Higher Order Active Contours

Marie Rochery, Ian H. Jermyn and Josiane Zerubia
Ariana (joint research group INRIA/I3S), INRIA, B.P. 93, 06902 Sophia Antipolis Cedex, France.
(`firstname.lastname@inria.fr`)

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.

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

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 $\Pr(\langle R, P \rangle | I, K)$ 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 $\Pr(\langle R, P \rangle | K)$) and about their relations to the data (as likelihood terms $\Pr(I | \langle R, P \rangle, K)$). Attempts to assert such propositions thus have to construct, if only implicitly, probability distributions on the space of regions, $\Pr(R)$, which may depend on other, known or unknown parameters, and the data.

One way of implicitly constructing such distributions is provided by active contours. 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

2

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 paper,[1] we introduce 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. 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. 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 shows an Euclidean invariant quadratic energy term:

$$E(C) \ = \ \int \int dp \ dp' \ \vec{\mathbf{t}}(p) \cdot \vec{\mathbf{t}}(p') \ \Psi(|C(p) - C(p')|) \ . \tag{1}$$

Here, p and p' are coordinates on $\mathrm{dom}\, C$, the domain of the curve C; $\vec{\mathbf{t}}(p) = C'(p)$ 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) 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 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.

---

[1] Portions of this work were published in (**?**).

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 **?** (**?**). 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 **?** (**?**) 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 (**?**; **?**; **?**; **?**) 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 **?** (**?**), **?** (**?**), and **?** (**?**), 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, $\Pr(x|a) \propto \exp(-x \cdot a)$. 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.

4

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. **?** (**?**) 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. **?** (**?**) 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. **?** (**?**) 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. **?** (**?**) 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 **?** (**?**). **?** (**?**) 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, **?** (**?**) 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 2.4, 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 (**?**). 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:

$$E(\alpha C_1 + \beta C_2) \;=\; \alpha E(C_1) + \beta E(C_2) \ . \tag{2}$$

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

$$E(C) \;=\; \int_{\partial R} A = \int_{\mathrm{dom}\, C} C^* A = \int dp \; \vec{\mathbf{t}}(p) \cdot A \ , \tag{3}$$

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;

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

$$\int_R F \;=\; \int_R dA_F = \int_{\partial R} A_F \ . \tag{4}$$

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

$$E(C) \;=\; \frac{1}{2} \int dp \; \vec{\mathbf{t}}(p) \times C(p) = \int dp \; \frac{\partial x}{\partial p} y(C(p)) \ , \tag{5}$$

where (x, y) are Euclidean coordinates. In consequence of equation (4), linear energies of the form (3) encompass all the forms of region energies in the literature. They are also used by **?**

(**?**) as part of a 'ratio energy', and by **?** (**?**) to find 'flux maximizing flows', while **?** (**?**) 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:

$$E(C) \;=\; \int_{\mathrm{dom}\,C} \star_{C^*g}\; C^*f = \int dp \; |\vec{\mathbf{t}}(p)|_g \; f(C(p)) \;\;.$$ (6)

Here, f is a function (0-form) on Ω; C^*g is the metric on $\mathrm{dom}\,C$ induced by C; and $|v|_g$ is the norm of the vector v in the metric g. The form of functional in equation (6) encompasses the remainder of the models mentioned above, including geometric and geodesic active contours, and most others that have appeared in the literature. A particular example is boundary length, in the metric g, which is given by $f = \mathbb{I}$.

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 paper make use of the linear structure of the chain space to go beyond 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.

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:

$$V \xrightarrow{\;\Delta_n\;} V^n \xrightarrow{\;\otimes\;} V^{\otimes n} \xrightarrow{\;E\;} \mathbb{R}$$ (7)

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. One can proceed in several ways, one of which is analogous to equation (6), while another is analogous to equation (3). We do not describe both possibilities here for lack of space, but instead focus on the latter. Given an n-form F on Ω^n, we pull it back to the domain of $C^{\otimes n}$ and integrate it, using the analogue of equation (3):

$$E(C) \;=\; \int_{(\partial R)^n} F = \int_{(\mathrm{dom}\,C)^n} (C^{\otimes n})^* F \;\;.$$ (8)

What is then required is an n-form on Ω^n. In what follows, we will focus on the quadratic case, $n = 2$, both for clarity and because this is what we will use in the application later in the paper.

