Variational snake theory

Agnès Desolneux
Lionel Moisan
Jean-Michel Morel

ABSTRACT In this chapter, we review briefly the theory of edge detection and its non local version, the "snake", or "active contour" theory. The snakes are curves minimizing locally a contrast and smoothness energy in the image. We show that simpler and simpler models have emerged and we support the idea that a very recent class of models proposed by Kimmel and Bruckstein is actually optimal. The snake energy simply is an average contrast across the curve, the contrast being measured as a function g of the image gradient through the snake, u_n. We discuss, however, this last model from two points of views : the shape of the contrast function g and the experimental blunders due to the varying contrast along the sought for boundary. This leads us to propose a very particular form of the contrast function in the mentioned snake model, as close as possible to a threshold function of the gradient. Eventually, we show by arguments and experiments that the resulting snakes simply coincide with the well contrasted level lines of the image. This is shown in two ways : first, we prove that all meaningful level lines of the image *hardly move* by the snake evolution equation and second that if we evolve snakes by their energy, they tend to follow well contrasted level lines. For a sake of completeness, we give all formal computations needed for deriving the main models, their evolution equation and steady state equation. Also, the very simple direct curve energy minimization which we use for experiments is described. It does not actually use the "level set method" of Osher and Sethian. Indeed, the big claimed advantage of these methods is to deal with topology changes of the snake during the minimization process. If, as we sustain, the snakes can be replaced in most practical cases by level lines, the topological changes are simply handled by using their nested inclusion tree.

1 Introduction

It seems to be as useful as ever to discuss the way salient boundaries, or edges, can be computed in an image. In the past 30 years, many methods have been proposed and none has imposed itself as a standard. All have in common a view of "edginess" according to which an edge is a curve in the image across which the image is contrasted. Hildreth and Marr [MH80] proposed to define the boundaries in a grey level image $u(x, y)$ as lines of

points where the Laplacian

$$\Delta u(x,y) = \frac{\partial^2 u}{\partial x^2}(x,y) + \frac{\partial^2 u}{\partial y^2}(x,y)$$

crosses zero. This was based on the remark, true for a one-dimensional function $u(x)$, that points with the highest gradient satisfy $u''(x) = 0$. Haralick [Har84] improved this a lot by simply proposing as "edge points" the points where the magnitude of the gradient $|Du|$ attains a maximal value along gradient lines. An easy computation shows that such points (x,y) satisfy $D^2u(Du,Du)(x,y) = 0$, where we take as a notation

$$D^2u(x,y) = \begin{pmatrix} \frac{\partial^2 u}{\partial x^2} & \frac{\partial^2 u}{\partial x \partial y} \\ \frac{\partial^2 u}{\partial y \partial x} & \frac{\partial^2 u}{\partial x^2} \end{pmatrix}, \quad Du = \begin{pmatrix} \frac{\partial u}{\partial x} \\ \frac{\partial u}{\partial y} \end{pmatrix}.$$

(If $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ and $\xi = \begin{pmatrix} v \\ w \end{pmatrix}$, we set $A(\xi,\xi) = av^2 + 2bvw + cw^2$.)
Thus, in the following, we shall talk about Hildreth-Marr's and Haralick's edge points respectively. The Haralick's edge points computation was proved by Canny [Can86] to be optimal for "step edges" under a Gaussian noise assumption. Now, image analysis aims at global structures. Edge points are from that point of view a poor basis for the further building up of geometric structures in image perception. This is why edge detection methods have evolved towards "boundary detection" methods, i.e. methods which directly deliver curves in the image along which the gradient is, in some sense, highest.

There is agreement about the following criteria : if the image contrast along a curve is high, so much the better. If in addition such a contrasted curve is smooth, its "edginess" is increased. The requirement that the curve be smooth must be tempered, however, since simple objects can be ragged or have corners.

Let us give some notations. A smooth curve in the image will be denoted by $\gamma(s)$, a one to one map from some real interval into the image plane. Unless otherwise specified, s is an Euclidean parameterization, which means by definition that

$$|\gamma'(s)| = 1.$$

We shall denote by $L(\gamma)$ the length of γ. When $v = (x,y)$ is a vector, we set $v^\perp = (-y,x)$, a vector orthogonal to v. The unit tangent vector to the curve is $\gamma'(s)$ and we set

$$\vec{n}(s) = \gamma'(s)^\perp,$$

a vector normal to the curve γ. We finally consider

$$\gamma''(s) = Curv(\gamma(s)), \tag{1.1}$$

which is a vector normal to the curve whose magnitude is proportional to its curvature. If $Curv(\gamma(s))$ is bounded, the curve behaves locally as a circle with radius $1/|Curv(\gamma(s))|$ and is therefore smooth. Thus, if some privilege has to be given to smooth curves, one is led to define an "edge" as a curve $\gamma(s)$ such that

$$\int_{[0,L(\gamma)]} g(|Du(\gamma(s))|)ds$$

is minimal (g being any decreasing function), and such that

$$\int_{[0,L(\gamma)]} (1 + |Curv(\gamma(s))|)ds$$

is minimal. Both requirements can be combined in a single energy functional : the Kass-Witkin-Terzopoulos "snake" original model [MKT87], where the minimum of

$$\int_0^{L(\gamma)} g(|Du(\gamma(s)|))ds + C \int_0^{L(\gamma)} (a + |Curv(\gamma(s))|)ds \tag{1.2}$$

is sought for. According to this formulation, a snake must be as smooth as possible and have a contrast as strong as possible.

This model has been generally abandoned for the derived "geodesic snakes" [CKS97]. This model proposed a geometric form of the snakes functional, where the minimum of

$$\int_0^{L(\gamma)} g(|Du(\gamma(s)|))ds \tag{1.3}$$

is sought for. Since $g > 0$, this last functional looks for a compromise between length and contrast and yields a well contrasted curve with bounded curvature. Indeed, if $g(0) > 0$, which is usual, the first term also gives a control on the length of the snake, and therefore actually forces the curvature to be bounded. Notice also that the minimization process therefore tends to decrease the length and forces the snake to shrink, which is not exactly what is wished !

This may explain why Fua and Leclerc [FL90] proposed to minimize, for edge detection, the average functional

$$\frac{1}{L(\gamma)} \int_0^{L(\gamma)} g(|Du(\gamma(s)|))ds. \tag{1.4}$$

Here again, g is decreasing. Minimizing this functional amounts to finding a curve in the image with, so to say, maximal average contrast. One of the main advances in the formulation adopted in [FL90] and [CKS97] is the

reduction of the number of model parameters or functions to a single one : the contrast function g. The Fua-Leclerc model is in our opinion better since it focuses on contrast only and is therefore apparently no more a hybrid combination of contrast and smoothness requirements.

Now, all above mentioned models were in back with respect to the first edge detectors, defined in the beginning of the seventies. Indeed, the Montanari [Mon71] and Martelli [Mar72] original boundary detection models were more accurate in one aspect : instead of $|Du(\gamma(s))|$, they used as contrast indicator a discrete version of

$$u_n(s) = \frac{\partial u}{\partial n}(s) = Du(\gamma(s)).\vec{n}(s), \tag{1.5}$$

that is, the *contrast of the image across the curve*. At a point $\gamma(s)$, we can see that $u_n(s)$ is larger if the magnitude of the gradient, $|Du(\gamma(s))|$, is larger, but also if this the gradient is as much as possible normal to the curve. Clearly, the above Kass, Witkin, Terzopoulos, the Fua-Leclerc and the Caselles-Kimmel-Sapiro contrast measures are worse at that point : they only take into account the the magnitude of the gradient, independently of its angle with the curve normal.

