INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Higher Order Active Contours

Marie Rochery — Ian Jermyn — Josiane Zerubia

N° 5656

Août 2005

Thème COG

Rapport de recherche

ISSN 0249-6399

Higher Order Active Contours

Marie Rochery , Ian Jermyn , Josiane Zerubia*

Thème COG — Systèmes cognitifs
Projet Ariana

Rapport de recherche n° 5656 — Août 2005 — 29 pages

Abstract: We introduce a new class of active contour models that hold great promise for region and shape modelling, and we apply a special case of these models to the extraction of road networks from satellite and aerial imagery. The new models are arbitrary polynomial functionals on the space of boundaries, and thus greatly generalize the linear functionals used in classical contour energies. While classical energies are expressed as single integrals over the contour, the new energies incorporate multiple integrals, and thus describe long-range interactions between different sets of contour points. As prior terms, they describe families of contours that share complex geometric properties, without making reference to any particular shape, and they require no pose estimation. As likelihood terms, they can describe multi-point interactions between the contour and the data. To optimize the energies, we use a level set approach. The forces derived from the new energies are non-local however, thus necessitating an extension of standard level set methods. Networks are a shape family of great importance in a number of applications, including remote sensing imagery. To model them, we make a particular choice of prior quadratic energy that describes reticulated structures, and augment it with a likelihood term that couples the data at pairs of contour points to their joint geometry. Promising experimental results are shown on real images.

Key-words: active contour, shape prior, geometric, higher-order, polynomial, quadratic, road network, remote sensing

The authors would like to thank CNES and IGN for the use of the images. This work was partially supported by by NATO/Russia CLG 980107; EU project MOUMIR (HP-99-108, http://www.moumir.org); and by EU project MUSCLE (FP6-507752).

* Email: FirstName.LastName@sophia.inria.fr

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

Contours Actifs d'Ordre Supérieur

Résumé : Nous introduisons une nouvelle classe de contours actifs qui offre des perspectives intéressantes pour la modélisation des régions et des formes, et nous appliquons un cas particulier de ces modèles à l'extraction de réseaux linéiques dans des images satellitaires et aériennes. Les nouveaux modèles sont des fonctionnelles polynômiales arbitraires sur l'espace des contours, et généralisent ainsi les fonctionnelles linéaires utilisées dans les modèles classiques de contours actifs. Alors que les fonctionnelles classiques s'écrivent avec de simples intégrales sur le contour, les nouvelles énergies sont définies comme des intégrales multiples, décrivant ainsi des interactions de longue portée entre les différents ensembles de points du contour. Utilisées comme des termes d'a priori, les fonctionnelles décrivent des familles de contours aux propriétés géométriques complexes, sans faire référence à une forme spécifique et sans nécessiter l'estimation de la position. Utilisées comme des termes d'attache aux données, elles permettent de décrire des interactions multi-points entre le contour et les données. Afin de minimiser ces énergies, nous adoptons la méthodologie des courbes de niveau. Les forces dérivées des énergies sont cependant non locales, et nécessitent une extension des méthodes de courbes de niveau standard. Les réseaux sont une famille de formes d'une grande importance dans de nombreuses applications et en particluier en télédétection. Pour les modéliser, nous faisons un choix particulier d'énergie quadratique qui décrit des structures branchées et nous ajoutons un terme d'attache aux données qui lie les données et la géométrie du contour au niveau des paires de points du contour. Des résultats d'extraction prometteurs sont montrés sur des images réelles.

Mots-clés : contour actif, a priori sur la forme, géométrique, ordre supérieur, polynomial, quadratique, réseau routier, télédétection

Contents

1 Introduction

The task of image processing algorithms is to assert propositions about images, propositions that typically concern not the images themselves, but the 'scene' of which the image is a representation. Amongst the many varieties of propositions one can make, one of the most common consists of those formed using the predicate 'The volume that projects to region R in the image domain has properties P'. Examples of properties include labels naming entities in the scene (person, John, car, road, forest, building) or physical parameters of those entities (depth, illumination, reflectance). In real applications, we possess significant prior knowledge K about both R and P, and their relationship to the image. Examples of K include knowledge of the smoothness or specific form of the depth map; knowledge of the textural and reflectance properties of the entity, and hence its appearance in the image; and, of particular interest here, knowledge of the likely shape of the region occupied in the image by an entity with a given label. The central quantity of interest is then the probability distribution $\mathrm{P}(\langle R, P\rangle | I, K)$ over these propositions given the image data I and all prior knowledge K. This distribution describes our knowledge of the propositions, from which, if required, point estimates of R and P can be extracted. The information K is typically crucial for solving real problems, so that as much knowledge as possible should be encoded, both about R and P (as prior terms $\mathrm{P}(\langle R, P\rangle | K)$) and about their relations to the data (as likelihood terms $\mathrm{P}(I | \langle R, P\rangle, K)$). Attempts to assert such propositions thus have to construct, if only implicitly, probability distributions on the space of regions, $\mathrm{P}(R)$, which may depend on other, known or unknown parameters, and the data.

One way of implicitly constructing such distributions is provided by active contours.[1] An 'energy' functional is defined on the space of regions, and then minimized. With some reservations, this can be regarded as the computation of a MAP estimate using the negative logarithm of the probability density. Previous active contour models can be divided into two classes, which are reviewed in sections 1.1.1 and 1.1.2. The first class includes only relatively trivial information about R: in the prior terms only its boundary length and area enter (and sometimes integrals of curvature), whereas in the likelihood terms the data is coupled to the geometry at one point only. In the second class, more information is introduced in the prior terms: specific shapes are modelled by defining a 'mean' shape and typical variations around it.

In this report, we describe a new class of active contour models that subsumes the first class of energy functionals as special cases (and perhaps the second as well), and greatly generalizes them.[2] The new models allow the incorporation of sophisticated prior information about region geometry, and the construction of likelihood terms that describe complex interactions between the geometry of the region and the data. This is achieved by a large generalization of the space of energy functionals considered. In a way to be made precise, previous functionals are *linear* on a certain space containing the space of curves:; they are expressed in terms of single integrals over the contour, and thus can incorporate only local interactions between contour points and hence only very weak prior information K about region geometry R or the relation of this geometry to the data.

[1]More generally, deformable models. In this report, we focus on the case of codimension one objects in two dimensions, but the techniques we use to construct functionals are not limited to this case. They can be applied to objects of any codimension in any number of dimensions.

[2]These models were first introduced by Rochery et al. (2003a, b), while subsequent developments are the subject of Rochery et al. (2004, 2005).

In contrast, the new functionals consist of *arbitrary polynomials* on this space; they include multiple integrals over the contour. These functionals can describe arbitrarily long-range interactions between subsets of points in the boundary, or between subsets of points and the data: quadratic energies describe interactions between pairs of points, cubic energies between triples, and so on. In the case of quadratic energies, for example, there are two integrals, which means that for every pair of points there is a contribution to the energy that depends on the geometry (and perhaps the data) at those two points. Equation 1.1 shows an Euclidean invariant quadratic energy term:

$$E(C) = \int\int dp\, dp'\, \mathtt{t}(p) \cdot \mathtt{t}(p')\ \Psi(|C(p) - C(p')|)\ . \tag{1.1}$$

Here, p and p' are coordinates on $\square C$, the domain of the curve C; $\mathtt{t}(p) = \partial_p C(p)$ is the tangent vector to C at p; $|x - y|$ is the Euclidean distance between points x and y in the image domain Ω; and the function Ψ weights the interactions between different points of the curve according to their distance, and must be chosen carefully since it defines the geometrical content of the model.

It is clear from equation (1.1) that even in the quadratic case, the use of higher-order energies opens up a much wider range of modelling possibilities than previously possible. Only two Euclidean invariant linear terms exist if curvature is not included: length and area. In contrast, equation 1.1 shows that there is a whole function space full of Euclidean invariant quadratic terms, and higher-order Euclidean invariant add a great deal more flexibility. Due to their inherent invariance, these obviate the necessity for pose estimation involved in the second class mentioned above, yet they can describe families of contours with complex shape properties. We will describe the general framework of the new models in section 2.

In order to minimize the new energies, we use a level set approach. The implementation of level sets for the new energies requires an extension of standard techniques, however, because the forces derived from the new energies are non-local: the speed of a point in the boundary depends on the whole of the boundary and not just on its infinitesimal neighbourhood. The resulting algorithms are described in section 3.

Networks constitute a specific family of shapes that share complex qualitative and quantitative properties and are of great importance to image processing problems in many domains, for example in remote sensing (road and hydrographic networks) and medical imaging (vascular and other physiological networks). We apply a specific instance of the new models to the problem of road network extraction in section 4, where we also discuss previous work on this problem. We construct a prior energy in the new class that describes network shapes, and likelihood terms in the new class that describe more complex relations between the contour and the data than previously possible. We conclude in section 5.