2.3. QUADRATIC ENERGIES

In the case $n = 2$, equation (8) becomes

$$E(C) = \int_{(\partial R)^2} F = \int_{(\operatorname{dom} C)^2} (C \otimes C)^* F \ . \qquad (9)$$

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

$$E(C) = \int \int dp \ dp' \ \vec{\mathbf{t}}(p) \cdot F(C(p), C(p')) \cdot \vec{\mathbf{t}}(p') \ , \qquad (10)$$

where $F(x, x')$, for each $(x, x') \in \Omega^2$, is a matrix. The operator F allows us to model a nontrivial interaction between different contour points. Note that this interaction is not Markov, even if the value of F tends to zero rapidly with increasing distance between its 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.

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. Note 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

$$E_g(C) = \mathcal{L}(C) + \alpha \mathcal{A}(C) - \beta \int \int dp \ dp' \ \vec{\mathbf{t}} \cdot \vec{\mathbf{t}}' \ \Psi(R(p, p')) \ , \qquad (11)$$

where $\mathcal{L}$ is the length of the boundary in the Euclidean metric on Ω, an energy of the form (6); and $\mathcal{A}$ is the area of its interior, an energy of the form (3). $R(p, p') = |C(p) - C(p')|$ 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 (1), introduces the interactions. We choose the following form for the function Ψ:

$$\Psi(x) = \begin{cases} 1 & \text{if } x < d_{\min} - \epsilon \ , \\ 0 & \text{if } x > d_{\min} + \epsilon \ , \\ \frac{1}{2}(1 - \frac{x - d_{\min}}{\epsilon} - \frac{1}{\pi} \sin(\pi \frac{x - d_{\min}}{\epsilon})) & \text{otherwise} \ . \end{cases} \qquad (12)$$

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

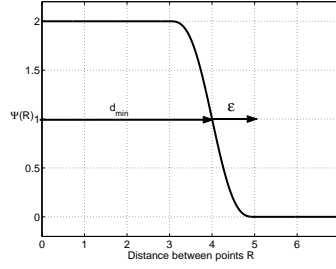*Figure 1.* The function Ψ

in Ω. The function Ψ is always positive, and so from equation (11), 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 (11) is minimized using gradient descent. Thus the contour evolution is determined by

$$\frac{\partial C}{\partial t} = -\frac{\delta E}{\delta C}(C) \; , \tag{13}$$

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

$$\hat{\mathbf{n}} \cdot \dot{C}(p) = -\kappa(p) - \alpha + 2\beta \int dp' \; (\hat{\mathbf{R}}(p,p') \cdot \hat{\mathbf{n}}(p')) \; \Psi'(R(p,p')) \; , \tag{14}$$

where $\hat{\mathbf{R}}(p,p') = (C(p) - C(p'))/|C(p) - C(p')|$. The component of $\partial C/\partial t$ 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 in it. Figure 2 shows examples of evolutions starting from a circle using equation (14). 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 (11). For certain ranges of angular frequency, small perturbations, rather
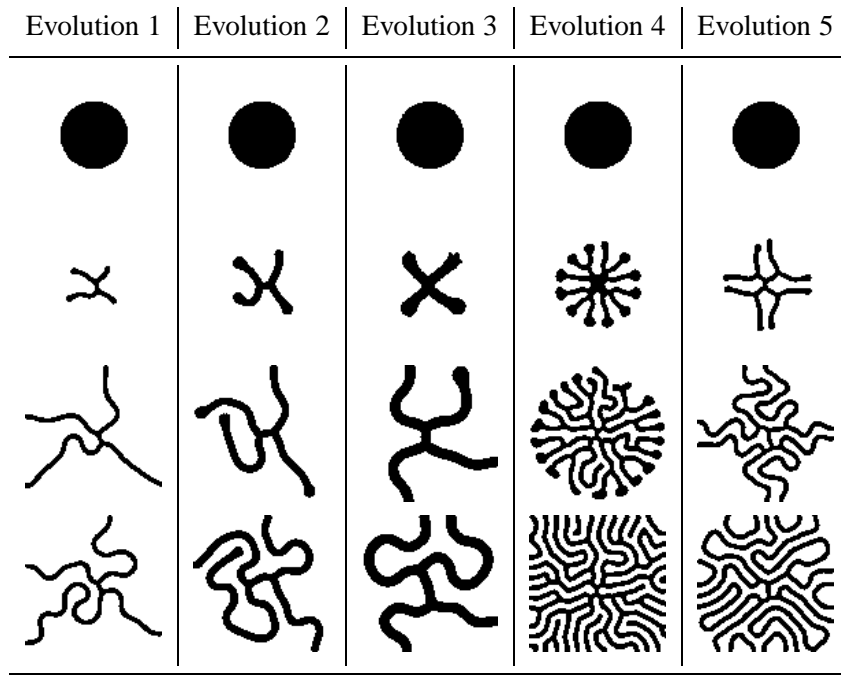
Evolution 1	Evolution 2	Evolution 3	Evolution 4	Evolution 5