As a general remark on all variational snakes, we must also notice that if g is nonnegative, which is usual, the best snake in the energetical sense is reduced to a single point at which the maximum magnitude of the gradient of the image is attained. Thus, in all snake models, local minima of the snake energy should be sought for, the global ones being irrelevant. Such local minima usually depend upon :

- the initial position of the snake,
- the variance of the usual preliminary smoothing of the gradient,
- the form of the contrast function g.

Recently, Kimmel and Bruckstein [KB01] [KB02] made several important advances on the formulation of edge detectors and the snakes method, which we shall discuss, and push a bit further in this chapter. We can summarize the Kimmel-Bruckstein results as follows.

- Maximizers of the contrast along γ, $\int_0^{L(\gamma)} u_n(s)ds$ satisfy $\Delta u(\gamma(s)) = 0$, provided u_n does not change sign along the curve. This yields a variational interpretation of Hildreth-Marr edges.
- Active contours can more generally be performed by maximizing a nonlinear function of contrast, $E(\gamma) = \int_0^{L(\gamma)} g(u_n(s))ds$, where g is even and increasing, a good example being $g(t) = |t|$. This is basically the energy (1.3) but where the isotropic contrast indicator $|Du(\gamma(s))|$ is replaced by the better term $u_n(s) = Du(\gamma(s)).\vec{n}(s)$

used in Montanari-Martelli. The case $g(t) = t^2$ was actually considered earlier, in the founding Mumford-Shah paper [MS89]. More precisely, Mumford and Shah considered the minimization of

$$E_\infty = \int_0^{L(\gamma)} \nu - u_n(s)^2 \, ds,$$

but they discovered that this functional has no minimizers because of the folding problem, which we shall address further on.

Also, Kimmel and Bruckstein consider maximizing the *average contrast*, namely

$$E(\gamma) = \frac{1}{L(\gamma)} \int_0^{L(\gamma)} g(u_n(s)) ds, \tag{1.6}$$

where g is some increasing function. We then get an improved version of the Fua-Leclerc functional.

All this looks good and fine : the energy functional is simpler, it does not enforce a priori smoothness constraints and it measures the real contrast.

- The evolution equation towards an optimum boundary can be written in much the same way as in the geodesic snake method.

- In [KB02], it is also shown that the Haralick operator can receive a global variational snake interpretation. In order to give this interpretation, let us define a (non local) orientation at each point $\mathbf{x}_0$ of the image plane in the following way : we consider the level line $\{\mathbf{x}, u(\mathbf{x}) = u(\mathbf{x}_0)\}$ passing by $\mathbf{x}_0$ and we set $O(\mathbf{x}) = +1$ if the level line surrounds $\mathbf{x} + \varepsilon Du(\mathbf{x})$ for $\varepsilon > 0$ small enough, $O(\mathbf{x}) = -1$ otherwise (all level lines are supposed to be closed and compact ; this can be ensured (e.g.) by taking $u = -\infty$ outside the image domain). Call $\Omega(\gamma)$ the domain surrounded by a Jordan curve γ and consider the functional

$$E(\gamma) = \int_0^{L(\gamma)} O(\gamma(s)) u_n(s) ds - \int\int_{\Omega(\gamma)} O(\mathbf{x}) |Du(\mathbf{x})| curv(u)(\mathbf{x}) d\mathbf{x}.$$

Maximizing the first term means finding a curve across which the image u has maximal contrast. As Kimmel and Bruckstein show, the second term is a topological complexity measure of the image inside the curve : when γ is a level line, this term is exactly 2π times the number of hills and dips of the image, weighted by their heights. Jordan curves along which $D^2u(Du, Du) = 0$ are critical curves for the preceding functional.

Our purpose in this chapter is to complete the above discussion. We shall start by explaining in detail some of the above mentioned results. In particular, we shall discuss the main point left out, namely the shape of g in (1.6) and prove that not all g's are convenient. More precisely, we shall prove that *the form of g is roughly fixed if one wants to avoid errors in the snake method due to disparities of contrast along the snake.* In continuation, we shall prove that in a normal quality image, *all snakes without reverse contrast can be defined as a subset of the level lines of the image.* To be more precise, we shall show that a *simple selection* of the level lines of the image, by choosing the best contrasted level lines in the level line tree of the image, gives back all perceptually expected contours. This evidence will be given by showing that when we try to maximize average contrast of the locally best contrasted level lines of the image by moving them as initial curves to the Kimmel-Bruckstein snake method, the curves do not move in general.

2 Curve flows maximizing the image contrast

In this section, we compute the Euler-Lagrange equations for the main snake energies introduced in the introduction. This derivation is made in [KB01], [KB02] ; we tried to make it a bit more explicit.

2.1 The first variation of the image contrast functional

Let $\gamma(s)$ be a closed Jordan curve parameterized by arc length, and let g be a real even function, increasing on $\mathbb{R}^+$, and C^2 except possibly at 0 (e.g. $g(t) = |t|^\alpha$). The Kimmel-Bruckstein energy is defined by

$$E(\gamma) = \int_0^{L(\gamma)} g(u_n(s))\, ds. \tag{1.7}$$

We now compute its first variation.

Proposition 1 *Set $h(t) = g(t) - tg'(t)$. The Gateaux derivative of $E(\gamma)$ with respect to perturbations ε is*

$$\nabla E_\varepsilon(\gamma) = \int_0^{L(\gamma)} \Big[\, (g'(u_n))' Du^\perp(\gamma) + g'(u_n)\Delta u(\gamma)\vec{n} \\ -(h(u_n))'\gamma' - h(u_n)Curv(\gamma) \, \Big].\varepsilon \, ds \tag{1.8}$$

We shall need the simple formal computation of the next lemma.

Lemma 1 *Let A be 2×2 symmetric matrix. Then for all pairs of two dimensional vectors v and w, $A(v, w^\perp) - A(v^\perp, w) = Trace(A)v.w^\perp$.*

Proof of proposition 1

We consider a perturbation $\lambda\varepsilon$ of γ, with $\lambda \in I\!R$ and $\varepsilon : [0, L(\gamma)] \to I\!R^2$. We want to compute

$$\nabla_\varepsilon E(\gamma) = \frac{d}{d\lambda} E(\gamma + \lambda\varepsilon)_{/\lambda=0}.$$

Since $|\gamma'(s) + \lambda\varepsilon'(s)|$ has no reason to be equal to 1, s is not an arc length parameterization of the curve $s \mapsto \gamma(s) + \lambda\varepsilon(s)$. Thus, we differentiate the general (non-arc-length) form of E from (1.7) and write

$$E(\gamma + \lambda\varepsilon) = \int_0^{L(\gamma)} g\left(Du(\gamma + \lambda\varepsilon).\frac{\gamma'^{\perp} + \lambda\varepsilon'^{\perp}}{|\gamma' + \lambda\varepsilon'|}\right) |\gamma' + \lambda\varepsilon'|\, ds.$$

