate σ_Ω and $\sigma_{\overline{\Omega}}$ using (23)
- Calculate diffusivity $g(\delta(f, x, x^-, t), \sigma_\Omega)$ using (18)
- Solve the reaction-diffusion equations (25)
- Update: $t \leftarrow t + 1$.

until $(t = t_{\max})$ or if $(f(x, t+1) = f(x, t)$, for all $x \in S)$

Figure 9: Object selection and reaction-diffusion algorithm.



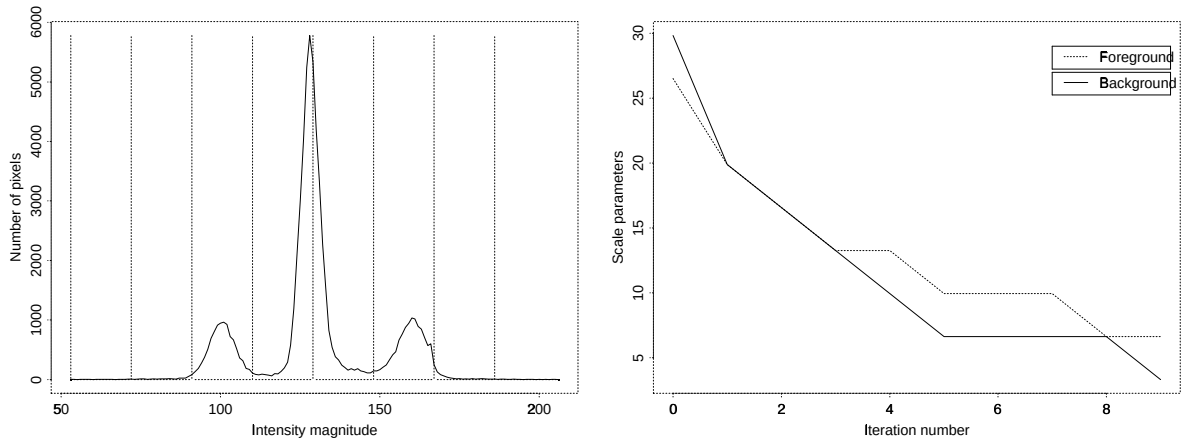
(a) Original image

(b) Noisy image

(c) Robust anisotropic diffusion ($\sigma = 26.52$) (Black *et al.* (1998))

(d) Geometry-driven diffusion-reaction

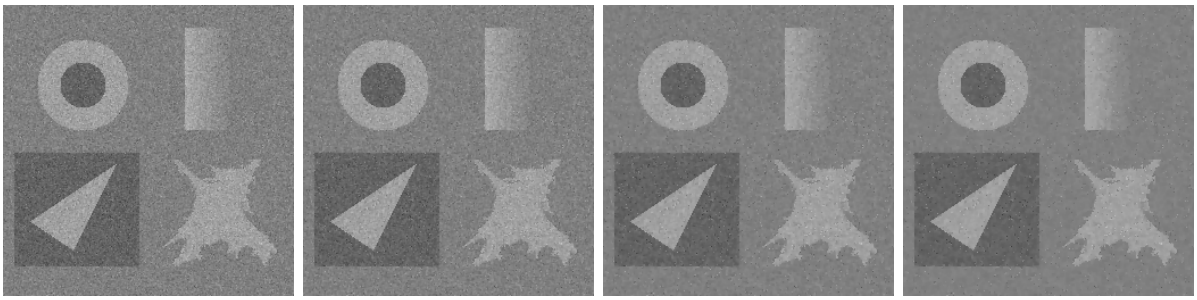
Figure 10: Results of different diffusion schemes after 10 iterations applied to a noisy version (gaussian noise of standard deviation 14.1) of the synthetic image "*gdr*".



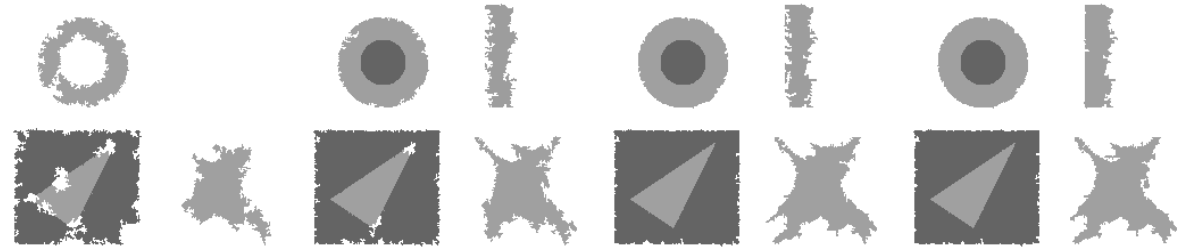
(a) Image histogram

(b) Scale parameters

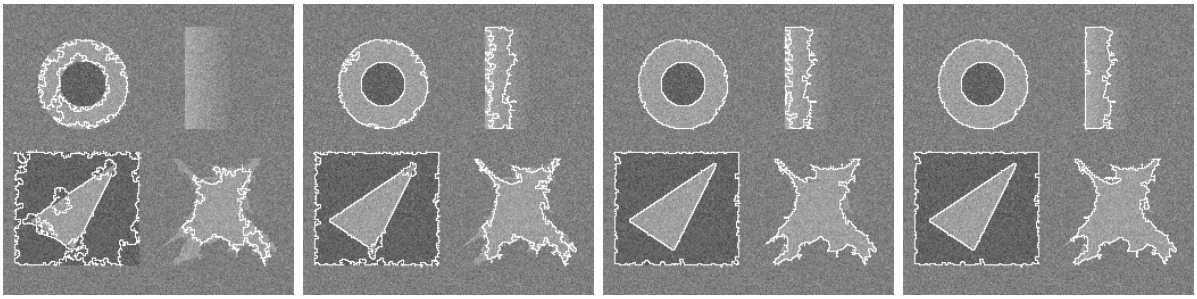
Figure 11: Left: histogram of the restored image after 10 iterations. Right: evolution of scale parameters $\sigma_{\overline{\Omega}}$ and σ_{Ω}.