1.1 Previous work

In this section, we review previous work, dividing it into the two classes mentioned in the Introduction, linear energies and shape modelling, and putting the emphasis on the type of prior information included.

1.1.1 Linear energies

The original paper on active contours was by Kass et al. (1988). The energy defined there is parameterization dependent, but if the parameterization is taken to be arc length, then the energy used is the sum of boundary length and the integral of boundary curvature, plus the negative of the integral of image gradient magnitude. 'Balloon forces' (a constant pressure, which can be viewed as generated by adding the region area to the energy) were introduced by Cohen (1991) to improve the stability of results by 'pushing' the region boundary past shallow local minima caused by weak image gradients. 'Geometric' or 'geodesic' active contours (Malladi et al., 1995; Caselles et al., 1993, 1997; Kichenassamy et al., 1995) removed the parameterization dependence of the early models by using as energy the length of the boundary in a non-Euclidean metric on Ω determined by the image. Most of these energies were written as the integrals of functions over the boundary of the region, but Chan and Vese (2001), Paragios and Deriche (2002), and Jehan-Besson et al. (2003), among others, introduced integrals of functions over the interior to facilitate the description of region properties and to reduce sensitivity to noise and clutter.

All the energy functionals used in the above work, both prior and likelihood terms, are representable as algebraic combinations of single integrals over the boundary of the region or over its interior. Such functionals are 'linear', for reasons to be explained in section 2. The limitation of such functionals is that they incorporate only local interactions. In the case of a finite-dimensional vector space X, this is clear. Linear functionals $x \cdot a$, $x \in X$, $a \in X^*$, lead to exponential probability distributions, $\mathrm{P}(x|a) \propto \exp(-x \cdot a)$. In any basis, this takes the form $\exp(-\sum_i x^i a_i) = \prod_i \exp(-x^i a_i)$. Thus x^i and x^j are independent for all $i \neq j$.

For linear functionals on the space of curves, the situation is similar. Linear functionals incorporate only local interactions, where local means constructed from derivatives of the curve at each point. This notion of locality is closely related to the property of Markovianity. In the discrete case, the dependence on derivatives means that interactions take place within fixed size neighbourhoods, as in a Markov random field. In addition, because the degree of the derivatives involved is typically small, the neighbourhoods are small.

The result of this limitation is impoverished modelling, especially in the prior terms. If we only allow first derivatives, then the only two Euclidean invariant terms are length and area. Thus any two boundaries that share length and area are equiprobable from the point of view of these models. The limitation imposes itself equally on likelihood terms, although there the lack of Euclidean invariance allows a wider variety. Nevertheless, such terms can only express the likely configurations of the data given the geometry at a single point of the curve.

1.1.2 Shape modelling

In order to get around this limitation on prior terms, various approaches have been taken to the incorporation of more sophisticated information. Leventon et al. (2000) represent shapes as signed distance functions, and use a Gaussian distribution on the principal components of variation around the mean distance function acquired from training data as a shape prior. Cremers et al. (2001) modify the Mumford-Shah functional to incorporate statistical shape knowledge. They use an explicit parameterization of the contour as a closed spline curve, and learn a Gaussian probability distribution for the spline control point vectors. The statistical prior restricts the contour deformations to the subspace of learned deformations. Paragios and Rousson (2002) propose a functional that can account for the global and local shape properties of the target object. A prior shape model is built using aligned training examples. A probabilistic framework uses the shape image and the variability of shape deformations as unknown variables. They seek a global transformation and a level set representation that maximizes the posterior probability given the prior shape model. Chen et al. (2001) define an energy functional depending on the gradient and the average shape of the target object. The prior shape term evaluates the similarity of the shape of the contour to that of the reference shape through the computation of a distance function using the Fast Marching method of Sethian (1996). Foulonneau et al. (2003) define shape descriptors with Legendre moments and introduce a geometric prior in the framework of region-based active contours, with a quadratic distance function between the set of moments of the contour and the set of moments of the reference object. In an interesting piece of work, Steiner et al. (1998) use a regularized inverse diffusion to exaggerate the properties of given shapes.

What the above models have in common, is that they are looking for a single instance of a specific shape in an image. Given one or more training examples, and a shape representation, a 'mean' shape is computed. The evolution of the contour is then constrained by this 'mean' shape and the possible deformations around this shape. This is effective in some circumstances, but these approaches rapidly become restrictive if there are several instances of the shape to detect in the image, or if the regions to be extracted cannot be defined as small variations around a 'mean' shape. Consider 'network' shapes. These possess complex geometric properties in common (they are composed of 'arms' of roughly parallel sides, perhaps of varying width, joined together in various ways), but their variability cannot be reduced to perturbations of a template shape parameterized by a few quantities. Nevertheless it is clearly important from a modelling point of view to incorporate the geometrical properties that they share; what might be called their 'family resemblance'.

2 New models: general framework

With the aim of modelling such families, and of extending the expressive power of active contour models more generally by introducing a coherent way to construct functionals of increasing complexity, we introduce a new class of active contour models. These have been described at an intuitive level in section 1. In this section, we present the new class of energies in detail. In section 2.1, we formalize the notion of linearity as it applies to the energies of section 1.1.1. In section 2.2, we use this notion to define higher-order polynomial functionals, with the aim both of demonstrating the degree to which the new functionals generalize the linear models, and of providing a language for the construction of higher-order models. In section 2.3, we specialize to the case of quadratic functionals and give detailed equations for the various forms these can take. In section 2.4, we discuss a particular example of a higher-order energy and illustrate its properties.

2.1 Linear functionals

As already stated, the energy functionals of section 1.1.1, are representable as algebraic combinations of single integrals over the boundary of the region or over its interior. Such integrals represent *linear* or *twisted linear* functionals on the spaces of 1-boundaries and 2-chains (Bossavit, 2002). Chains are equivalence classes of formal linear combinations of differentiable embeddings of rectangles, *e.g.* the interval (1-chains) or the unit square (2-chains). 'Boundaries' in a generalized sense are then defined by the action of a boundary operator ∂ taking n-chains to $(n-1)$-chains. In the plane, 1-boundaries (1-chains in the image of ∂) are equivalent to closed 1-chains (those in the kernel of ∂, and thus without boundary). Consequently, we will reserve the term 'boundary' for the geometric boundary of a region, and use the word 'closed' to indicate boundaries in this generalized sense.

The utility of these formal objects is to characterize properties of curves and curve functionals in algebraic terms. A functional on chains is 'linear' in the standard sense: given a linear combination of chains $\alpha C_1 + \beta C_2$, the value of the functional is given by the same linear combination of the values of the two chains:

$$E(\alpha C_1 + \beta C_2) = \alpha E(C_1) + \beta E(C_2) \; . \tag{2.1}$$

Note that by definition two embeddings C_1 and C_2 with the same domain D represent the same chain if $C_2 = C_1 \epsilon$, for some diffeomorphism $\epsilon : D \to D$. Functionals defined on the space of embeddings must therefore be invariant under diffeomorphisms in order to project to well-defined functionals on chains. This invariance requirement means that differential forms are the natural language in which to represent such functionals. Linear functionals on 1-chains thus take the form

$$E(C) = \int_{\partial R} A = \int_{\Box C} C^* A = \int dp \, \mathbf{t}(p) \cdot A \; , \tag{2.2}$$

where A is a 1-form on Ω; $v \cdot A$ denotes the evaluation ('inner product') of the 1-form A on the vector v; and C^* is pullback by C. This form of functional can be given a physical interpretation as the energy of a current-carrying wire in an external electromagnetic potential A (although of course this potential is not limited by Maxwell's equations).

Using the generalized Stokes theorem, such functionals can be rewritten as integrals over R. Equally importantly, since in two dimensions every 2-form is closed and in the plane every closed form is exact because the cohomology is trivial, the reverse is true. For every 2-form F there exists a 1-form A_F such that $F = dA_F$, meaning that every energy of the form $\int_R F$, where R is a region, or more generally a 2-chain, can be rewritten as

$$\int_R F = \int_R dA_F = \int_{\partial R} A_F \; . \tag{2.3}$$

The area of the interior of a closed 1-chain provides one example of this process. In this case, $F = \star_g \mathbb{I}$, where $\mathbb{I}$ is the function identically equal to one everywhere, g is a metric on Ω, and $\star_g$ is the Hodge operator that converts functions to 2-forms. In an Euclidean metric, this becomes

$$E(C) = \frac{1}{2} \int dp \, \mathbf{t}(p) \times C(p) = \int dp \, \partial_p x(p) \, y(p) \; , \tag{2.4}$$

where (x, y) are Euclidean coordinates. In consequence of equation (2.3), linear energies of the form (2.2) encompass all the forms of region energies in the literature. They are also used by Jermyn and Ishikawa (2001) as part of a 'ratio energy', and by Vasilevskiy and Siddiqi (2002) to find 'flux maximizing flows', while Kimmel and Bruckstein (2003) show the relation between certain instances of such energies and some common edge detectors.