Now, since $|\gamma'| = 1$ and $\vec{n} = \gamma'^{\perp}$, we have

$$\frac{d}{d\lambda}|\gamma' + \lambda\varepsilon'|_{/\lambda=0} = \varepsilon'.\gamma' \quad \text{and} \quad \frac{d}{d\lambda}\left(\frac{\gamma'^{\perp} + \lambda\varepsilon'^{\perp}}{|\gamma' + \lambda\varepsilon'|}\right)_{/\lambda=0} = \varepsilon'^{\perp} - (\gamma'.\varepsilon')\vec{n}.$$

Hence,

$$\begin{aligned}
&\nabla_\varepsilon\ E(\gamma) \\
&= \int_0^L g'(u_n) Du(\gamma).\left(\varepsilon'^{\perp} - (\gamma'.\varepsilon')\vec{n}\right) + g'(u_n) D^2u(\gamma)(\vec{n}, \varepsilon) + g(u_n)\gamma'.\varepsilon' \\
&= \int_0^L g'(u_n) Du(\gamma).\varepsilon'^{\perp} + h(u_n)\gamma'.\varepsilon' + g'(u_n) D^2u(\gamma)(\vec{n}, \varepsilon),
\end{aligned}$$

since $Du(\gamma).(\gamma'.\varepsilon')\vec{n} = u_n\gamma'.\varepsilon'$ and $h(t) = -u_n g'(u_n) + g(u_n)$. We notice that all terms are assumed to be C^2 and are integrated on a closed contour. Thus, we can perform all integrations by parts we wish without boundary terms. This yields

$$\begin{aligned}
\nabla_\varepsilon E(\gamma) \quad = \quad & \int_0^L -(g'(u_n))' Du(\gamma).\varepsilon^{\perp} - g'(u_n) D^2u(\gamma)(\gamma', \varepsilon^{\perp}) \\
& -(h(u_n))'\gamma'.\varepsilon - h(u_n) Curv(\gamma).\varepsilon + g'(u_n) D^2u(\gamma)(\vec{n}, \varepsilon).
\end{aligned}$$

We conclude by using the fact that $Du(\gamma).\varepsilon^{\perp} = -Du(\gamma)^{\perp}.\varepsilon$ and that

$$-D^2u(\gamma)(\gamma', \varepsilon^{\perp}) + D^2u(\gamma)(\vec{n}, \varepsilon) = \Delta u(\gamma)\, \vec{n}.\varepsilon$$

thanks to Lemma 1. □

Proposition 1 permits to deduce the *maximization gradient flow* for $E(\gamma)$, namely $\frac{\partial\gamma}{\partial t} = \nabla E(\gamma)$. As usual with curve evolution, we can restrict ourselves to depict the evolution by its velocity in the direction of the normal $\gamma'^{\perp} = \vec{n}$ to the curve γ.

Corrolary 1 *The curve flow maximizing the curve contrast (1.7) is*

$$\frac{\partial \gamma}{\partial t} = (g'(u_n))'(Du^{\perp}(\gamma).\vec{n})\vec{n} + g'(u_n)\Delta u(\gamma)\vec{n} - h(u_n)Curv(\gamma). \qquad (1.9)$$

Notice that the term $-(h(u_n))'\gamma'$ disappears in the normal flow formulation because this tangential motion does not influence the geometric motion of γ. It can instead be interpreted as a motion of the parameterization of the curve $\gamma(s)$ itself. For the same reason, we only kept the component of $Du^{\perp}(\gamma)$ normal to the curve, namely $(Du^{\perp}(\gamma).\vec{n})$ instead of $Du^{\perp}(\gamma)$. The last two terms of the above expression have been left unaltered, since they both are terms orthogonal to the tangent to the curve, γ'.

A special case of (1.9) is obtained when $g(t) = |t|$. In this case, we have $h(t) = 0$ and the normal flow equation boils down to

$$\frac{\partial \gamma}{\partial t} = ((g'(u_n))'(Du^{\perp}(\gamma).\vec{n})\vec{n} + g'(u_n)\Delta u(\gamma)\vec{n}. \qquad (1.10)$$

If u_n has constant sign on γ, i.e. if u does not reverse contrast along γ, we have $g'(u_n) = \text{sign}(u_n) = \pm 1$ and therefore $(g'(u_n))' = 0$. Thus, the equation becomes

$$\frac{\partial \gamma}{\partial t} = \text{sign}(u_n)\Delta u(\gamma)\vec{n}. \qquad (1.11)$$

Thus, one obtains a *variational interpretation* of the Hildreth-Marr edge detector. Those edges can be obtained as steady states for a *boundary contrast maximizing flow.*

Proposition 2 *(Kimmel-Bruckstein) The curves of an image which are local extrema of the contrast* $E(\gamma) = \int_\gamma |u_n|$ *satisfy* $\Delta(\gamma(s)) = 0$ *and are therefore Hildreth-Marr* edges*, provided* u_n *does not change sign along the curve.*

2.2 A parameterless edge equation

Following Fua and Leclerc [FL90] and Kimmel and Bruckstein [KB01], we now consider a second, probably better, definition of a boundary as a closed curve along which the *average* contrast is maximal. Thus, we consider the energy

$$F(\gamma) = \frac{1}{L(\gamma)} \int_0^{L(\gamma)} g(u_n)\, ds = \frac{E(\gamma)}{L(\gamma)}, \qquad (1.12)$$

where $L(\gamma)$ is the length of γ. Again, we can compute the first variation of F with respect to C^1 perturbations of γ, which yields

$$\nabla_\varepsilon F(\gamma) = \frac{1}{L(\gamma)} \left(\nabla_\varepsilon E(\gamma) - F(\gamma)\nabla_\varepsilon L(\gamma) \right).$$

Now,

$$\nabla_\varepsilon L(\gamma) = \frac{d}{d\lambda}\left(\int_0^{L(\gamma)} |\gamma' + \lambda\varepsilon'|\, ds\right)_{/\lambda=0} = \int_0^{L(\gamma)} \gamma'.\varepsilon' = -\int_0^{L(\gamma)} Curv(\gamma).\varepsilon.$$

Thus, we can again write an evolution equation for the average contrast maximizing flow,

$$\frac{\partial\gamma}{\partial t} = ((g'(u_n))'(Du^\perp(\gamma).\vec{n}) + g'(u_n)\Delta u(\gamma))\vec{n} + (F(\gamma) - h(u_n))Curv(\gamma). \tag{1.13}$$

If we again consider the important case where the sign of u_n is constant along γ and if we choose $g(t) = |t|$, we obtain, since $g'(t) = \text{sign}(t)$,

$$\frac{\partial\gamma}{\partial t} = \text{sign}(u_n)\Delta u(\gamma)\vec{n} + F(\gamma)Curv(\gamma). \tag{1.14}$$

The steady states of the preceding equation yield an interesting equilibrium relation, if we notice that $F(\gamma)$ is nothing but the average contrast along the boundary.