Piecewise-constant approximation



Boundary sets



(a) iteration 1 (b) iteration 2 (c) iteration 5 (d) iteration 10

Figure 12: Different stages of the object selection and reaction-diffusion process applied to a noisy version (gaussian noise of standard deviation 14.1) of the synthetic image "*gdr*". The columns, from left to right, show the diffused images after 1, 2, 5 and 10 iterations respectively. The first row show the diffused images. The second row shows the piecewise constant approximation of objects where the background is shown using a "*white*" label. The third row shows the boundary sets superimposed on the input noisy image.

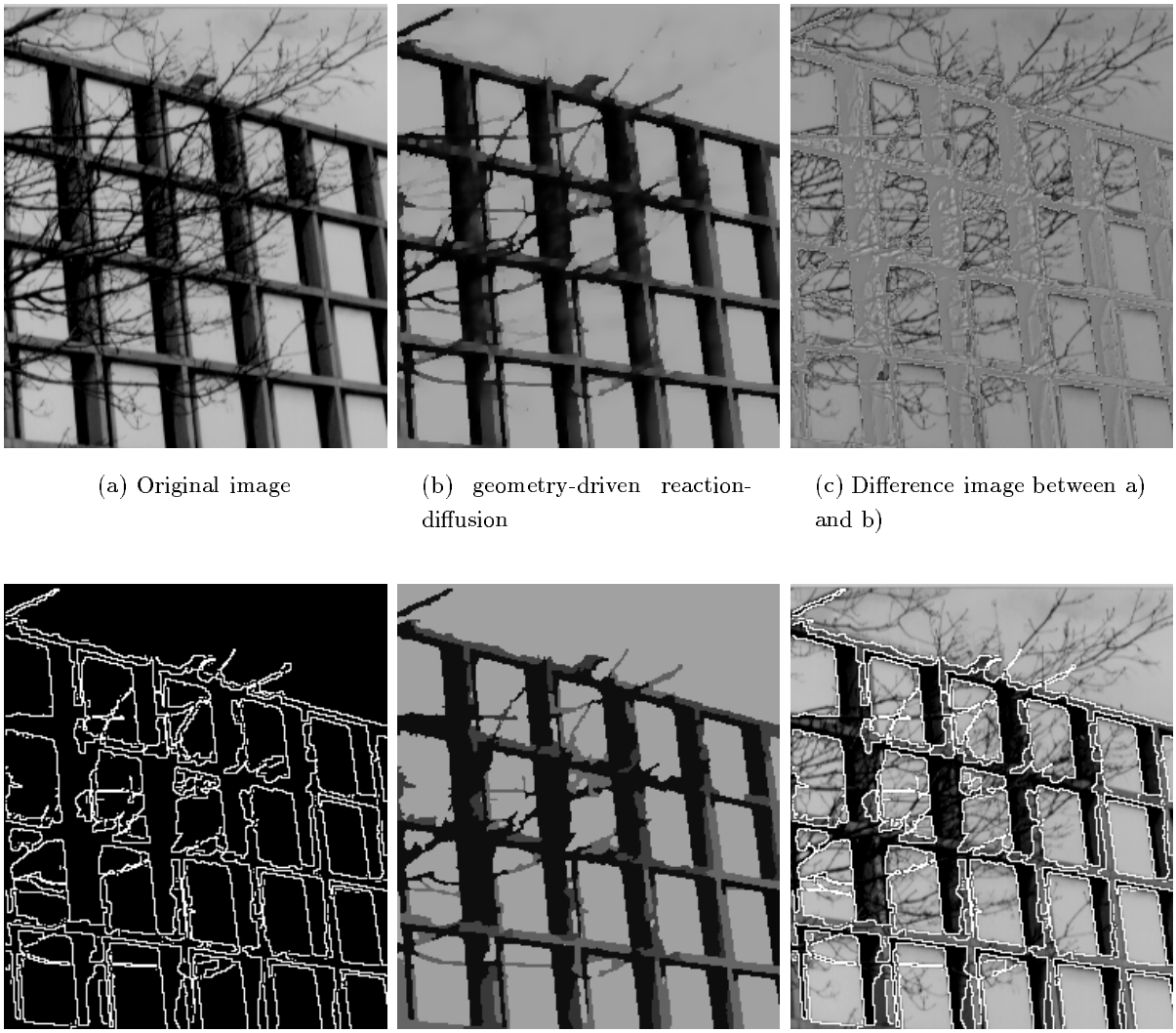
Synthetic data In order to assess the performance of the proposed and other common diffusion schemes, we have used a synthetic 256×256 image where a ground truth is known, coming from the GdR ISIS corrupted with a white gaussian noise. This image shown in Figs. 10 and 12 is composed of six objects against a background and includes two main difficulties: the irregularity of the shape on bottom right-hand side and a grey-level gradation on top right-hand side. Figure 10 shows the results after 10 iterations of two diffusion processes. The first row shows from left to right, the original synthetic "*gdr*" image, the image after adding noise (standard deviation 14.1), the image processed with the robust anisotropic diffusion (Black *et al.* (1998)) (the scale parameter $\sigma = 26.52$ is robustly estimated using (19)), and the image processed with our geometry-driven reaction-diffusion scheme. The second row shows the magnified intensity gradient norm corresponding to each of images on the row above. Both diffusion schemes smooth homogeneous regions but geometry-based approach has a better enhancing effect on the object boundaries after 10 iterations as we can see on the gradient images. Our method locates them with satisfying visual precision and sharpness as shown in Fig. 10. In Fig. 12, the segmentation maps at iteration $t = 1, 2, 5, 10$ are displayed. On this figure, the background and objects are respectively represented by a "white" label and the average values of f computed over each object domain (second row in Fig. 12). The third row shows object boundaries represented by "white" solid lines. After the first iteration, we see that object surfaces are not adequately regularized and regions are not complete (Fig. 12). We retrieve the meaningful objects after 5 iterations using $L = 8$ and $\lambda = 0.001 \times 255^2$ but the linear gradation on top right-hand side is partially merged in the background when $t > 10$. The histogram of the restored image after 10 iterations and the evolution of scale parameters σ_Ω and $\sigma_{\overline{\Omega}}$ as a function of the iteration number are plotted in Fig. 11. The computational time is 20 seconds per iteration on a 296 Mhz workstation for the complete segmentation-filtering procedure.