Rather than define twisted linear functionals in general, we simply give the form appropriate to our context:

$$E(C) = \int_{\Box C} \star_{C^* g} \, C^* f = \int dp \, |\mathtt{t}(p)|_g \, f(C(p)) \; . \tag{2.5}$$

Here, f is a function (0-form) on Ω; $C^* g$ is the metric on $\Box C$ induced by C; and $|v|_g$ is the norm of the vector v in the metric g. The form of functional in equation (2.5) encompasses the remainder of the models mentioned above, including geometric and geodesic active contours, and most others that have appeared in the literature. It also can be given a physical interpretation, as the energy of a charged wire in an external electric potential f (again this potential is not limited by Maxwell's equations).

A particular example of this form of functional is boundary length in the metric g, which is given by setting $f = \mathbb{I}$.

In the particular case of prior terms, much more can be said. Prior terms should be Euclidean invariant in general. This forces f to be constant, g to be Euclidean, and A to calculate the interior area. Thus there are only two linear prior terms compatible with Euclidean invariance: length and area.

2.2 Higher-order functionals

The new, higher-order models that are the subject of this report make use of the linear structure of the chain space to go beyond linear functionals (we will use the word 'linear' to refer to both linear and twisted linear functionals) to polynomial functionals in a clear and structured way. This can be thought of as a coherent way of generating functionals of increasing complexity, or as the expansion of an arbitrary functional. The power of this approach can be seen in the fact that an arbitrary functional can express arbitrary information about the geometry of a region. The constructions we will give will be applied to the case of 1-chains in two dimensions, but as we have noted, the same constructions apply equally to the general case of p-chains in d dimensions, for example, 2-chains (surfaces) in three dimensions.

To construct polynomial functionals, it suffices to construct monomials. By definition, a monomial function of order n on a vector space V is the composition of three maps:

$$V \xrightarrow{\Delta_n} V^n \xrightarrow{\otimes} V^{\otimes n} \xrightarrow{E} \mathbb{R}$$

$$v \rightsquigarrow (v, v, \cdots) \rightsquigarrow v \otimes v \cdots \rightsquigarrow E(v \otimes v \cdots)$$

where: Δ_n is the diagonal map from V to its n-fold Cartesian product V^n; $\otimes$ is the projection from this latter space to the n-fold tensor product of V, $V^{\otimes n}$; and E is a linear functional on the latter. Note that setting $V = \mathbb{R}$ gives normal monomials, ax^n, $x \in \mathbb{R}$. In our context, $V = \mathcal{C}_1(\Omega)$, the space of 1-chains in Ω, and our task boils down to constructing linear functionals E on tensor products of $\mathcal{C}_1(\Omega)$ with itself. Fortunately, $\mathcal{C}_1(\Omega)^{\otimes n}$ is a subspace of $\mathcal{C}_n(\Omega^n)$, the space of n-chains in Ω^n, so that a linear functional on the latter is also a linear functional on the former. Linear functionals on the latter are easy to create however, as we now describe.

The only two ways to proceed for a generic n-chain are precisely analogous to the situation with linear functionals on 1-chains as shown in equations (2.2) and (2.5). In equation (2.2), a 1-form is pulled back to the domain of the 1-chain, and integrated. In equation (2.5), a function is pulled back to the domain of the chain, its Hodge dual is taken, and then it is integrated. Analogously, in order to construct linear functionals on n-chains we can:

1. (analogy of equation (2.2)) construct an n-form on Ω^n, pull it back to the domain of the n-chain and integrate it;
2. (analogy of equation (2.5)) construct a function on Ω^n, pull it back to the domain of the n-chain, take its Hodge dual (with respect to the induced metric, which is simply the n-fold product of the induced metric in the previous section), and integrate it.

However, in our case these are not the only possibilities because we are not dealing with generic n-chains. Instead, we are dealing with n-chains that are tensor products of 1-chains, whose domains are therefore products of the domains of these 1-chains, and whose induced metrics will be product metrics. This means that we can, for example, construct a p-form that only has components in p of the factors in Ω^n. When it is pulled back to the domain of the n-chain, it will be a p-chain on p of the factors in the domain of the n-chain, and a function on the other $n - p$ factors. Using the fact that the induced metric is a product, the Hodge dual can be taken on these $n - p$ factors alone, thus converting the p-form to a $p+(n-p) = n$-form, which can then be integrated. For a general p-form on Ω^n, the same can be achieved by splitting its pullback into components. We will nevertheless concentrate on the two simplest forms enumerated above.

In the first case, we are given an n-form F on Ω^n. We pull it back to the domain of C^n and integrate it, using the analogue of equation (2.2):

$$E(C) = \int_{(\partial R)^n} F = \int_{\square C^n} (C^n)^* F \; . \tag{2.6}$$

In the second case, we are given a 0-form (function) f on Ω^n. We pull it back to the domain of C^n, take its Hodge dual, and integrate it:

$$E(C) = \int_{\square C^n} \star_{(C^n)^*(g^n)} (C^n)^* f \; . \tag{2.7}$$

where, note, $(C^n)^*(g^n) = (C^* g)^n$.

The coordinate expressions for these functionals are complex. In what follows, we focus on the quadratic case, $n = 2$, both for clarity and because this is what we will use in the application later in the report.

2.3 Quadratic functionals

2.3.1 Equation 2.6

In the case $n = 2$, the first form of functional, equation (2.6), becomes

$$E(C) = \int_{(\partial R)^2} F = \int_{\square C^2} (C^2)^* F \; . \tag{2.8}$$

The product structures of C^2 and $\square C^2$ mean that this form of functional can always be written (in terms of coordinates (p, p') on $\square C^2$) as

$$E(C) = \int\int dp\, dp'\; \mathbf{t}(p) \cdot F(C(p), C(p')) \cdot \mathbf{t}(p') \; , \tag{2.9}$$

where $F(x, x')$, for each $(x, x') \in \Omega^2$, is a matrix (actually a bitensor with one index at x and one at x').

For prior terms, when the 2-form F does not depend on the image, we require the energy to be Euclidean invariant. This results in the form given in equation 1.1, which we reproduce here:

$$E(C) = \int\int dp\, dp'\; \mathbf{t}(p) \cdot \mathbf{t}(p')\; \Psi(|C(p) - C(p')|) \; . \tag{2.10}$$

2.3.2 Equation 2.7

In the case $n = 2$, the second form of functional, equation (2.7) becomes

$$E(C) = \int_{\square C^2} \star_{(C^2)^*(g^2)}\; (C^2)^* f \; . \tag{2.11}$$

The product structures of C^2 and $\square C^2$ mean that this form of functional can always be written (in terms of coordinates (p, p') on $\square C^2$) as

$$E(C) = \int\int dp\, dp'\; |\mathbf{t}(p)|_g\; |\mathbf{t}(p')|_g\; f(C(p), C(p')) \quad = \int\int ds\, ds'\; f(C(s), C(s')) \; , \tag{2.12}$$

where s is the arc length parameter.

For prior terms, the function $f(x, x')$ can only depend on the distance between x and x':

$$E(C) = \int\int ds\, ds'\; f(|C(s) - C(s')|) \; . \tag{2.13}$$

2.3.3 Other forms

In the $n = 2$ case, the other forms of linear functional discussed in the previous subsection reduce to one. In coordinates (p, p') on $\square C^2$, this form of functional is given by

$$E(C) = \int\int dp\, dp'\; |\mathbf{t}(p)|_g\; \mathbf{t}(p') \cdot v(C(p), C(p')) \; , \tag{2.14}$$

where $v(x, x')$, for each $(x, x') \in \Omega^2$, is a tangent vector at x'.

2.3.4 Discussion

The operator F, the function f, and the vector field v allow us to model non-trivial interactions between different contour points, which in turn allows the incorporation of non-trivial prior knowledge about region geometry, and the relation between region geometry and the data.

As with linear functionals on 1-forms, the above functionals can be interpreted in physical terms. The form in equation (2.10) can be thought of as the energy of a current-carrying wire in its induced magnetic field, or equivalently as the (self-)interaction between current-carrying wires via the magnetic field. Equation (2.13) can be though of as the energy of a charged wire in its induced electric field, or equivalently as the (self-)interaction between charged wires via the electric field. However, in neither case is the kernel of the interaction limited to that occurring in electromagnetism.