We get what we can qualify a parameterless edge or image boundary equation, namely

$$\text{sign}(u_n)\Delta u(\gamma(s))\vec{n} + F(\gamma)Curv(\gamma(s)) = 0 \tag{1.15}$$

(this equation is derived in [KB01, KB02]). Notice that in this equation, $F(\gamma)$ is a global term, not depending upon s. The preceding equation is valid if $Du.\vec{n}$ does not change sign, which means that we observe no reverse contrast along γ. Recall also that $Curv(\gamma) = \gamma''(s)$ (s being the arclength). Since $F(\gamma) \geq 0$, Equation (1.14) is well posed and its implementation is very similar to the implementation of the *geodesic snake* equation, which also has a propagation term and a curvature term :

$$\frac{\partial\gamma}{\partial t} = f(\gamma)Curv(\gamma) - (Df.\vec{n})\vec{n}, \tag{1.16}$$

where here $f(\mathbf{x})$ denotes an "edge map", i.e. a function vanishing (or close to 0) where u has edges. Typically, $f(\mathbf{x}) = (1 + |Du(\mathbf{x})|)^{-1}$. We can interpret the geodesic snake equation by saying that the "snake" is driven by the second term $Df.\vec{n}$ towards the smaller values of f, while the first term roughly is a length minimizing term, but tuned by f. Thus, the snake tends by this term to shrink, but this smoother slows down where f is small, namely in a neighborhood of high gradient points for u.

2.3 Numerical scheme

We now describe in detail a numerical scheme implementing the maximization of (1.12). For a non-Euclidean parameterization $\gamma(p) : [a, b] \to I\!R^2$, the

energy we want to maximize writes

$$F(\gamma) = \frac{\int_a^b g\left(Du.\frac{\gamma'(p)^\perp}{|\gamma'(p)|}\right)|\gamma'(p)|\,dp}{\int_a^b |\gamma'(p)|\,dp}. \tag{1.17}$$

Rather than writing the Euler equation for (1.17) and *then* discretizing it, we discretize the energy and compute its exact derivative with respect to the discrete curve. Let us suppose that the snake is represented by a polygonal curve $M_1..M_n$ (either closed or with fixed endpoints). For the curve length, we can take

$$L = \sum_i \|\Delta_i\| \qquad \text{with} \qquad \Delta_i = M_{i+1} - M_i.$$

Now the discrete energy can be written $F = \frac{E}{L}$, with

$$E = \sum_i g(t_i)\|\Delta_i\|,$$

$$t_i = w_i \cdot \frac{\Delta_i}{\|\Delta_i\|}, \quad w_i = Du^\perp(\Omega_i), \quad \text{and} \quad \Omega_i = \frac{M_i + M_{i+1}}{2}.$$

Differentiating E with respect to M_k, we obtain

$$\nabla_{M_k} F = \frac{1}{L}\left(\nabla_{M_k} E - F\nabla_{M_k} L\right)$$

with

$$\nabla_{M_k} L = \frac{\Delta_{k-1}}{\|\Delta_{k-1}\|} - \frac{\Delta_k}{\|\Delta_k\|},$$
$$\nabla_{M_k} E = v_k + v_{k-1} + z_{k-1} - z_k,$$
$$v_i = \frac{1}{2}g'(t_i)D((Du^\perp)^T)(\Omega_i)\Delta_i,$$
$$z_i = g'(t_i)w_i + h(t_i)\frac{\Delta_i}{\|\Delta_i\|}$$

and $h(t) = g(t) - tg'(t)$. Note that

$$D((Du^\perp)^T) = D(-u_y \;\; u_x) = \begin{pmatrix} -u_{xy} & u_{xx} \\ -u_{yy} & u_{xy} \end{pmatrix}.$$

Numerically, we compute Du at integer points with a 3×3 finite differences scheme, and D^2u with the same scheme applied to the computed components of Du. This introduces a slight smoothing of the derivatives, which counterbalances a little the strong locality of the snake model. These estimations at integer points are then extended to the whole plane using a bilinear interpolation.

To compute the evolution of the snake, we use a two-steps iterative scheme.

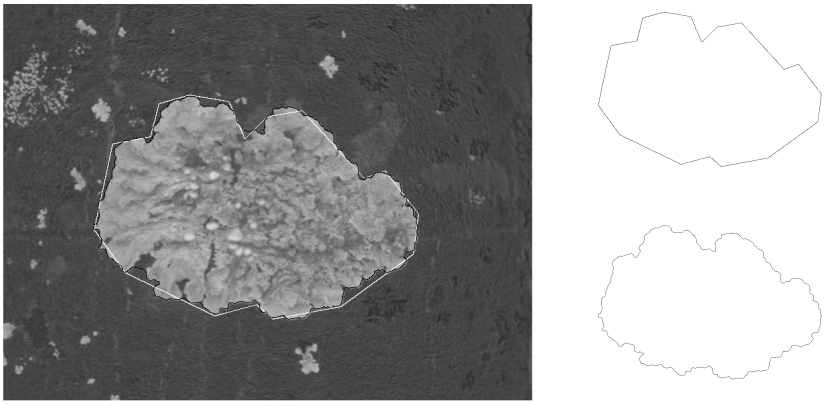

FIGURE 1. An initial contour drawn by hand on the lichen image (curve upright, in white on left image) and its final state (curve on the bottom right of left image, in black) computed with the average contrast snake model (1.12) for $g(t) = |t|$. *The evolution allows an important improvement in the localization of the boundary of the object, as illustrated by the energy gain (360%, from 8.8 to 40.8). This shows the usefulness of the snake model as an interactive tool for contour segmentation.*

1. The first step consists in a reparameterization of the snake according to arc length. It can be justified in several ways : aside from bringing stability to the scheme, it guarantees a *geometric* evolution of the curve, it ensures an homogeneous estimate of the energy, and it prevents singularities to appear too easily. However, we do not prevent self-intersections of the curve.

2. The second step is simply a gradient evolution with a fixed step. If $(M_i^n)_i$ represents the (polygonal) snake at iteration n and $(\tilde{M}_i^n)_i$ its renormalized version after step 1, then we set

$$M_i^{n+1} = \tilde{M}_i^n + \delta \nabla_{\tilde{M}_i^n} F.$$

A numerical experiment realized with this scheme is shown on Figure 1.

2.4 Choice of the function g

In this part, we shall show that the shape of the contrast function g is extremely relevant. Let us consider the average contrast functional (1.12), where g is an increasing function. This energy, which has to be maximized, is an arc length weighting of the contrast $g(u_n)$ encountered along the curve. In particular, it increases when the curve is lengthened by "adding" a high-contrasted part, or when it is shortened by "removing" a low-contrasted part (here, the qualities "high" and "low" are to be considered with respect

to the average contrast of the curve). Of course, these operations are not so easy to realize, because the snake must remain a Jordan curve, but we shall see now that this remark has consequences on the local and the global behavior of the snake.

It is all the more easy to increase the energy by shrinking the curve in low-contrasted parts that the function g increases more quickly. This means that if we want to be able to reach object contours presenting strong variations of contrast, we must choose a slowly increasing function for g. Let us illustrate this by a little computation, associated to the numerical experiment of Figure 2.