FIGURE 13 TO BE PLACED HERE

FIGURE 14 TO BE PLACED HERE

FIGURE 15 TO BE PLACED HERE

Clutter removal In many applications, distortions in images are not gaussian. For example, the tree branches in Fig. 13 may be regarded as clutter that cause occlusions of the building. Modeling such clutter is a challenging problem which can be solved by learning Gibbs distributions for each set of objects, i.e. trees and buildings, and maximizing a posterior distribution using a variant of simulated annealing (Langevin equation) (Zhu and Mumford (1997)). This situation can be partially handled by the reaction-diffusion filter described in the previous section. We assume herein that clutter corresponds to non-constant areas in images with respect to homogeneous objects. These inhomogeneous areas are then labeled as background and can



(a) Original image

(b) geometry-driven reaction-diffusion

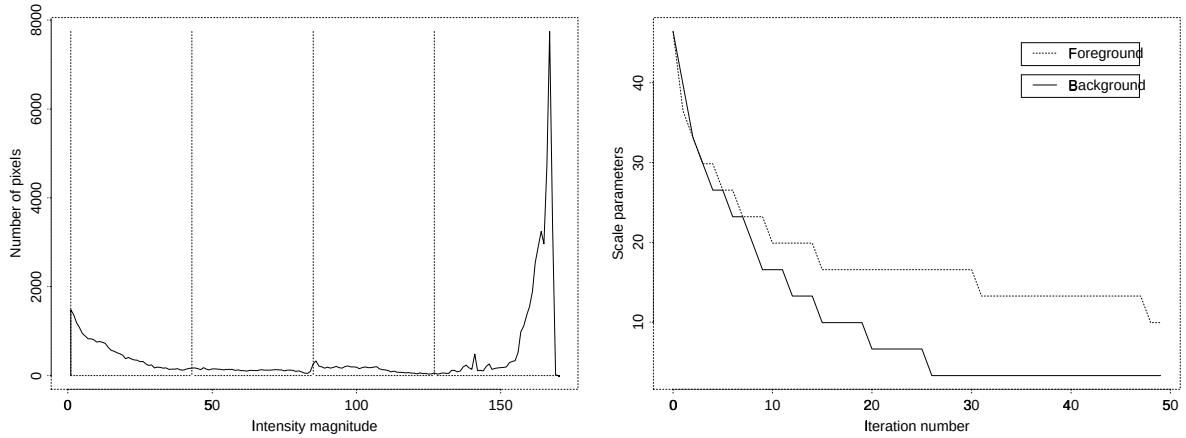
(c) Difference image between a) and b)

(d) Topographic map

(e) Piecewise-constant approximation

(f) Boundary set

Figure 13: Results of the geometry-driven reaction-diffusion scheme after 50 iterations applied to a optical image showing a building and branches.



(a) Image histogram

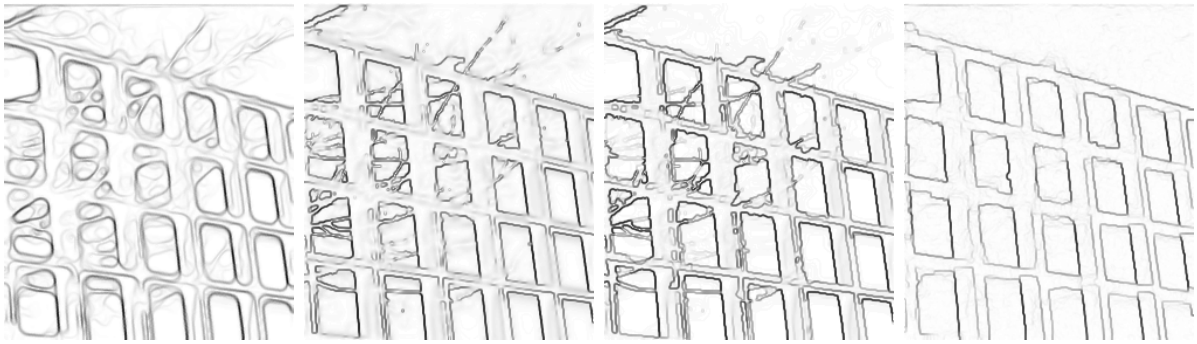
(b) Scale parameters

Figure 14: Left: histogram of the restored image after 50 iterations. Right: evolution of scale parameters $\sigma_{\overline{\Omega}}$ and σ_{Ω}.

ORIGINAL IMAGE

RESULTS OF DIFFERENT DIFFUSION SCHEMES

MAGNITUDE OF THE INTENSITY GRADIENT



(a) Affine morphological scale space

(b) Robust anisotropic diffusion ($\sigma = 59.67$)

(c) Geometry-driven reaction-diffusion

(d) Gibbs reaction-diffusion equations (Zhu and Mumford (1997))

Figure 15: Results of different diffusion schemes applied to a confocal image. First row: original image. Second row, left to right: affine morphological scale space, robust anisotropic diffusion, geometry-driven reaction-diffusion, Gibbs reaction-diffusion equations (Zhu and Mumford (1997)). Third row: magnitude of the intensity gradient corresponding to each image in the row above.

be eliminated using the reaction-diffusion process. However, labeling background seems a challenging problem in image segmentation as we have addressed in Section 2.

Applying our approach to the image (248×288) in Fig. 13, that is setting $L = 4$, $\lambda = 0.025 \times 255^2$ and $|\Omega_{\min}| = 0.001 \times |S|$ pixels, produces visually appealing constant images with straight edges. As we can observe on Figs. 13 and 15, trees branches are partially eliminated from the image and boundaries are enhanced. As a comparison, we run the affine morphological scale-space and the robust anisotropic diffusion (constant scale parameter $\sigma = 59.67$) processes in Figs. 15a and 15b. This corresponding results have been introduced just to demonstrate the performance of this method after 50 iterations when applied on the same data. To go further in the comparisons, the restored image using the Gibbs reaction-diffusion equations (Zhu and Mumford (1997)) is displayed in Fig. 15d. The underlying image formation model is clearly quite different. However, we aims also at discriminating two regions/classes (buildings in the foreground and trees in the background). It appears that the forcing term in the model (22) acts to better eliminate branches without degrading building edges. The image eventually becomes oversmoothed using the two first diffusion schemes with continued iterations whereas our reaction-diffusion process is stopped before 50 iterations in this experiment. The histogram of the restored image after 50 iterations and the evolution of scale parameters σ_Ω and $\sigma_{\overline{\Omega}}$ for increasing number of iterations are plotted in Fig. 14.