Note that the interactions described by the functionals are not Markov, even if the values of F, f, and v tend to zero rapidly with increasing distance between their arguments. Since the interaction is mediated by the embedding rather than the embedded space, interactions can occur between arbitrarily separated pieces of the contour if they approach each other in Ω.

Note also that unlike the shape models described in section 1.1.2, the new energies incorporate Euclidean invariance naturally without requiring the estimation of position or rotation, since they are not mixture models over these variables. Note also, however, that this does not constrain the minimum energy configurations to be Euclidean invariant, although the set of such minima will be; the symmetry is 'broken' in general.

2.4 An example of a geometric quadratic energy

In this section, we study a particular case of an Euclidean invariant quadratic energy. We will use this particular case later on to model road networks, but we use it here to illustrate the possibilities inherent in higher-order energies.

The energy is a combination of two linear terms (length and area) to which are added a quadratic term characteristic of the new class of energies. It takes the form

$$E_g(C) = \mathcal{L}(C) + \alpha \mathcal{A}(C) - \beta \int\int dp\, dp'\, \mathbf{t}(p) \cdot \mathbf{t}(p')\, \Psi(R(p,p'))\ , \tag{2.15}$$

where $\mathcal{L}$ is the length of the boundary in the Euclidean metric on Ω, an energy of the form (2.5); and $\mathcal{A}$ is the area of its interior, an energy of the form (2.2). $R(p,p') = |C(p) - C(p')|$ is the Euclidean distance between $C(p)$ and $C(p')$. The length term acts as a regularizer. The area term is introduced to control the expansion of the region. The Euclidean invariant quadratic term, of the form (2.10), introduces the interactions. We choose the following form for the function Ψ:

$$\Psi(x) = \begin{cases} 1 \text{ if } x < d_{\min} - \epsilon\ , \\ 0 \text{ if } x > d_{\min} + \epsilon\ , \\ \frac{1}{2}(1 - \frac{x - d_{\min}}{\epsilon} - \frac{1}{\pi}\sin(\pi \frac{x - d_{\min}}{\epsilon}))\ \text{ else}\ . \end{cases} \tag{2.16}$$

This function is shown in figure 1, where the parameters $d_{\min}$ and ϵ are also illustrated. A point p on the contour interacts with other points within a certain distance $d_{\min} + \epsilon$, measured in Ω. The

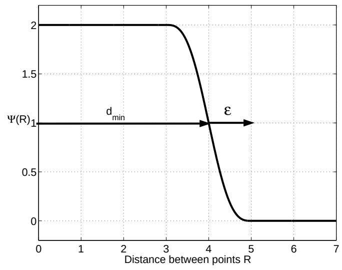

Figure 1: The function Ψ

function Ψ is always positive, and so from equation (2.15), the quadratic part of the energy is a minimum when the points interacting with one another have parallel tangent vectors. The quadratic energy thus favours straight boundaries. On the other hand, for pairs of points with antiparallel tangent vectors, the quadratic part of the energy is zero unless the points approach closer than a distance of $d_{\min} + \epsilon$, when it starts to increase rapidly. The quadratic energy therefore acts as a softened 'hard-core' potential, preventing the points from approaching much closer than $d_{\min}$ by making such points mutually repelling.

The energy in equation (2.15) is minimized using gradient descent. Thus the contour evolution is determined by

$$\partial_t C(p) = -\frac{\delta E}{\delta C(p)} , \tag{2.17}$$

where $\delta E/\delta C$ is the functional derivative of E with respect to C. [3] The resulting descent equation is then

$$\hat{\mathbf{n}} \cdot \partial_t C(p) = -\kappa(p) - \alpha + 2\beta \int dp' \, \hat{\mathbf{R}}(p, p') \cdot \mathbf{n}(p') \, \dot{\Psi}(R(p, p')) , \tag{2.18}$$

where $\dot{}$ indicates derivative, $\hat{\mathbf{R}}(p, p') = (C(p) - C(p'))/|C(p) - C(p')|$, and $\mathbf{n}$ ($\hat{\mathbf{n}}$) is the (normalized) outward pointing normal vector. The component of $\partial_t C$ along the normal has been taken, movement along the tangent direction being equivalent to a diffeomorphism of the domain of C, and thus irrelevant.

The precise behaviour of this energy depends on the parameter values, and in particular on the size of the parameter β describing the strength of the quadratic term. By making this parameter large, we can exaggerate the effect of the new term in order to make clear the information contained

[3] Note that strictly speaking, $\delta E/\delta C$ are the components of a 1-form on the space of boundaries, and that to generate a gradient we should map it to the tangent bundle using a metric. By using it as is, we are effectively assuming that the metric on the space of boundaries is Euclidean in the point basis; this is common practice. The choice of a metric is difficult; there are good arguments for saying that it should be determined not *a priori*, but by the measurements we intend to make, that is, by the likelihood in Bayes' theorem (Jermyn, 2005).

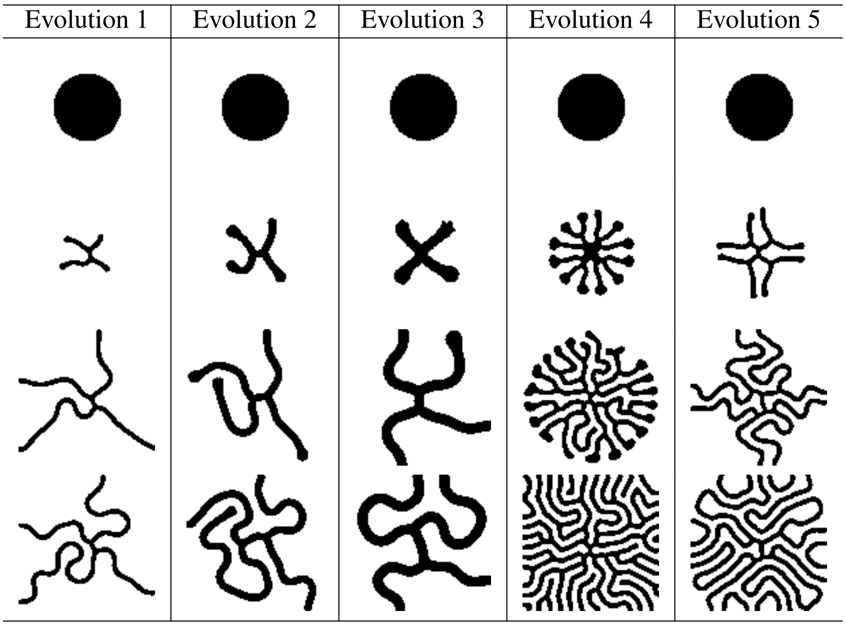

Figure 2: Examples of gradient descent using the energy in equation (2.15). The first three columns correspond to different values of $d_{\min}$, while the last two correspond to different values of α.

in it. Figure 2 shows examples of evolutions starting from a circle using equation (2.18). All the evolutions show the formation of fingered structures with parallel-sided arms of constant width. The width is controlled by the parameter $d_{\min}$ in the Ψ function, and the first three rows of figure 2 show evolutions for different values of this parameter ($d_{\min} = 3, 5, 7$); the fingers formed are indeed of the correct width. The last two rows illustrate the role of the parameter α. In the fourth row, $\alpha = 0.05$, while $\alpha = 0.1$ in the fifth row. The larger the value of α, the fewer the number of arms that form at the beginning of the evolution.

The growth away from a circle towards a labyrinthine structure with elongated 'arms' can be understood in two stages. A linear analysis of the stability of the circle to small sinusoidal perturbations shows that for β larger than a certain threshold, the circle becomes a saddle-point of the energy (2.15). For certain ranges of angular frequency, small perturbations, rather than being damped back to zero as in the linear case, are amplified. The maximally unstable angular frequency controls the motion away from the initial circle. Thus instead of smoothing all irregularities, as in the linear case, this energy allows some of them to develop, and hence encourages complex shapes. It is important to note that high frequencies are still damped, and thus there is no development of uncontrolled noise on the contour, as would be the case with evolution according to a negative length term, for example. Intuitively this can be seen by realizing that the 'bumps' corresponding to two peaks in the sinusoid cannot approach closer than $d_{\min}$ due to their mutual repulsion.

The second stage occurs when fledgling arms have developed. With the value of β used in these experiments, each unit length of arm adds a negative amount of energy to the total. The arms, once created, thus elongate. The parallel sides of the arms are stable to small perturbations, but their tips, which are roughly semicircular, possess the same kinds of instability as the original circle, and can thus branch, with a branching number again controlled by β. Since the arms possess a negative energy, in an infinite domain the energy is not bounded below, and the arms will continue to grow and to ramify indefinitely. In a finite domain such as an image, this cannot happen due to the repulsion between the arms, and stable configurations are eventually reached.