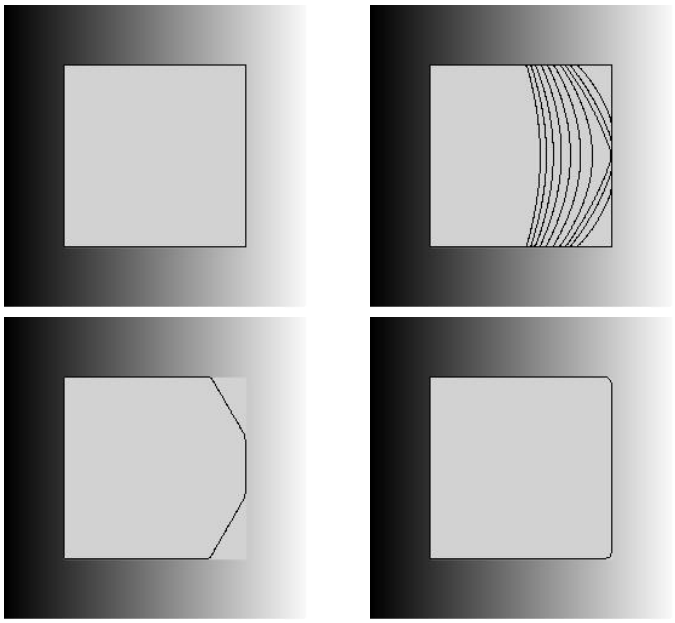

FIGURE 2. Influence of the function g for a synthetic image. *The snake model is applied to a synthetic image made of a bright square on a ramp background. Up left: original contour (can be detected as the unique maximal meaningful boundary of this image, see Section 3). Top, right: for $g(t) = |t|$, the snake collapses into a "flat curve" enclosing the left side of the square (some intermediate states are shown). Down, left: for $g(t) = |t|^{0.85}$, the snake converges to an intermediate state. Downright: for $g(t) = |t|^{0.5}$, the snake hardly moves, which means that the initial contour is nearly optimal despite the large difference of contrast between the left side and the right side of the square. Contours with large variations of contrast are more likely to be optimal curves for low powers, for which the energy is less sensitive to outliers.*

Consider a white square with side length 1 superimposed to a simple background image whose intensity is a linear function of x, varying from

black to light gray. If a and b are the values of u_n on the left and right sides of the square (we assume that $b < a$), and if γ is the boundary of the square, we have

$$F(\gamma) = \frac{E}{L}, \text{ with } L = 4 \text{ and } E = g(a) + g(b) + 2\int_0^1 g\big(a + (b-a)t\big)\, dt.$$

Now we would like to know if γ is an admissible final state of the snake model (that is a local maximum of F) or not. If we shrink a little the curve γ by "cutting" by $\varepsilon > 0$ the two right corners at 45, we obtain a curve γ_ε whose energy is

$$F_\varepsilon(\gamma) = \frac{E_\varepsilon}{L_\varepsilon}, \text{ with } L_\varepsilon = L - 4\varepsilon + 2\varepsilon\sqrt{2}$$

$$\text{and } E_\varepsilon = E - 2\int_{1-\varepsilon}^1 g\big(a + (b-a)t\big)\, dt - 2\varepsilon g(b) = E - 4\varepsilon g(b) + o(\varepsilon).$$

Since ε may be arbitrarily small, we know that γ cannot be optimal as soon as

$$\frac{L - L_\varepsilon}{L} > \frac{E - E_\varepsilon}{E}$$

for $\varepsilon > 0$ small enough. Using the previous estimates of E_ε and L_ε and the fact that $E > 3g(b) + g(a)$, we can see that this condition is satisfied for $\varepsilon > 0$ small enough as soon as

$$\frac{4 - 2\sqrt{2}}{4} > \frac{4g(b)}{3g(b) + g(a)},$$

which can be rewritten

$$\frac{g(a)}{g(b)} > \frac{2 + 3\sqrt{2}}{2 - \sqrt{2}} = 10.65...$$

Hence, the ratio $g(a)/g(b)$ must be kept (at least) below that threshold, and as small as possible in general in order to avoid the shrinkage of low-contrasted boundaries. If we choose a power function for g (that is $g(t) = |t|^\alpha$), this is in favor of a small value of α, as illustrated on Figure 2.

More generally, all this is in favor of a function g increasing as slowly as possible, that is to say almost constant. On the other hand, it is well known that all digital images have a level of noise (e.g. quantization noise to start with) which makes unreliable all gradients magnitudes below some threshold θ. This leads us to the following requirement.

Flatness contrast requirement for snakes. *The contrast function for snake energy must satisfy, for some θ :*

- *if $t \leq \theta$, $g'(t)$ is high ;*

- *if $t \geq \theta$, $g(t)$ is flat and $g(t) \to g(\infty)$, with $g(\infty) < \infty$.*

In the next section, we describe a way to compute θ as a meaningfulness threshold for the gradient, in function of the length of the curve.

Now we shall focus on the special family of power functions $g(t) = |t|^\alpha$. For any of these functions, the snakes method is zoom invariant, namely, up to time scale change, the snake's evolution remains the same when we zoom both the image and the initial contour. Conversely, this identity of snake evolutions leads to ask that the energy of a snake and the one of its zoomed counterpart are proportional. It is easy to prove that this implies that $g(\lambda t) = f(\lambda)g(t)$ for some function $f(\lambda)$. If g is continuous and even, this implies that $g(t) = C|t|^\alpha$.

According to the requirements above, we can predict the following behaviors.

Experimental predictions. *Consider the average contrast energy*

$$F(\gamma) = \frac{1}{L(\gamma)} \int_0^{L(\gamma)} g(u_n(\gamma(s))ds, \qquad \text{with} \qquad g(t) = |t|^\alpha. \tag{1.18}$$

- *When α is large, all snakes present straight parts and shrink from the parts of the boundaries with weaker gradient.*
- *The smaller α is the better : snakes will be experimentally more stable and faithful to perceptual boundaries when $\alpha \to 0$.*

We checked these predictions on several numerical experiments. First, in a synthetic case (a white "comb" on a ramp background), we can see the evolution of the snake (Figure 3) and its final state in function of the power α (Figure 4). As expected, some teeth of the comb are not contrasted enough to be kept for $\alpha = 1$, and a smaller value is required to have an accordance between the snake and the perceptual boundary.

Our predictions remain true when the snake model (1.18) is applied to the "real" image of a bird (see Figure 5) : as expected, we notice that the low contrasted parts of the contour are kept only when the power α is small and replaced by straight lines when α is large.

The trend to favor high contrasted parts in the snake model, which becomes very strong for large powers, has some consequences on the numerical simulations. As we noticed before, if the contrast is not constant along the curve one can always increase the average contrast (1.12) by lengthening the curve in the part of the curve which has the highest contrast. If the curve γ was not required to be a regular Jordan curve, there would be no nontrivial local maximum of the functional $F(\gamma)$, since the curve could "duplicate" itself infinitely many times in the highest contrasted part by creating cusps. This formation of cusps was proven by Mumford and Shah [MS89] in the case $g(t) = t^2$. In the model we presented, cusps may be

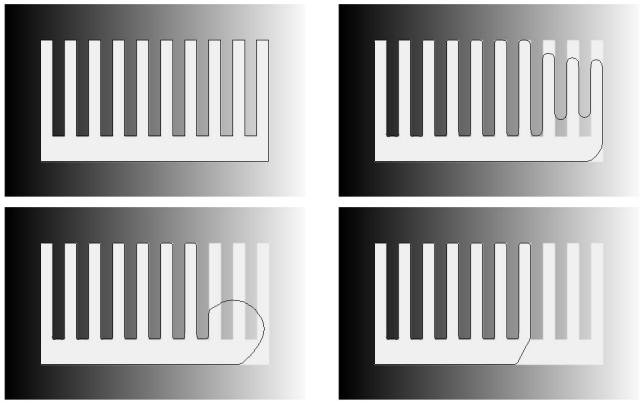

FIGURE 3. Evolution of the snake for $g(t) = |t|$. *The snake model is applied to a synthetic image made of a bright comb on a ramp background. Upleft: original contour (detected as the unique maximal meaningful boundary of this image).*