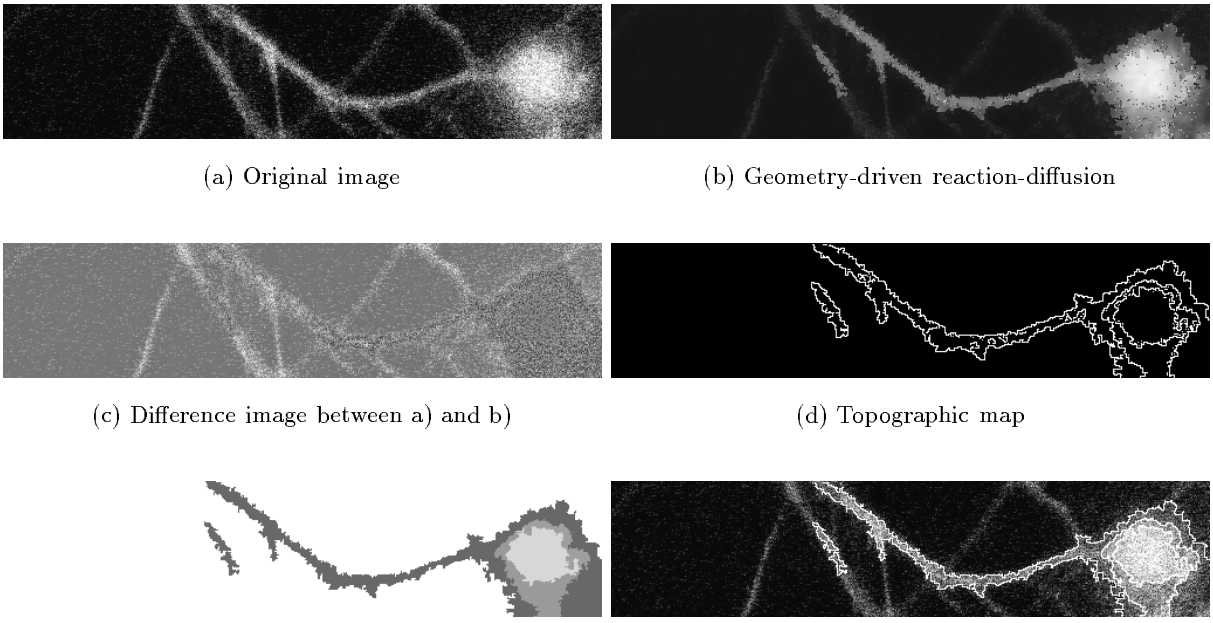
FIGURE 16 TO BE PLACED HERE

FIGURE 17 TO BE PLACED HERE

FIGURE 18 TO BE PLACED HERE

Confocal microscopy imagery Confocal systems offer the chance to image thick biological tissue in 2D+t or 3D dimensions. They operate in the bright-field and fluorescence modes, allowing the formation of high-resolution images with a depth of focus sufficiently small that all the detail which is imaged appears in focus and the out-of-focus information is rejected. Some of the current applications in biological studies are in neuron research. The biological study aims at examining the effect of intracellular calcium concentration (Ca^2) on neurite outgrowth in individual neuronal cells. Indeed, Ca^2 is known to be implicated as an important regulator of neurite extension.

We have tested the proposed algorithm on a 2D confocal microscopy 115×512 image (Fig. 16, courtesy of INSERM 413 IFRMP $n^o 23$ (Rouen, France)) depicting a neurite in cultured cerebellar granule cells. In Figure 16a, high grey-level values correspond to elevated calcium concentration. Accordingly, our goal is twofold: extracting the axon and nucleus in the foreground and segmenting these objects of interest into significant sub-regions corresponding to different levels of calcium concentration. Additionally, undesirable structures (occluded axons) in the



(a) Original image

(b) Geometry-driven reaction-diffusion

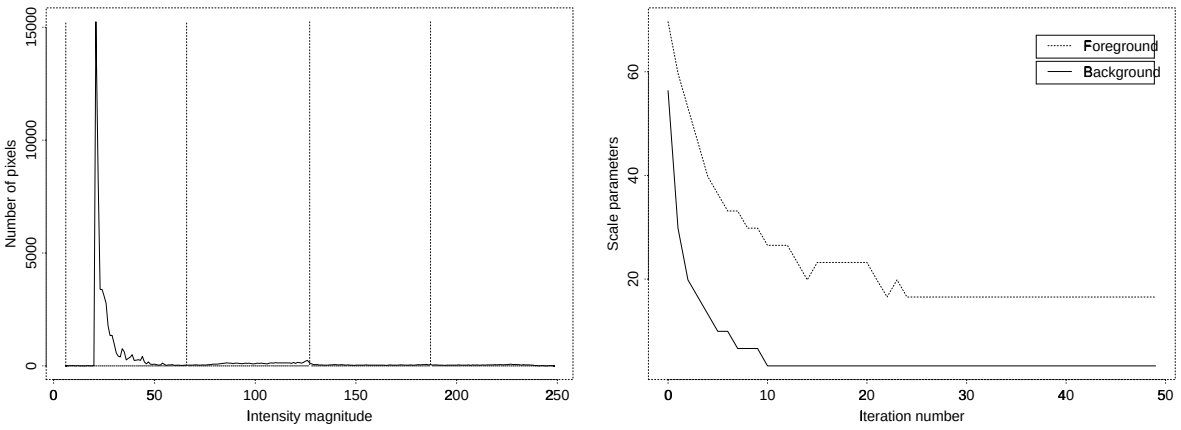
(c) Difference image between a) and b)

(d) Topographic map

(e) Piecewise-constant approximation

(f) Boundary set

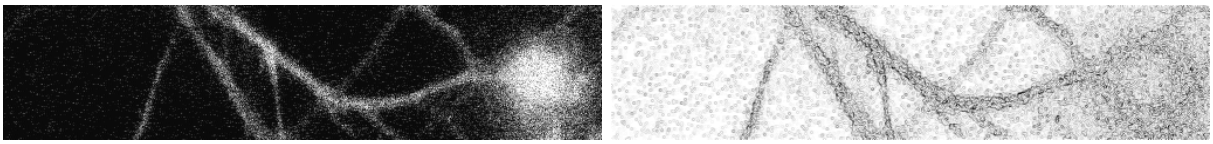
Figure 16: Results of the geometry-driven reaction-diffusion scheme applied to a confocal image.