In the absence of data, and with the parameter values used in figure 2, it is deviations from exact circularity caused by the discretization drive the evolution away from the initial conditions, and it is clear that small changes in the initial conditions (perhaps caused by the discretization) will result in convergence to possibly very different shapes, although these shapes will have many qualitative and quantitative properties in common. The latter is a consequence of the large amount of symmetry built into the problem. The energy is invariant under rotations and translations, but it is clear that there is also a high degree of local symmetry: one can contort the arms in many ways and leave the energy unchanged. This is as it should be: road networks (or any other type of network) have a large number of configurations that while differing in their detail, are equally reasonable as networks *a priori*. The prior energy quite properly does not distinguish amongst these possibilities, and with these parameter values, small perturbations effectively choose amongst them.

While the experiments serve to illustrate the greater complexity of information contained within quadratic energies as compared to linear energies, and to show that the specific energy in equation (2.15) is well-suited to modelling network structures, it is very important to note that the value of β used in these experiments is not the same as that used in the extraction of road networks. In the experimental results that we show later, β is adjusted so that each unit length of an arm adds a small but positive amount to the energy, and so that the circle is marginally stable. The result is that the data drives the production, growth, and branching of arms, the effect of the prior term being to favour such arms, and network shapes in general, with respect to other shapes by reducing the energy of network configurations.

However, even with the parameter values set similarly to those used to generate figure 2, in the presence of data, the situation changes dramatically. The data now defines 'preferred directions' with a strength that far outweighs any perturbations produced by the discretization or otherwise. Indeed perturbations that do not fit the data will be rapidly damped (although of course this is dependent on the relative magnitudes of the likelihood and prior terms). In fact, the presence of the higher-order terms now has the opposite effect: far from increasing the sensitivity of the minimum found to the initial conditions, the higher-order terms reduce it, precisely because the incorporation of more sophisticated prior knowledge eliminates many local minima.

3 Minimization of the energy

Although the linear space of chains is useful for constructing and describing functionals, the minimization of the energy takes place not over the space of (closed) 1-chains, but over the space of region boundaries. In order to minimize the energy, we use gradient descent, evolving the contour using the level set framework introduced by Osher and Sethian (1988). This framework, and its advantages for contour evolution are by now well known. Here we just note that if the contour propagates along the outward normal direction with speed F, *i.e.* $\hat{\mathbf{n}} \cdot \partial_t C(p) = F[C](p)$, then the level set function on the contour must obey

$$\partial_t \phi = -\nabla\phi \cdot F\hat{\mathbf{n}} = F\nabla\phi \cdot \nabla\phi/|\nabla\phi| = F|\nabla\phi| \ . \tag{3.1}$$

In principle, we would like ϕ to evolve as a signed distance function, to which it is usually initialized, but this is hard to guarantee in general. However, since the exact evolution of ϕ off the contour is of no consequence provided it is well enough behaved, it can be chosen for convenience, reinitializing every so often if necessary to restore the signed distance function. A typical choice is to apply the expression for F to each level set, and evolve the function ϕ accordingly.

As can be seen from equation (2.18), the evolution equations derived from quadratic energies contain nonlocal terms, and this creates new difficulties. Following the procedure of applying the expression for F to every level set is impractical, since it means extracting the level set belonging to each point of the discretized version of Ω and integrating over it. In order to construct the speed at all points of Ω from the speed on the contour, we therefore use the technique of 'extension velocities' (Adalsteinsson and Sethian, 1999). The level set function is thus evolved in four steps. First ϕ is (re)initialized (section 3.1), then the zero level set is extracted and the speed on the contour computed (section 3.2). The speed is then extended from the zero level set to Ω (section 3.3), and finally ϕ is updated (section 3.4). Figure 3 depicts the four steps. In the next few subsections, we describe these steps in more detail.

3.1 (Re)initialization

In order to (re)initialize ϕ as a signed distance function, we use the approach described by Sussman et al. (1994), where the PDE

$$\partial_\tau \phi = \text{sign}(\phi_0)\,(1 - |\nabla\phi|) \text{ ,with} \qquad \phi(x,0) = \phi_0(x) \tag{3.2}$$

is solved for this purpose. We found, however, that the zero level set moved during the numerical solution of this equation, an effect which manifested itself as a loss of area when we attempted to simulate an area-preserving flow for example. This is a recognized problem, to which Sussman and Fatemi (1997) have proposed a solution. A local area conservation constraint is imposed by modifying equation (3.2) in each cell Ω_{ij} of Ω to

$$\partial_\tau \phi = \text{sign}(\phi_0)\,(1 - |\nabla\phi|) + \lambda_{ij}\dot{H}(\phi)|\nabla\phi| \ , \tag{3.3}$$

where

$$\lambda_{ij} = \frac{-\int_{\Omega_{ij}} \dot{H}(\phi)\text{sign}(\phi_0)\,(1 - |\nabla\phi|)}{\int_{\Omega_{ij}} \dot{H}(\phi)^2|\nabla\phi|} \ . \tag{3.4}$$

The initial condition for equation (3.3) is the current value of ϕ, except for the initialization of the evolution, when ϕ is set to $+1$ inside the contour and -1 outside.

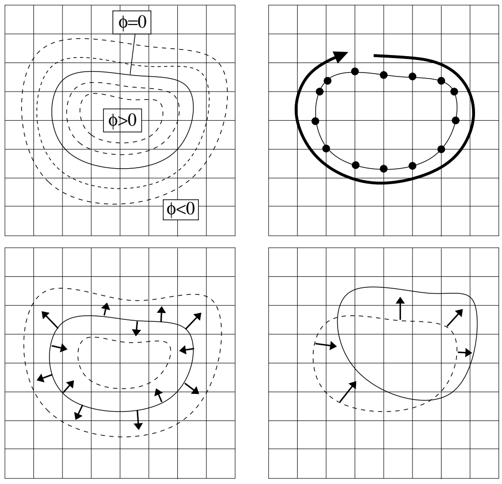

Figure 3: The four evolution steps. Top-left, step (1): (re)initialization; top-right, step (2): contour extraction and computation of the speed on-contour; bottom-left, step (3): extension of the speed to the Narrow Band; bottom-right, step (4): evolution of ϕ.

3.2 Contour extraction and computation of F on the contour

In order to compute accurately the speed F on the zero level set, we first locate the intersections of this set with the grid using ENO interpolation (Siddiqi et al., 1997). After interpolation, the boundary is extracted using the contour tracing algorithm shown in table 1 (Pavlidis, 1982). At each step, we start from the current point and consider six possible directions for the next point. These directions are adapted to the different possible configurations, as shown in figure 4. We obtain an ordered set of points $\{C(p_i); i = 1, \ldots, n\}$ representing the boundary.

In fact, the situation is more complicated than this, because some configurations are ambiguous, as shown in figure 4. To deal with these, it is necessary to adopt a convention: either the interior or the exterior, but not both, can have subcellular width. We choose the former.

Having extracted the boundary, and after interpolating the necessary values from the grid, we compute the speed F for each extracted point by performing a numerical integration over the contour.

3.3 Computation of F on all points of the domain

The speed is needed for all points of Ω. As mentioned at the beginning of this section, in order to do this efficiently, we use the method of 'extension velocities', as proposed by Adalsteinsson and Sethian (1999). To initialize the process, the grid points closest to the extracted boundary inherit the

Table 1: Tracing algorithm.

1. Choose a starting point A in a set of points R. Set current point $C = A$ and search direction $S = 6$.
2. While C is different from A or *first*$= 1$, do steps 3 to 9.
3. *found*$= 0$.
4. While *found*$= 0$, do steps 5 to 8, at most 3 times.
5. If B, the neighbour $(S-1)$ of C is in R; $C = B$, $S = S - 2$, *found*$= 1$.
6. Else, if B, the neighbour S of C is in R, $C = B$ and *found*$= 1$.
7. Else, if B, the neighbour $(S+1)$ of C is in R, $C = B$ and *found*$= 1$.
8. Else $S = S + 2$.
9. *first*$= 0$.