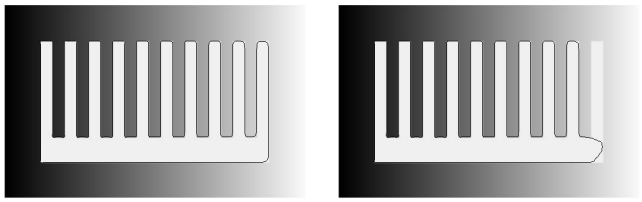

FIGURE 4. Influence of the function g for a synthetic image. *The snake model is applied to a synthetic image made of a bright comb on a ramp background.* $g(t) = |t|^{0.4}$ *(left),* $g(t) = |t|^{0.55}$ *(right),* $g(t) = |t|$ *(bottom, right of Figure 3).*

avoided because the duplication of the curve cannot be realized directly for a geometric curve evolution, which consists in infinitesimal *normal* deformations. In general, the creation of a cusp with a normal curve evolution will not be compatible with the need for the energy to increase during the evolution, so that in many cases the snake will not be able to fold itself during the energy maximization process.

Numerically, this effect is more difficult to avoid since the curve does not evolve continuously but step by step, so that the energy gap required to develop a cusp may be too small to stop the process. This is especially true for large powers, for which self-folding is often observed in practice (though, of course, this phenomenon depends a lot on the time step used in the gradient maximization scheme). In the bird image for example, one may notice a "thick part" of the curve for $\alpha = 3$ (Figure 6), which corresponds to such a self-folding. The numerical experiment of Figure 7, which is in some way the "real case" analog of Figure 2, shows that curve duplication may also be attained continuously in the energy maximization process (the

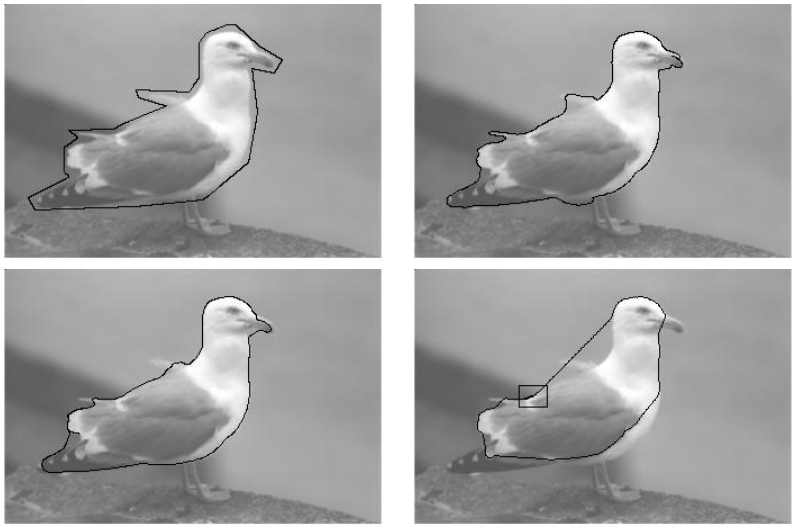

FIGURE 5. Contour optimization in function of g for the bird image. *An initial contour (up, left) was first designed by hand. Then, it was optimized by the snake model for different functions* g *:* $g(t) = |t|^{0.5}$ *(upright),* $g(t) = |t|$ *(down, left), and* $g(t) = |t|^3$ *(downright). As the power increases, the snake becomes more sensitive to edges with high contrast and smooth (or cut) the ones with low contrast. (original bird image from F. Guichard)*

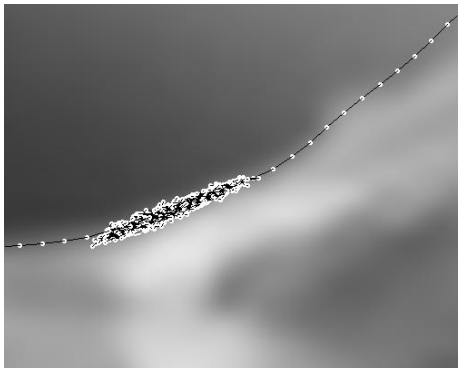

FIGURE 6. Zoom on the curve duplication that appears for $g(t) = |t|^3$ in the highest contrasted part (rectangle drawn on the bottom, right image of Figure 5). The discretization of the snake (black curve) is shown by white dots.

initial region enclosed by the snake collapses into a zero-area region enclosed by a "flat curve").

In the next section, we introduce a boundary detector which we recently developed and which we shall prove to be theoretically close and experimen-

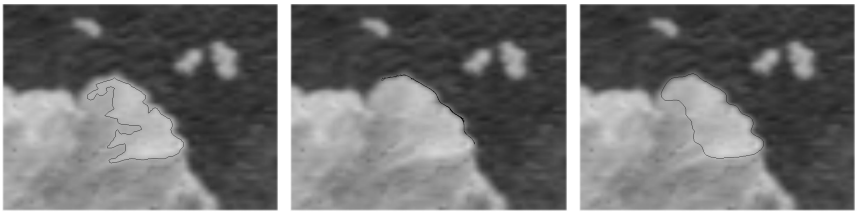

FIGURE 7. Influence of the function g in the self-folding phenomenon. *The initial boundary (left) is a maximal meaningful boundary of the lichen image. Like for the square image (see Figure 2), the contrast along this curve present strong variations. The snake collapses into a self-folded "flat curve" with two cusps for* $g(t) = |t|$ *(middle) but remains a Jordan curve for* $g(t) = \sqrt{|t|}$ *(right).*

tally almost identical to the Kimmel-Bruckstein detector, when g satisfies the flatness requirement.

3 Meaningful boundaries

In very much the same way as different species of animals can converge to the same morphology by the way of the evolutive pressure[1], two image analysis structures with very different origins, namely the variational snakes and the maximal meaningful level lines, arrive at almost exactly the same numerical results. Level lines for image representation have been proposed in [CCM99] as an efficient *contrast invariant* representation of any image. This representation stems from Mathematical Morphology [Ser82] where connected components of level sets are extensively used as image features and indeed are contrast invariant features (level lines are nothing but the boundaries of level sets). Level lines are closed curves, like the snakes. They are a complete representation of the image. They have a tree structure which permits a fast computation, the so called Fast Level Set Transform [MG00]. In addition, they satisfy the easy topological change requirement : their topology changes effortless at saddle points and permit level lines to merge or to split numerically by just evolving their level. The only drawback of level lines is : they are many. Let us now describe a pruning algorithm proposed in [DMM01] to reduce drastically the number of level lines without changing the image aspect and having thus an image representation *with only a few closed curves*. The result consists roughly in giving all essential (in some information theoretical sense) best contrasted level lines. Two examples of this zero-parameter method are given in Fig-

[1] In Australia, some marsupials have evolved into a wolf-like predator species : the Tasmanian Thylacine (Thylacinus cynocephalus). This species unfortunately disappeared in 1936, but good drawings and photographs are available.

ures 8 and 9.