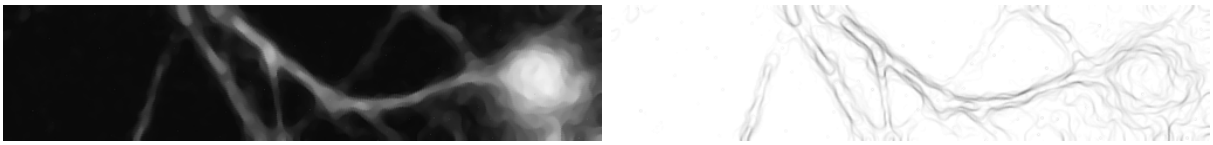
(a) Image histogram

(b) Scale parameters

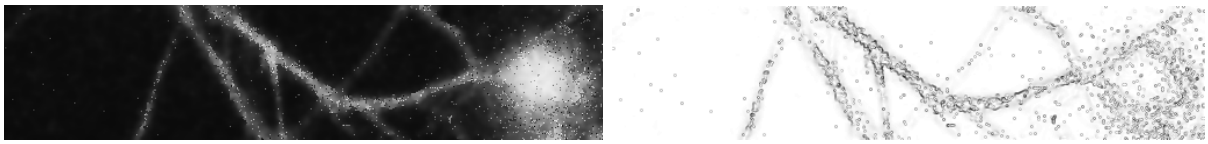
Figure 17: Left: histogram of the restored image after 50 iterations. Right: evolution of scale parameters $\sigma_{\overline{\Omega}}$ and σ_{Ω}.



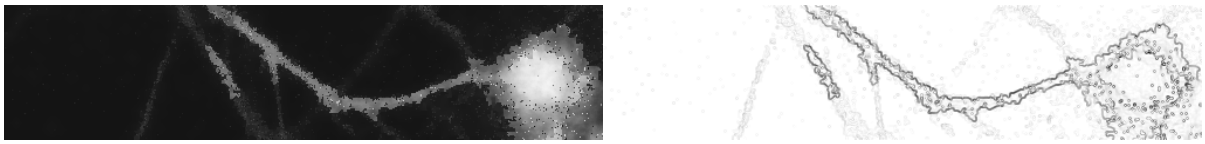
(a) Original image



(b) Affine morphological scale space



(c) Robust anisotropic diffusion ($\sigma = 62.99$)



(d) Geometry-driven reaction-diffusion

Figure 18: Results of the different diffusion schemes applied to a confocal image.Results of the different diffusion schemes applied to a confocal image.

background should be ideally eliminated when the segmentation is performed. The boundaries are quite accurately delineated in Fig. 16 ($L = 4$, $\lambda = 0.05 \times 255^2$, $|\Omega_{\min}| = 0.005 \times |S|$) after 50 iterations. The level lines (of the restored image) corresponding to $L = 4$ are shown on Fig. 16d. On this figure, the background and objects are respectively represented by a "white" label (see second row of Fig. 16). The axon and nucleus have been roughly extracted and segmented into sub-regions associated to mainly three classes of calcium concentration.

This application (Figs. 16 and 18) illustrates the robustness of our method with respect to noise and partial clutter. As we mentioned before, this shows that the proposed method performs reasonably well in presence of non-significant geometric structures. Fig. 17 displays the histogram of the restored image after 10 iterations and the decreasing of scale parameters σ_Ω and $\sigma_{\overline{\Omega}}$. In Fig. 18, if we compare the different diffusion processes and observe that boundaries are better maintained and enhanced using our approach. Noise and irrelevant features are eliminated.

6 Conclusion and perspectives

In this paper, we have presented a level line selection approach for extracting structures in images. We proved that the minimizer of our segmentation energy can be directly characterized. We introduced a geometry-driven reaction diffusion process to preserve sharper boundaries of objects and regularize level lines and object surfaces in the image. A total CPU time of a few seconds for segmenting a 256×256 image on a 296 Mhz workstation makes the method attractive for many time-critical applications. The contribution of this approach has been illustrated on synthetic as well as real-world images.

Several promising directions may be explored for continued research. In particular, some open mathematical questions must be addressed. On the one hand, given the image under consideration, mathematical conditions and constraints to ensure that no pathological minima exists, must be investigated ; we just mention the Morse theory which can be the relevant functional framework as it has been already investigated in image segmentation by Olsen (1996). On the other hand, the underlying global energy functional including segmentation and diffusion is not actually well-defined due to the dependency of $\{\Omega_i, \bar{\Omega}, \overline{f}_{\Omega_i}, \overline{f}_{\bar{\Omega}}, \sigma_\Omega, \sigma_{\bar{\Omega}}\}$ and should be examined in future works. Moreover, the proposed discrete scheme is a naive scheme and probably a more efficient scheme would be suitable. The influence of parameters $\triangle t$, α_Ω and $\alpha_{\bar{\Omega}}$, manually adjusted in our experiments, must be also studied. In other respects, we currently examine other energy functionals which minimizers can be characterized as well, and study an entropy-based quantization technique instead of the uniform quantization used at present to estimate the objects. Finally, an other direction for future work is to extend the proposed approach to operate on multi-spectral 3D images. In this setting, the structure of the algorithm would be largely the same, although there are a number of points which would need to be examined closely.

Acknowledgments

A part of this work has been done using the MegaWave2 image processing environment (Georges KOEPFLER – *Copyright (C)1993-1199 CMLA, ENS Cachan, 94235 Cachan cedex, France* – `http://www.ceremade.dauphine.fr` – All rights reserved –). The authors would like also to thank the GdR ISIS (Information, Signal, Images et viSion), `http://www-isis.enst.fr` of CNRS for the "*gdr*" test image.

References

Acton, S. and Mukherjee, D. 2000. Scale space classification using area morphology. *IEEE Trans. Image Processing*, 9(4): 623–635.

Alvarez, L. and Esclarin, J. 1997. Image quantization using reaction-diffusion equations. *SIAM Journal of Applied Mathematics*, 57(1): 153–1755.

Alvarez, L., Lions, P. and Morel, J. 1992. Image selective smoothing and edge detection by non-linear diffusion - 2. *SIAM J. Numerical Analysis*, 29(3): 845–866.

Alvarez, L., Guichard, F., Lions, P. and Morel, J. 1993. Axioms and fundamental equations in image processing. *Arch. Rational Mech. Aanalysis*, 16(9): 200–257.