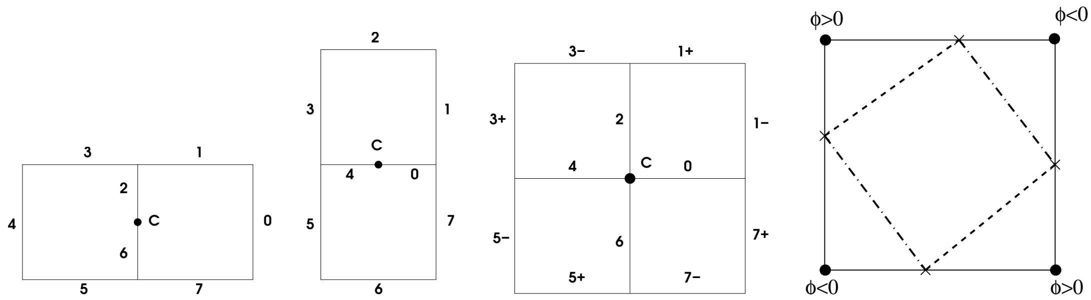

Figure 4: Leftmost three: configurations encountered in the contour tracing algorithm. Right: an ambiguous configuration.

speed of the closest extracted boundary point. We then solve the PDE

$$\partial_\tau F + \text{sign}(\phi) \frac{\nabla \phi}{|\nabla \phi|} \cdot \nabla F = 0 \ .$$

with initial condition $F = 0$ except at these points. Note that on the contour, when $\phi = 0$, the force F does not change, thus preserving the boundary condition. For the implementation, this translates into conservation of the values of F at these nearby grid points. At convergence, when $F_\tau = 0$, the solution satisfies $\nabla \phi \cdot \nabla F = 0$, which means that F is constant along the normals to the level sets. Every level set then evolves with the same speed, meaning that the distance between each level set is preserved in principle, thus preventing ϕ from being badly behaved.

3.4 Evolution of ϕ

In practice, it is not necessary to compute the evolution of the level set function over the whole of Ω. Computational efficiency can be increased by restricting computation to a band around the zero level set, known as the 'Narrow Band' (Sethian, 1999), defined by $|\phi(x)| < t$, where t is a threshold. When the zero level set comes too close to the edge of the Narrow Band, the level set function is reinitialized as described above, and the Narrow Band is reconstructed.

4 Application: line network extraction

The extraction of line networks, and especially road networks, from remote sensing imagery has been studied for the last twenty years at least, and a wide variety of methods have been developed to attack this problem. Despite all this attention, the automatic extraction of line networks remains a challenge because of the great variability of the objects concerned, and the consequent difficulty in their characterization. The intensity of a road can vary significantly from one road to another, for example, while the presence of trees and buildings in high resolution data can obscure the network; junctions can be highly complex; networks do not possess exactly the same properties in rural and urban areas; and so on.

Previous work can be characterized in a number of ways. Some methods extract the network as a one-dimensional object, whereas others extract the network as a region. Some restrict the network topologies that can be found, generally to linear structures with no junctions. Some are semi-automatic, and require information about the road location in the image as initialization (end-points, or initialization very close to the road). Others aim at being fully automatic, although to our knowledge there is no fully automatic method: parameters at least always need to be set.

Methods that restrict the topology and find 1D structures include those that find an optimal path between two endpoints. Fischler et al. (1981), for example, combine the results of applying several specially designed operators into an array of costs inversely related to the likelihood of the presence of a road, and then find an optimal path through this array. Merlet and Zerubia (1996) define a path cost depending on the contrast, grey-level and curvature along a path between two endpoints, and then minimize the cost using dynamic programming. Geman and Jedynak (1996) propose a tree search method for road tracking (i.e. only a start point need be given) based on reducing as much as possible the uncertainty in the road position.

Methods that do not restrict the topology, but find 1D structures include MRF and marked point process models. Tupin et al. (1998) first generate a number of candidate line segments using two different line detectors. The segments are then connected together using a Markov random field defined on a graph with vertices the segments, thus allowing complex topologies. Stoica et al. (2004) and Lacoste et al. (2002) model thin networks, including roads, as ensembles of line segments embedded in the image domain. Marked point processes (with line segments as marks) control network parameters such as connectivity and curvature via interactions between the segments.

All these methods find a connected set of points or segments, but do not extract the borders of the road (although Stoica et al. (2004) consider a model that includes segment width as a variable). With increase of image resolution, the width of networks can become significant, and it then makes more sense to consider the network as a region. Barzohar and Cooper (1996) propose an automatic approach that first finds MAP estimates of the road configuration in small windows using dynamic programming, and then combines these window estimates, again using dynamic programming. The model used explicitly includes the road borders.

Active contour methods also find regions, but all previous applications of active contours to road network extraction find only linear structures, and require initialization very close to the road to be found. Neuenschwander et al. (1997) introduce 'ziplock snakes'. From an initial and a final point, forces derived from the image are progressively used to adjust the position of the active contour. The endpoints are positioned on either side of the road, and both borders of the road are extracted. Fua and Leclerc (1990) and Laptev et al. (2000) model roads using 'ribbon snakes', active contours with a certain width associated to each point.

In contrast, the higher-order active contour models described here find regions, with no restriction on the topology, and can be initialized in a generic way without reference to the true road position.

4.1 Proposed model

The model has to take into account two fundamental aspects of the entity to be detected: the geometry and the radiometry, corresponding to prior and likelihood terms. The energy thus contains two parts:

$$E(C) = E_g(C) + \lambda E_i(C) \;, \tag{4.1}$$

where λ balances the contributions of the geometric part E_g and the data part E_i. The geometric part E_g is given by equation (2.15), and is described in section 2.4. The image part E_i is composed of two terms:

$$\begin{aligned} E_i(C) &= \int_{\partial R} \star dI - \int_{(\partial R)^2} (\Psi \circ R)(dI \star dI') \\ &= \int_{\Box C} C^*(\star dI) - \int_{\Box C^2} (C \times C)^* \left[(\Psi \circ R)(dI \star dI')\right] \\ &= \int dp\, \hat{\mathbf{n}} \cdot \nabla I - \int\!\!\int dp\, dp'\, \mathbf{t} \cdot \mathbf{t}'\, (\nabla I \cdot \nabla I')\, \Psi(R(p,p')) \;, \end{aligned} \tag{4.2}$$

where we use primed and unprimed variables to designate quantities evaluated at points p (or $C(p)$) and p' (or $C(p')$) respectively. The first, linear term has the form (2.2), while the quadratic term takes the general form (2.8).

The linear term favours situations in which the outward normal is opposed to the image gradient, or in other words, in which the road is lighter than its environment. When this is the case, it also favours larger gradients under the contour. The second term is an example of a quadratic likelihood term: it describes a relation between the contour and the data that cannot be incorporated into a linear functional. Its effect is to favour the two situations illustrated in figure 5. First, it favours configurations in which pairs of points whose tangent vectors are parallel and that are not too distant

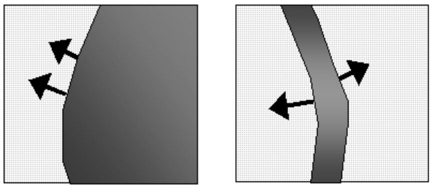

Figure 5: The two configurations favoured by the quadratic image term.

from each other (*i.e.* points on the same side of a road) lie on image gradients that point in the same direction and are large. Second, it favours configurations in which pairs of points whose tangent vectors are antiparallel (*i.e.* points on opposite sides of a road) lie on image gradients that point in opposite directions and are large. This latter is important, as it allows the model to capture the *joint* behaviour exhibited by the opposing sides of a road.

The energy in equation (4.1) is minimized using gradient descent implemented via level sets as described in section 3. The resulting descent equation is

$$\hat{\mathbf{n}} \cdot \partial_t C = -\kappa - \alpha - \lambda \nabla^2 I + 2\lambda \int dp' \, (\nabla I' \cdot \nabla\nabla I \cdot \mathbf{n}') \, \Psi(R(p,p')) \\ + 2 \int dp' \, (\hat{\mathbf{R}} \cdot \mathbf{n}')(\beta + \lambda \nabla I \cdot \nabla I') \, \dot{\Psi}(R(p,p')) \; . \quad (4.3)$$

4.2 Parameters and initialization

The above models have parameters, but as with all variational methods, there exist no good ways of assigning values to most of these parameters. In the experiments below, the parameters were all set empirically to optimize the results, with the exception of $d_{\min}$ and ϵ, which have clear physical meanings and can be set using the resolution of the image and an examination of the roads it contains.

Initialization is an important issue for gradient descent methods. The results may depend heavily on the initialization chosen, and indeed a number of the methods used for the detection of roads rely on an initialization very close to the network. In all the results shown below, however, the region used to initialize the gradient descent was a rounded rectangle lying just inside Ω. This is possible because the greater specificity of the model eliminates many candidate contours from consideration, thus removing many local minima. All experiments were run until convergence.

4.3 Experimental results with the first model

We tested the above model on real satellite and aerial images. Two such images are shown in the first column of figure 6. The images present several difficulties. There are regions of high gradient corresponding to the borders of fields rather than to roads, and fields also exhibit parallel sides. In the first image, there is a discontinuity in the road. The gradient descent procedure and results are shown in the second to fifth columns of figure 6. In both images, the roads are perfectly extracted.