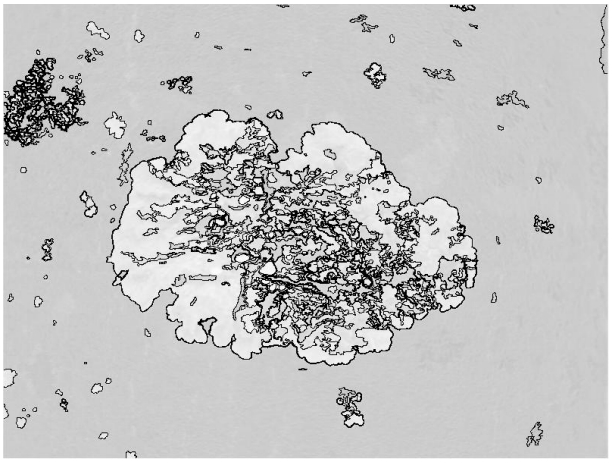

FIGURE 8. The maximal meaningful boundaries of the lichen image (superimposed in grey, see Figure 1).

FIGURE 9. The maximal meaningful boundaries of the bird image (superimposed in grey, see Figure 5).

Let u be a discrete image, of size $N \times N$. We consider the level lines at quantized levels. The quantization step is chosen in such a way that level lines make a dense covering of the image: if e.g. this quantization step is 1 and the natural image ranges 0 to 255, we get such a dense covering of the image. A level line can be computed as a Jordan curve contained in the

boundary of an upper (or lower) level set with level λ,

$$\chi_\lambda = \{\mathbf{x},\ u(\mathbf{x}) \leq \lambda\} \quad \text{and} \quad \chi^\lambda = \{\mathbf{x},\ u(\mathbf{x}) \geq \lambda\}.$$

It can also be given a simple linear spline description if one uses a bilinear interpolation : in that case, level lines with level λ are solved by explicitly solving the equation $u(\mathbf{x}) = \lambda$ [LMR].

For obvious stability reasons, we consider in the following only level lines along which the gradient is not zero. What follows is a very fast summary of the theory developed in [DMM01].

Let $\mathcal{L}$ be a level line of the image u. We denote by l its length counted in number of "independent" points. In the following, we will consider that points at a geodesic distance (along the curve) larger than 2 are independent. Let $\mathbf{x}_1$, $\mathbf{x}_2$,... $\mathbf{x}_l$ denote the l considered points of $\mathcal{L}$. For a point $\mathbf{x} \in \mathcal{L}$, we will denote by $c(\mathbf{x})$ the contrast at $\mathbf{x}$. It is defined by

$$c(\mathbf{x}) = |Du|(\mathbf{x}), \tag{1.19}$$

where Du is computed by a standard finite difference on a 2×2 neighborhood. For $\mu > 0$, we consider the event : *"for all $i = 1..l$, $c(\mathbf{x}_i) \geq \mu$, i.e. each point of $\mathcal{L}$ has a contrast larger than μ"*. From now on, all computations are performed in the Helmholtz framework explained in [DMM00]: we make all computations as though the contrast observations at $\mathbf{x}_i$ were mutually independent. If the gradient magnitudes of the l points were independent, the probability of this event would be $P[c(\mathbf{x}_1) \geq \mu] \cdot P[c(x_2) \geq \mu] \cdot \ldots \cdot P[c(x_l) \geq \mu] = H(\mu)^l$, where $H(\mu)$ is the empirical probability for a point on any level line to have a contrast larger than μ. Hence, $H(\mu)$ is given by the image itself,

$$H(\mu) = \frac{1}{M} \#\{\mathbf{x} \,/\, |Du|(\mathbf{x}) \geq \mu\}, \tag{1.20}$$

where M is the number of pixels of the image where $Du \neq 0$. In order to define a meaningful event, we have to compute the expectation of the number of occurrences of this event in the observed image. Thus, we first define the number of false alarms.

Definition 1 (Number of false alarms) *Let $\mathcal{L}$ be a level line with length l, counted in independent points. Let μ be the minimal contrast of the points $\mathbf{x}_1$,... $\mathbf{x}_l$ of $\mathcal{L}$. The number of false alarms of this event is defined by*

$$NF(\mathcal{L}) = N_{ll} \cdot [H(\mu)]^l, \tag{1.21}$$

where N_{ll} is the number of level lines in the image.

Notice that the number N_{ll} of level lines is provided by the image itself. We now define ε-meaningful level lines.

Definition 2 (ε-meaningful boundary) *A level line $\mathcal{L}$ with length l and minimal contrast μ is ε-meaningful if $NF(\mathcal{L}) \leq \varepsilon$.*

We can summarize the method in the following way : not all level lines are meaningful ; some can cross flat regions where noise predominates. In order to eliminate such level lines, we first compute the minimum gradient, μ, on a given level line $\mathcal{L}$ with length l. We then compute the probability of the following event : a level line in a white noise image with the same gradient histogram as our image has contrast everywhere above μ. This probability is the probability of a "contrasted level line happening by chance". We then compute the false alarm rate, namely the product of this probability by the number of tested level lines. If the false alarm rate is less than ε, the level line is said to be ε-meaningful. In practice, one takes $\varepsilon = 1$, since one does not care much having on the average one wrong among the many meaningful level lines.

Because of the image blur, contrasted level lines form bundles of nearly parallel lines along the edges. We call *interval of level lines* a set of level lines such that each one is enclosed in only one, and contains only another one.

Definition 3 *We say that a level line is maximal meaningful if it is meaningful and if its number of false alarms is minimal among the level lines in the same interval.*

With this definition, one should retain in the simple case of a contrasted object again a uniform background, a single maximal meaningful level line. In all experiments below, we only display maximal meaningful level lines.

4 Snakes versus Meaningful Boundaries

In this section, we would like to compare the snake model and the meaningful boundaries model (abbreviated MB).

In terms of *boundary detection*, the MB model has one serious advantage : it is fully automatic. Comparatively, realizing boundary detection with the snake model is much more difficult and an automatic algorithm (that is, with no parameters to set) seems unreachable. Indeed, there are so many ways to change the large set of local maxima of the snake functional that an important user interaction is needed to ensure a good compromise between false detections and missed contours. Among the parameters to set are :

- the function g (we have shown that not many possibilities are left to obtain reliable detections. We were led to choose $g(t) = |t|^\alpha$ with some small α) ;
- the initial contour, which represents a high number of parameters. Starting with a set of fixed contours (e.g. a covering of the image with circles with different radii) requires a real multi-scale strategy

and a strong smoothing of the image since the actual contours may be quite different from the ones fixed *a priori*;

- the parameters of the numerical scheme used to implement the snake, including the gradient step and the above-mentioned initial smoothing, required for non-interactive detection. The set of maxima of the discrete snake functional depends heavily on them in general, as shown (for the gradient step) on Figure 10.

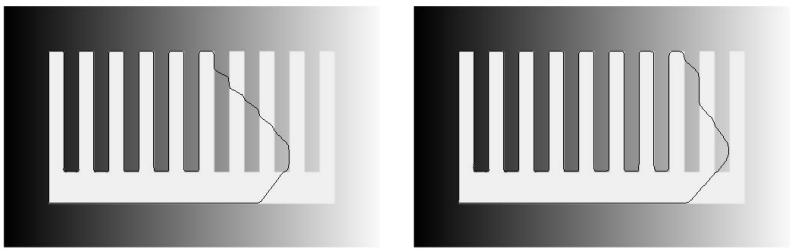

FIGURE 10. Sensitivity of the snake model to numerical parameters (here the time step used for the gradient descent). *The snake model* ($g(t) = |t|$) *is applied for several values of the gradient step* δ *:* $\delta = 1$ *(left),* $\delta = 3$ *(Figure 3, downright),* $\delta = 10$ *(right). Due to the huge number of local maxima of the snake functional, the final result is very sensitive to the numerical implementation of the model, in particular to* δ.