Alvarez, L., Gousseau, Y. and Morel, J. 1999. Scales in natural images and a consequence on their bounded variation. In *Int. Conf. on Scale-Space Theories Comp. Vis.*, Kerkyra, Greece, pp. 247–258.

Alvarez, L. and Mazorra, L. 1994. Signal and image restoration using shock filters and anisotropic diffusion. *SIAM Journal of Numerical Analysis*, 31(2): 590–605.

Amadieu, O., Debreuve, E., Barlaud, M. and Aubert, G. 1999. Simultaneous inward and outward curve evolution. In *Int. Conf. on Image Processing*, Kobe, Japan.

Beaulieu, J. and Goldberg, M. 1989. Hierarchy in picture segmentation: a stepwise optimization approach. *IEEE Trans. Patt. Anal. and Mach. Int.*, 11(2): 150–163.

Black, M., Sapiro, G., Marimont, D. and Heeger, D. 1998. Robust anisotropic diffusion. *IEEE Trans. Image Processing*, 7(3): 421–432.

Black, M. and Rangarajan, A. 1996. On the unification of line processes, outlier rejection, and robust statistics with applications in early vision. *Int. J. Computer Vision*, 19(1): 57–92.

Black, M. and Sapiro, G. 1999. Edges as outliers: smoothing using local image statistics. In *Int. Conf. on Scale-Space Theories Comp. Vis.*, Kerkyra, Greece, pp. 259–270.

Blake, A. and Zisserman, A. 1987. *Visual Reconstruction*, MIT Press: Cambridge, Mass.

Caselles, V., Kimmel, R. and Sapiro, G. 1997. Geodesic active contours. *Int J. Computer Vision*, 22(1): 61–79.

Caselles, V., Morel, J., Sapiro, G. and Tannenbaum, A. 1998. Special issue on partial differential equations and geometry-driven diffusion in image processing and analysis. *IEEE Trans. Image Processing*, 7(3): 269–473.

Caselles, V., Coll, B. and Morel, J. 1999. Topographic maps and local contrast changes in natural images. *Int J. Computer Vision*, 33(1): 5–27.

Catte, F., Lions, P., Morel, J. and T., C. 1992. Image selective smoothing and edge detection by non-linear diffusion. *SIAM J. Numerical Analysis*, 29(1): 182–193.

Chan, T. and Vese, L. 1999. Active contour model without edges. In *Int. Conf. on Scale-Space Theories Comp. Vis.*, Kerkyra, Greece, pp. 141–151.

Charbonnier, P., Blanc-Feraud, L., Aubert, G. and Barlaud, M. 1997. Deterministic edge-preserving regularization in computed imaging. *IEEE Trans. Image Processing*, 6(2): 298–311.

Chesnaud, C., Réfrégier, P. and Boulet, V. 2000. Statistical region snake-based segmentation adapted to different physical models. *IEEE Trans. Patt. Anal. and Mach. Int.*, 21(11): 1145–1157.

Chu, C. and Agarwal, J. 1993. The integration of image segmentation maps using region and edge information. *IEEE Trans. Patt. Anal. and Mach. Int.*, 15: 1242–1252.

Cohen, L. 1991. On active contour models and balloons. *CVGIP: Image Understanding*, 53(2): 211–218.

Cohen, L. 1996. Deformable curves and surfaces in image analysis. In *Int. Conf. Curves and Surfaces*, Chamonix, France.

Darrell, T. and Pentland, A. 1995. Cooperative robust estimation using layers of support. *IEEE Trans. Patt. Anal. and Mach. Int.*, 17(5): 474–487.

Deriche, R. 1987. Using canny's criteria to derive a recursively implemented optimal edge detector. *Int J. Computer Vision*, 2(1): 167–187.

Figueiredo, M., Leitao, J. and Jain, A. 2000. Unsupervised contour representation and estimation using b-plines and mininim description length criterion. *IEEE Trans. Image Processing*, 6(9): 1075–1087.

Froment, J. 2000. Perceptible level lines and isoperimetric ratio. In *Int. Conf. on Image Processing*, Vancouver, Canada.

Geiger, D. and Yuille, A. 1991. A common framework for image segmentation. *Int. J. Computer Vision*, 6(3): 1186–1191.

Geman, S. and Geman, D. 1984. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Trans. Patt. Anal. and Mach. Int.*, 6(6): 721–741.

Grenander, U. and Miller, M. 1994. Representations of knowledge in complex systems. *J. Royal Statistical Society, series B*, 56(4): 549–603.

Ishikawa, H. and Geiger, D. 1998. Segmentation by grouping junctions. In *Comp. Vis. and Patt. Rec.*, Santa Barbara, CA.

Istas, J. 1997. *Statistics of processes and signal-image segmentation.* University of Paris VII.

Jermyn, I. and Ishikawa, H. 1999. Globally optimal regions and boundaries. In *Int. Conf. on Comp. Vis.*, Kerkyra, Greece, pp. 904–910.

Kass, M., Witkin, A. and Terzopoulos, D. 1987. Snakes: active contour models. *Int J. Computer Vision*, 12(1): 321–331.

Kichenesamy, S., Kumar, A., Olver, P. and Yezzi, A. 1996. Conformal curvature flows: from transition to active contours. *Archive for Rational Mechanics and Anal.*, 134: 275–301.

Koepfler, G., Lopez, C. and Morel, J. 1994. A multiscale algorithm for image segmentation by variational method. *SIAM J. Numerical Analysis*, 31(1): 282–299.

Kornprobst, P., Deriche, R. and Aubert, G. 1997. Image coupling, restoration and enhancement via pde's. In *Int. Conf. on Image Processing*, Santa Barbara, USA, pp. 458–461.

Leclerc, Y. 1989. Constructing simple stable descriptions for image partitioning. *Int J. Computer Vision*, 3: 73–102.

Malladi, M. and Sethian, J. 1996. Image processing: Flows under min/max curvature and mean curvature. *Graphical Models and Image Processing*, 58(2): 127–141.

Matheron, G. 1975. *Random Sets and Integral Geometry*, John Wiley: New York.

Merriman, B., Bebce, J. and Osher, S. 1994. Motion of multiple junctions: A level set approach. *J. Comput. Physics*, 112(2): 334–363.