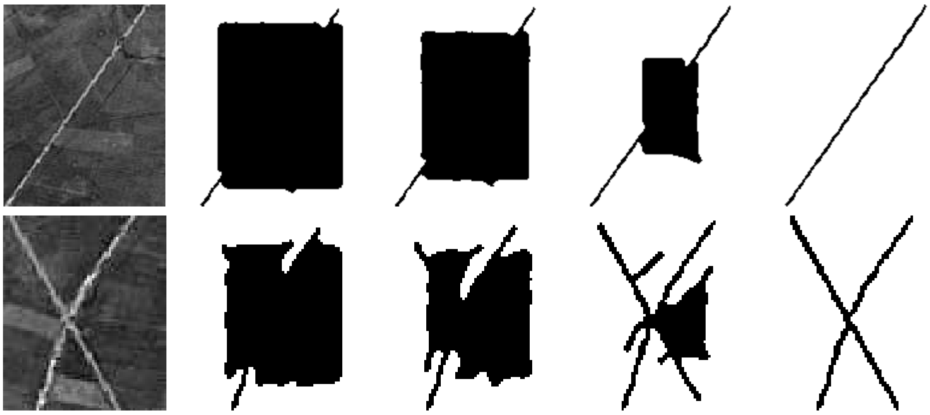

Figure 6: Gradient descent on the two SPOT satellite images in the first column.

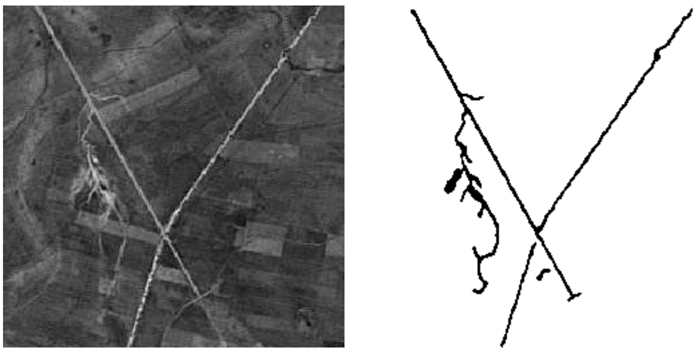

Figure 7: Result on a larger piece of the SPOT image.

Figure 7 shows another result on a larger, more complex piece of the same satellite image. The result is not perfect but very encouraging. We are able to detect both straight and 'windy' portions of the network, and areas where the road width varies.

The likelihood term, although it takes into account some aspects of the appearance of road networks in images, can nevertheless be improved. For instance, isolated edges are occasionally detected. In the next section, we add another image term to our model, more specific to the radiometry of a line network.

4.4 A more specific image term

Consider a function G on Ω that is representative of the entity to detect, in our case the line network. For instance, it could be the log probability that each point of Ω belongs to the network. Then one can define the following energy of the form (2.2) and (2.3):

$$E(C) = -\int_R \star G = -\int_{\partial R} A_G = -\int_{\square C} C^* A_G \;, \tag{4.4}$$

where $dA_G = \star G$. The functional derivative is given by

$$\frac{\delta E}{\delta C(p)} = G(C(p))\hat{\mathbf{n}}(p) \;. \tag{4.5}$$

In the following subsections, we describe two ways of constructing G. The first method uses oriented filtering, while the second uses hypothesis tests.

4.4.1 Oriented filtering

Define the function

$$\mathcal{F}_\theta = (\hat{\mathbf{v}}_\theta \cdot \nabla)^2 N_\sigma \;,$$

where N_σ is a rotationally symmetric Gaussian with standard deviation σ, and $\hat{\mathbf{v}}_\theta$ is the unit vector in direction θ. Then G is given by

$$G(x) = Q(\min_{\theta\in\Theta}(\mathcal{F}_\theta * I(x))) \;,$$

where $*$ indicates convolution. The rotations are chosen from the set $\Theta = \left\{0, \frac{\pi}{8}, \ldots, \frac{7\pi}{8}\right\}$. The function Q maps the values into the interval $[-1, 1]$:

$$Q(x) = \begin{cases} 1 & \text{if } x < s_1 \;, \\ 1 - 2\frac{x-s_1}{s_2-s_1} & \text{if } s_1 \le x \le s_2 \;, \\ -1 & \text{if } x > s_2 \;, \end{cases} \tag{4.6}$$

where s_1 and s_2 are two thresholds, chosen empirically.

4.4.2 Hypothesis tests

Lacoste et al. (2002) used Student t-tests for line network detection. Here we adapt their approach to our context. We suppose that roads are homogeneous and contrasted with respect to their environment. A t-test on sets of pixels from inside a potential road will test the homogeneity criterion, while a t-test on sets of pixels from inside and outside a potential road will test the contrast criterion. In order to compute the test, we use the mask shown in figure 8.

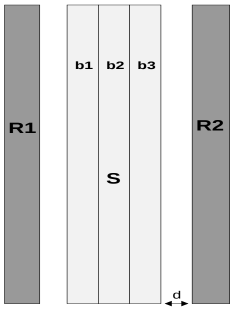

Figure 8: Mask for Student tests.

The Student t-test computes

$$\text{t-test}(x,y) = \frac{|\bar{x} - \bar{y}|}{\sqrt{\frac{\sigma_x}{n_x} + \frac{\sigma_y}{n_y}}} ,$$

where $\bar{\cdot}$, σ and n represent respectively the mean, the standard deviation, and the number of observations. When the result of the test is above a certain threshold, we can consider that the two sets of pixels belong to different populations (implicitly, Gaussian with different means and variances). Given a mask location and orientation, (x, θ), we test the homogeneity and contrast criteria by computing the quantity

$$T_\theta(x) = Q\left(\frac{H_2}{\min\{1, H_1\}}\right) .$$

where

$$H_1 = \max_{j,k \in \{1,\ldots,n_b\}, j \neq k} [\text{t-test}(b_j, b_k)] \text{and} \qquad H_2 = \min_{l \in \{1,2\}} [\text{t-test}(R_l, S)] .$$

The function G is then defined by

$$\theta_{\max}(x) = \arg\max_{\theta \in \Theta} |T_\theta(x)| \text{and} \qquad G(x) = T_{\theta_{\max}(x)}(x) .$$

Both functions act as simple linear structure detectors, picking up elongated structures for which the interior has an average intensity different to that of the exterior neighbourhood, but the second is more subtle. The first essentially calculates differences of means, while the second compares these differences to the data variances.

4.5 Experimental results with the second model

We add this new energy (4.4) to the model (4.1), and test the model on the high-resolution aerial image shown in figure 9. The image presents several difficulties because of high gradients that do

Figure 9: Aerial image. (©IGN.)

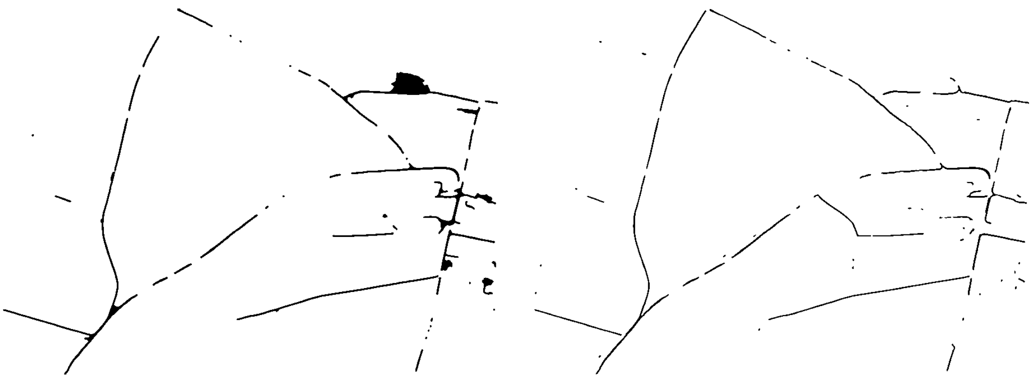

Figure 10: Results of extraction with the two functions G.

not correspond to sides of roads and because of occlusions due to the presence of trees next to the road network. We obtain two extraction results corresponding to the two functions G above. The results are similar, and are shown in figure 10.

The main part of the network is extracted, and field borders and other geometric noise are eliminated. In the top-right in one result, a road encircling a house is extracted as a solid area. This happens because 'holes' cannot form in the centre of a region with the current formulation. The main problem, however, is that occlusions due to trees disrupt the network. We are currently addressing this issue using a quadratic energy that causes two road 'tips' to attract one another, and thus close such gaps.

5 Conclusions

We have introduced a new class of active contour energy functionals that greatly generalizes the energies used in previous work. Previous energies are linear on the space of 1-chains, being expressed as single integrals over the contour. The new energies are arbitrary polynomials on the space of 1-chains, adding multiple integrals to the linear terms. The new energies describe arbitrarily long-range interactions between sets of contour points, and thus can incorporate both sophisticated prior geometric information and complex multi-point interactions between the contour and the data. The prior terms can easily be made Euclidean invariant, thus obviating the need for pose estimation usual for active contour shape models.