In terms of *boundary optimization*, that is, the refinement of a raw automatically detected or interactively selected contour, the snake model does not bring very substantial improvements compared to the MB model. We checked this by applying the snake model to the contours detected by the MB model (Figure 11 and 12). The very little changes brought in these experiments by the snake evolution prove that the curves detected with the MB model are very close to local maxima of the snake model. This is not very surprising since first, the curves delivered by the MB model are level lines (that is, curves for which Du is colinear to the normal $\vec{n}$ of the curve at each point), and second, the MB criterion is, like the snake model, based on gradient maximization. This is all the more true that as we discussed previously, the function g used in the snake model should be close to the *gradient thresholding* (that is, a step function selecting points with large enough gradient) realized by the MB model.

In our opinion, these experiments tend to prove that the snake model should be only used when interactive contour selection and optimization is required and when, in addition, the sought object presents contrast inversions. In all other cases and in particular for automatic boundary detection, the meaningful boundaries method appears to be much more practicable.

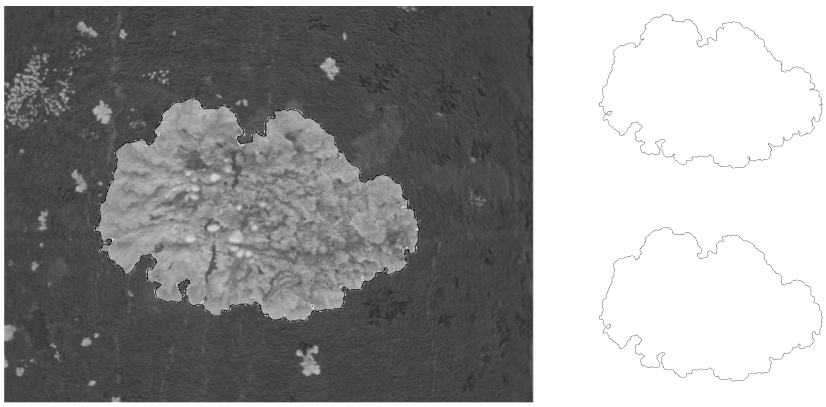

FIGURE 11. Optimization of a maximal meaningful boundary by the snake model. *The contour optimization brought by the snake model (here $g(t) = |t|$) is generally low when the contour is initialized as a contrasted level line (here a maximal meaningful boundary). In this experiment, the total energy is only increased by 17%, from 34.6 to 40.6.*

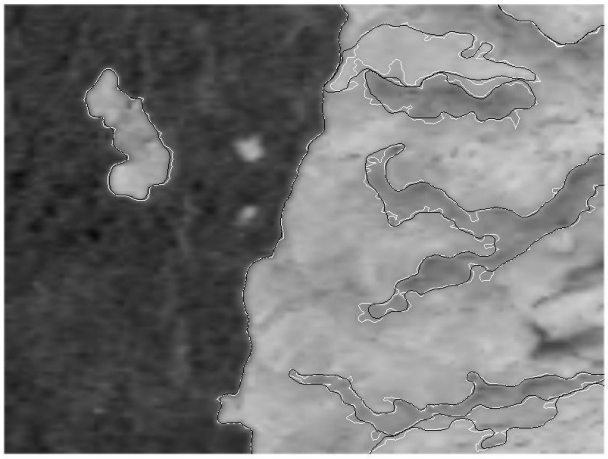

FIGURE 12. Optimization of all maximal meaningful boundaries ($g(t) = \sqrt{|t|}$). *The "objects" in this image are well detected with the maximal meaningful boundaries (in white). The optimization of these contours brought by the snake model (in black) is quite low, as shown by the little gain obtained for the total energy (sum of the energy of each curve) : 7%, from 35.9 to 38.5.*

Acknowledgments

Work partially supported by Office of Naval Research under grant N00014-97-1-0839, Centre National d'Etudes Spatiales, Centre National de la Recherche Scientifique et Ministère de la Recherche et de la Technologie. We thank the Fondation des Treilles for its hospitality during the writing of this paper.

Agnes.Desolneux@cmla.ens-cachan.fr
Lionel.Moisan@cmla.ens-cachan.fr
Jean-Michel.Morel@cmla.ens-cachan.fr
Ecole Normale Supérieure de Cachan,
61 avenue du président Wilson,
94235 Cachan Cedex, France.

5 References

[Can86] J.F. Canny. A computational approach to edge detection. *IEEE Trans. on PAMI*, 8:679–698, November 1986.

[CCM99] V. Caselles, B. Coll, and J.-M Morel. Topographic maps and local contrast changes in natural images. *Int. J. of Computer Vision*, 33(1):5–27, 1999.

[CKS97] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International Journal of Computer Vision*, 1(22):61–79, 1997.

[DMM00] A. Desolneux, L. Moisan, and J.-M. Morel. Meaningful alignments. *International Journal of Computer Vision*, 40(1):7–23, 2000.

[DMM01] A. Desolneux, L. Moisan, and J.-M. Morel. Edge detection by Helmholtz principle. *Journal of Mathematical Imaging and Vision*, 14(3):271–284, May 2001.

[FL90] P. Fua and Y.G. Leclerc. Model driven edge detection. *Machine Vision and Applications*, 3:45–56, 1990.

[Har84] R. Haralick. Digital step edges from zero crossing of second derivatives. *IEEE Trans. PAMI*, 6(1):58–68, January 1984.

[KB01] R. Kimmel and A. Bruckstein. Regularized laplacian zero crossings as optimal edge integrators. In *Proceedings of Image and Vision Computing, IVCNZ01, New Zealand*, november 2001.

[KB02] R. Kimmel and A. Bruckstein. On edge detection, edge integration and geometric active contours. In *Proceedings of Int.*

Symposium on Mathematical Morphology, ISMM 2002, Sydney, New South Wales, Australia, April 2002.

[LMR] J.-L. Lisani, P. Monasse, and L. Rudin. Fast shape extraction and applications. submitted to PAMI, july 2001.

[Mar72] A. Martelli. Edge detection using heuristic search methods. *Comp. Graphics Image Processing*, 1:169–182, 1972.

[MG00] P. Monasse and F. Guichard. Fast computation of a contrast-invariant image representation. *IEEE Transactions on Image Processing*, 9(5):860–872, 2000.

[MH80] D. Marr and E. Hildreth. Theory of edge detection. *Proc. Royal Soc. Lond.*, B 207:187–217, 1980.

[MKT87] A. Witkin M. Kass and D. Terzopoulos. Snakes: active contour models. *Int. Comp. Vis. Conf., IEEE*, n$^{\text{O}}$ 777, 1987.

[Mon71] U. Montanari. On the optimal detection of curves in noisy pictures. *CACM*, 14(5):335–345, May 1971.

[MS89] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics*, XLII(4):577–685, 1989.

[Ser82] J. Serra. *Image Analysis and Mathematical Morphology*. Academic Press, 1982.