Monasse, P. and Guichard, F. 2000. Fast computation of a contrast-invariant image representation. *IEEE Trans. Image Processing*, 9(5): 860–872.

Morel, J. and Solimini, S. 1994. *Variational methods in image segmentation*, Birkhauser.

Mumford, D. 1994. The bayesian rationale for energy functionals. In *Geometry-Driven Diffusion in Computer Vision*, Bart Romeny ed., Kluwer Academic, pp. 141–153.

Mumford, D. and Shah, J. 1989. Optimal approximations by piecewise smooth functions and variational problems. *Com. on Pure and applied Mathematics*, 42(5): 577–685.

Nitzberg, N. and Mumford, D. 1990. The 2.1d sketch. In *Int. Conf. on Comp. Vis.*, Osaka, Japan, pp. 138–144.

Nordström, K. 1990. Biased anisotropic diffusion: a unified regularization and diffusion approach to edge detection. *Image and Vision Computing*, 8(4): 318–327.

Olsen, O. 1996. Multi-scale segmentation of grey-scale images. In *Technical Report DIKU Nr. 96/30, ISSN 0107-8283*, Department of Computer Science, University of Copenhagen.

Osher, S. and Sethian, J. 1988. Fronts propagating with curvature dependent speed: algorithms based on the hamilton-jacobi formulation. *J. Computational Physics*, 79: 12–49.

Paragios, N. and Deriche, R. 2000. Coupled geodesic active regions for image segmentation: a level set approach. In *Euro. Conf. on Comp. Vis.*, Dublin, Ireland.

Pavlidis, T. and Liow, Y. 1990. Integrating region growing and edge detection. *IEEE Trans. Patt. Anal. and Mach. Int.*, 12: 225–233.

Perona, P. and Malik, J. 1990. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Patt. Anal. and Mach. Int.*, 12(7): 629–639.

Rougon, N. and Preteux, F. 1995. Controlled anisotropic diffusion. In *Proc. SPIE Conf. on Nonlinear Image Processing VI*, 2424: 329–340, San Jose, CA.

Rougon, N. and Preteux, F. 1998. Directional adaptive deformable models for segmentation. *J. of Electronic Imaging*, 7(1): 231–256.

Rousseeuw, P. and Leroy, A. 1987. *Robust regression and outlier detection.* Wiley Series in Probability and Mathematical Statistics, Wiley.

Salembier, P. and Serra, J. 1995. Flat zones filtering, connected operators, and filters by reconstruction. *IEEE Trans. Image Processing*, 4(8): 1153–1160.

Samson, C., Blanc-Féraud, L., Aubert, G. and Zerubia, J. 1999. A level set model for image classification. In *Int. Conf. on Scale-Space Theories Comp. Vis.*, Kerkyra, Greece, pp. 306–317.

Schnörr, C. 1998. A study of a convex variational diffusion approach for image segmentation and feature extraction. *J. Math. Imaging and Vision*, 3(8): 271–292.

Sethian, J. 1996. *Level Sets Methods: Evolving Interfaces in Geometry, Fluid Mechanics, Computer Vision, and Material Science*, Cambridges University Press.

Ter Haar Romeny, B. 1994. *Geometry-Driven Diffusion in Computer Vision.* Kluwer Academic Publishers: Dordrecht, The Netherlands.

Vincent, L. and Soille, P. 1991. Watershed in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Trans. Patt. Anal. and Mach. Int.*, 13(6): 583–598.

Wang, J. 1998. Stochastic relaxation on partitions with connected components and its application to image segmentation. *IEEE Trans. Patt. Anal. and Mach. Int.*, 20(6): 619–636.

Weickert, J. 1998a. *Anisotropic Diffusion in Image Processing.* Teubner-Verlag: Stuttgart, Germany.

Weickert, J. 1998b. Fast segmentation methods based on partial differential equations and the watershed transformation. In *Mustererkennung 1998*, P. Levi, R.-J. Ahlers, F. May and M. Schwanz (Eds.): Springer, Berlin, pp. 93–100.

Yezzi, A., Tsai, A. and Willsky, A. 1999. A statistical approach to snakes for bimodal and trimodal imagery. In *Int. Conf. on Comp. Vis.*, Kerkyra, Greece, pp. 898–903,.

You, Y., Xu, W., Tannenbaum, A. and Kaveh, M. 1996. Behavioral analysis of anisotropic diffusion in image processing. *IEEE Trans. Image Processing*, 5: 1539–1553.

Younes, L. (2000. Calibrating parameters of cost functionals. In *Euro. Conf. on Comp. Vis.*, Dublin, Ireland.

Zhu, S. and Mumford, D. 1997. Prior learning and gibbs reaction-diffusion. *IEEE Trans. Patt. Anal. and Mach. Int.*, 19(11): 1236–1250.

Zhu, S. and Yuille, A. 1996. Region competition: unifying snakes, region growing, and bayes/mdl for multiband image segmentation. *IEEE Trans. Patt. Anal. and Mach. Int.*, 18(9): 884–900.

A Proof of Lemma 2

First we consider the usual distance between two closed subsets A and B is the Hausdorff distance, defined by

$$d_\infty(A,B) \;=\; \max\left\{\sup_{x\in A} d(x,B),\, \sup_{x\in B} d(x,A)\right\}, \tag{26}$$

where $d(x,A)$ denotes, as usually, the distance of a point x to the set A

$$d(x,A) \;=\; \inf_{y\in A} |x-y|. \tag{27}$$

For two sets A and B such that $B \subseteq A$, denote $\displaystyle\int_{A\setminus B} f \stackrel{\text{def}}{=} \int_A f - \int_B f$.