We studied a particular form of quadratic energy that describe 'networks': structures composed of 'arms' of roughly parallel sides, perhaps of varying width, joined together in various ways. Using this energy as a base, we designed an energy functional including both quadratic prior and likelihood terms and applied it to the extraction of road networks from satellite and aerial imagery. Simulations prove the efficacy of the model and illustrate the effect of the incorporation of non-trivial geometrical interactions between points of the contour and between the contour and the data. The enhanced specificity of the prior eliminates many local minima, thus enabling an automatic initialization step. Algorithmically, these models presented new challenges also, in particular the need for a maximum of precision in the calculation of the speed and the evolution of the contour.

Immediate future work is focused on the solution of the problems mentioned in connection with figure 9, where occlusions disrupted the network. We have designed a quadratic 'gap closure' force that overcomes the repulsion introduced by the existing quadratic term in certain circumstances, leading road 'tips' to attract one another and fill in gaps in the network, something that is impossible using classical techniques. Incorporating such a force into an energy framework is challenging, as it involves higher-order derivatives that create numerical difficulties. We are currently working on resolving these.

Many open questions and research directions remain to be explored, the most important being the construction of a functional for a given family of shapes. Others include: higher-than-quadratic functionals; the extension to surfaces; a probabilistic formulation; improving computational efficiency; and applications to other domains, in particular to medical imagery, where it is clear that the geometry is often similar.

References

D. Adalsteinsson and J. A. Sethian. The fast construction of extension velocities in level set methods. *J. Comp. Phys.*, 148:2–22, 1999.

M. Barzohar and D. Cooper. Automatic finding of main roads in aerial images by using geometric-stochastic models and estimation. *IEEE Trans. Patt. Anal. Mach. Intell.*, 18(7):707–721, 1996.

A. Bossavit. Applied differential geometry: A compendium, 2002. http://www.icm.edu.pl/edukacja/mat/Compendium.php.

V. Caselles, F. Catte, T. Coll, and F. Dibos. A geometric model for active contours. *Numerische Mathematik*, 66:1–31, 1993.

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

T. F. Chan and L. A. Vese. Active contours without edges. *IEEE Trans. Im. Proc.*, 10-2:266–277, 2001.

Y. Chen, S. Thiruvenkadam, H. D. Tagare, F. Huang, D. Wilson, and E.A. Geiser. On the incorporation of shape priors into geometric active contours. *Proc. IEEE Workshop VLSM*, pages 145–152, 2001.

L. D. Cohen. On active contours and balloons. *CVGIP: Image Understanding*, 53:211–218, 1991.

D. Cremers, C. Schnorr, and J. Weickert. Diffusion-snakes: combining statistical shape knowledge and image information in a variational framework. *Proc. IEEE Workshop VLSM*, pages 137–144, 2001.

M. A. Fischler, J. M. Tenenbaum, and H. C. Wolf. Detection of roads and linear structures in low-resolution aerial imagery using a multisource knowledge integration technique. *Comp. Graph. and Im. Proc.*, 15:201–223, 1981.

A. Foulonneau, P. Charbonnier, and F. Heitz. Geometric shape priors for region-based active contours. *Proc. IEEE ICIP.*, 3:413– 416, 2003.

P. Fua and Y. G. Leclerc. Model driven edge detection. *Mach. Vis. and Appl.*, 3:45–56, 1990.

D. Geman and B. Jedynak. An active testing model for tracking roads in satellite images. *IEEE Trans. Patt. Anal. Mach. Intell.*, 18:1–14, 1996.

S. Jehan-Besson, M. Barlaud, and G. Aubert. DREAM2S: Deformable regions driven by an Eulerian accurate minimization method for image and video segmentation. *Int'l J. Comp. Vis.*, 53:45–70, 2003.

I. H. Jermyn. Invariant bayesian estimation on manifolds. *Ann. Stat.*, 33(2):583–605, April 2005. URL `http://dx.doi.org/10.1214/009053604000001273`.

I. H. Jermyn and H. Ishikawa. Globally optimal regions and boundaries as minimum ratio cycles. *IEEE Trans. Patt. Anal. Mach. Intell. (Special Section on Graph Algorithms and Computer Vision)*, 23(10):1075–1088, October 2001.

M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *Int'l J. Comp. Vis.*, pages 321–331, 1988.

S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi. Gradient flows and geometric active contour models. *Proc. ICCV*, pages 810–815, 1995.

R. Kimmel and A. M. Bruckstein. On regularized laplacian zero crossings and other optimal edge integrators. *Int'l J. Comp. Vis.*, 53(3):225–243, 2003.

C. Lacoste, X. Descombes, and J. Zerubia. A comparative study of point processes for line network extraction in remote sensing. Research Report 4516, INRIA, France, August 2002.

I. Laptev, T. Lindeberg, W. Eckstein, C. Steger, and A. Baumgartner. Automatic extraction of roads from aerial images based on scale space and snakes. *Mach. Vis. and Appl.*, 12:23–31, 2000.

M. E. Leventon, W. E. L. Grimson, and O. Faugeras. Statistical shape influence in geodesic active contours. *Proc. IEEE CVPR*, 1:316–322, 2000.

R. Malladi, J. A. Sethian, and B. C. Vemuri. Shape modeling with front propagation: A level set approach. *IEEE Trans. Patt. Anal. Mach. Intell.*, 17:158–175, 1995.

N. Merlet and J. Zerubia. New prospects in line detection by dynamic programming. *IEEE Trans. Patt. Anal. Mach. Intell.*, 18(4):426–431, 1996.

W. M. Neuenschwander, P. Fua, L. Iverson, G. Székely, and O. Kubler. Ziplock snakes. *Int'l J. Comp. Vis.*, 25(3):191–201, 1997.

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

N. Paragios and R. Deriche. Geodesic active regions: A new framework to deal with frame partition problems in computer vision. *Journal of Visual Communication and Image Representation*, 13: 249–268, 2002.

N. Paragios and M. Rousson. Shape priors for level set representations. *Proc. ECCV*, pages 78–92, 2002.

T. Pavlidis. *Algorithms for Graphics and Image Processing*, chapter 7. Computer Science Press, Inc., 1982.

M. Rochery, I. H. Jermyn, and J. Zerubia. Étude d'une nouvelle classe de contours actifs pour la détection de routes dans des images de télédétection. In *Proc. GRETSI*, Paris, France, September 2003a.

M. Rochery, I. H. Jermyn, and J. Zerubia. Higher order active contours and their application to the detection of line networks in satellite imagery. In *Proc. IEEE Workshop VLSM*, at ICCV, Nice, France, October 2003b.

M. Rochery, I. H. Jermyn, and J. Zerubia. Gap closure in (road) networks using higher-order active contours. In *Proc. IEEE ICIP.*, Singapore, October 2004.

M. Rochery, I. H. Jermyn, and J. Zerubia. New higher-order active contour energies for network extraction. In *Proc. IEEE ICIP.*, Genoa, Italy, September 2005.

J. A. Sethian. *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Geometry Fluid Mechanics, Computer Vision and Materials Science*. Cambridge University Press, 1999.

J. A. Sethian. Fast marching methods. *SIAM Rev.*, 41-2:199–235, 1996.

K. Siddiqi, B. B. Kimia, and C-W. Shu. Geometric shock-capturing ENO schemes for subpixel interpolation, computation and curve evolution. *Graphical Models and Image Processing*, 59: 278–301, 1997.

A. Steiner, R. Kimmel, and A. M. Bruckstein. Shape enhancement and exaggeration. *Graphical Models and Image Processing*, 60(2):1112–1124, 1998.

R. Stoica, X. Descombes, and J. Zerubia. A Gibbs point process for road extraction from remotely sensed images. *Int'l J. Comp. Vis.*, 57(2):121–136, May 2004.

M. Sussman and E. Fatemi. An efficient, interface-preserving level set redistancing algorithm and its application to interfacial incompressible fluid flow. *SIAM J. Sci. Comp.*, 20(4):1165–1191, 1997.

M. Sussman, P. Smereka, and S. Osher. A level set approach for computing solutions to incompressible 2-phase flow. *J. Comp. Phys.*, 114:146–159, 1994.

F. Tupin, H. Maitre, J-F. Mangin, J-M. Nicolas, and E. Pechersky. Detection of linear features in SAR images: Application to road network extraction. *IEEE Trans. Geoscience and Remote Sensing*, 36(2):434–453, 1998.

A. Vasilevskiy and K. Siddiqi. Flux maximizing geometric flows. *IEEE Trans. Patt. Anal. Mach. Intell.*, 24(12):1565–1578, December 2002.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399