Now, let Ω_δ be a variation of a set Ω such that $\Omega \subseteq \Omega_\delta$, i.e. the Hausdorff distance $d_\infty(\Omega_\delta,\Omega) \leq \delta$. To prove Lemma 2, we assume that, for any connected perturbation of Ω such $d_\infty(\Omega_\delta,\Omega) \leq \delta$, two neighboring sets Ω and Ω' do not merge into one single set $\Omega \cup \Omega'$ and, for any connected perturbation of Ω such $d_\infty(\Omega_\delta,\Omega) \leq \delta$, Ω does not split into two new sets. This corresponds to prohibited topological changes. Notice the Morse theory can be the relevant functional framework to ensure that no topological changes does occur or not Olsen (1996). Without loss of generality, we prove Lemma 2 for one object Ω and a background $\overline{\Omega}$, that is the closure of the complementary set of Ω. Then, we have

$$\int_{\Omega_\delta\setminus\Omega} 1 \stackrel{\text{def}}{=} |\Omega_\delta| - |\Omega| \quad \text{and} \quad \left(\int_{\Omega_\delta} f\right)^2 - \left(\int_{\Omega} f\right)^2 = 2\int_\Omega f \int_{\Omega_\delta\setminus\Omega} f + \left(\int_{\Omega_\delta\setminus\Omega} f\right)^2. \tag{28}$$

The difference between the involved energies is equal to $\Delta E_\lambda(f,\Omega,\overline{\Omega}) = E_\lambda(f,\Omega_\delta,\overline{\Omega}_\delta) - E_\lambda(f,\Omega,\overline{\Omega})$ where

$$\left\{\begin{array}{lcl} E_\lambda(f,\Omega,\overline{\Omega}) &=& \displaystyle\int_\Omega (f(x)-\overline{f}_\Omega)^2\,dx \;+\; \int_{\overline{\Omega}} (f(x)-\overline{f}_{\overline{\Omega}})^2\,dx \;+\; \lambda\,|\Omega| \\ \\ E_\lambda(f,\Omega_\delta,\overline{\Omega}_\delta) &=& \displaystyle\int_{\Omega_\delta} (f(x)-\overline{f}_{\Omega_\delta})^2\,dx \;+\; \int_{\overline{\Omega}_\delta} (f(x)-\overline{f}_{\overline{\Omega}_\delta})^2\,dx \;+\; \lambda\,|\Omega_\delta|. \end{array}\right. \tag{29}$$

Therefore, $\Delta E_\lambda(f,\Omega,\overline{\Omega}) = T_1+T_2+T_3+T_4+T_5$ with

$$\begin{array}{lcll} T_1 &=& \displaystyle\int_{\Omega_\delta} f^2 - \int_\Omega f^2, & T_2 = -\frac{1}{|\Omega_\delta|}\left(\int_{\Omega_\delta} f\right)^2 + \frac{1}{|\Omega|}\left(\int_\Omega f\right)^2, \\ T_3 &=& \displaystyle\int_{S\setminus\Omega_\delta} f^2 - \int_{S\setminus\Omega} f^2\,, & T_4 = -\frac{1}{|S|-|\Omega_\delta|}\left(\int_{S\setminus\Omega_\delta} f\right)^2 + \frac{1}{|S|-|\Omega|}\left(\int_{S\setminus\Omega} f\right)^2 \\ T_5 &=& \displaystyle\int_{\Omega_\delta\setminus\Omega} \lambda\,. & \end{array} \tag{30}$$

Denote $\Delta|\Omega| = |\Omega_\delta| - |\Omega|$. Using (28), and passing to the limit $\Delta|\Omega| \to 0$, i.e. $|\Omega_\delta| \simeq |\Omega|$, we

obtain (higher order terms are neglected)

$$T_1 = -T_3 = \int_{\Omega_\delta \setminus \Omega} f^2\,, \quad T_2 = -\frac{2}{|\Omega|}\int_{\Omega_\delta\setminus\Omega} f \int_\Omega f - \frac{1}{|\Omega|}\left(\int_{\Omega_\delta\setminus\Omega} f\right)^2 + \frac{1}{|\Omega|^2}\int_{\Omega_\delta\setminus\Omega} 1 \left(\int_\Omega f\right)^2,$$

$$T_4 = \frac{2}{|S|-|\Omega|}\int_{\Omega_\delta\setminus\Omega} f \int_{S\setminus\Omega} f - \frac{1}{|S|-|\Omega|}\left(\int_{\Omega_\delta\setminus\Omega} f\right)^2 - \frac{1}{(|S|-|\Omega|)^2}\int_{\Omega_\delta\setminus\Omega} 1 \left(\int_{S\setminus\Omega} f\right)^2, \quad (31)$$

and $T_5 = \lambda \int_{\Omega_\delta\setminus\Omega} 1$. Define the following image moments

$$m_0 = \int_\Omega 1\,, \quad m_1 = \int_\Omega f\,, \quad M_0 = \int_S 1\,, \quad M_1 = \int_S f\,. \quad (32)$$

Using the *mean value theorem for double integral*, which states that if f is continuous and a connected subset E is bounded by a simple curve, then for some point x_0 in E we have $\int_E f(x)dE = f(x_0)\cdot|E|$ where $|E|$ denotes the area of E, it follows that

$$\Delta E_\lambda(f,\Omega,\overline{\Omega}) = \overbrace{\left[\frac{m_1^2}{m_0^2} - \frac{(M_1-m_1)^2}{(M_0-m_0)^2} + \lambda\right]}^{a_0}\int_{\Omega_\delta\setminus\Omega} 1 + \overbrace{\left[-\frac{2m_1}{m_0} + \frac{2(M_1-m_1)}{M_0-m_0}\right]}^{a_1} f(x_0)\int_{\Omega_\delta\setminus\Omega} 1$$
$$- \left[\frac{1}{m_0} + \frac{1}{M_0-m_0}\right] f(x_0)^2 \left(\int_{\Omega_\delta\setminus\Omega} 1\right)^2. \quad (33)$$

Let x_b be a fixed point of the boundary $\partial\Omega$. Choose Ω_δ such that $\partial\Omega_\delta = \partial\Omega$ except on a small neighborhood of x_b. The energy having a minimum for Ω, $f(x_b)$ needs to be solution of the following equation

$$\frac{\Delta E_\lambda(f,\Omega,\overline{\Omega})}{\Delta|\Omega|} = a_0 + a_1 f(x_b) + O(\Delta|\Omega|) = 0. \quad (34)$$

By passing to the limit $\Delta|\Omega| \to 0$, we obtain

$$a_0 + a_1 f(x_b) = 0. \quad (35)$$

Equation (35) has one single solution. The coefficients a_0 and a_1 do depend on neither x_b nor $f(x_b)$, and $a_0 \neq 0$. The function f is continuous $S \subset \mathbb{R}^2$ and $\partial\Omega$ is a connected curve. Therefore $f(x_b)$ is constant when x_b covers $\partial\Omega$. In other words, $\partial\Omega$ is a level line of f. This completes the proof. □