Figure 2. Examples of gradient descent using the energy in equation (11). The first three columns correspond to different values of $d_{\min}$, while the last two correspond to different values of α.

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 con-

sequence 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 (11) 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 **?** (**?**). 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.* $\hat{\mathbf{n}} \cdot \dot{C}(p) = F[C](p)$, then the level set function on the contour must obey

$$\dot{\phi} \;=\; -\nabla\phi \cdot F\hat{\mathbf{n}} = F\nabla\phi \cdot \nabla\phi/|\nabla\phi| = F|\nabla\phi| \;. \tag{15}$$

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 (14), 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' (**?**). 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).

3.1. (Re)initialization

In order to (re)initialize ϕ as a signed distance function, we use the approach described by **?** (**?**), where the PDE

$$\phi_t = \text{sign}(\phi_0) \ (1 - |\nabla\phi|) \ , \ \text{with} \ \phi(p, 0) = \phi_0 \tag{16}$$

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 **?** (**?**) have proposed a solution. A local area conservation constraint is imposed by modifying equation (16) in each cell Ω_{ij} of Ω to

$$\phi_t = \text{sign}(\phi_0) \ (1 - |\nabla\phi|) + \lambda_{ij} H'(\phi)|\nabla\phi| \ , \tag{17}$$

where

$$\lambda_{ij} = \frac{-\int_{\Omega_{ij}} H'(\phi)\text{sign}(\phi_0) \ (1 - |\nabla\phi|)}{\int_{\Omega_{ij}} H'(\phi)^2 |\nabla\phi|} \ . \tag{18}$$

The initial condition for equation (17) 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.

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 (**?**). After interpolation, the boundary is extracted using the contour tracing algorithm shown in table I (**?**). 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 3. 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 3. 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 **?** (**?**). To

Table I. 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 3. Leftmost three: configurations encountered in the contour tracing algorithm. Right: an ambiguous configuration.

initialize the process, the grid points closest to the extracted boundary inherit the speed of the closest extracted boundary point. We then solve the PDE

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

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' (**?**), defined by $|\phi(x,y)| < t$, 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 (endpoints, 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. **?** (**?**), 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. **?** (**?**) 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. **?** (**?**) 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. **?** (**?**) 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. **?** (**?**) and **?** (**?**) 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 **?** (**?**) 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. **?** (**?**) 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. **?** (**?**) 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. **?** (**?**) and **?** (**?**) model roads using 'ribbon snakes', active contours with a certain width associated to each point.
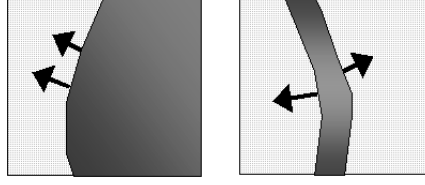
14

Figure 4. The two configurations favoured by the quadratic image term.

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:

$$E(C) = E_g(C) + \lambda E_i(C) \ , \tag{19}$$

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 (11), and is described in section 2.4. The image part E_i is composed of two terms:

$$
\begin{aligned}
E_i(C) &= \int_{\partial R} \star dI - \int_{(\partial R)^2} (\Psi \circ R)(dI \star dI') \\
&= \int dp \ \hat{\mathbf{n}} \cdot \nabla I - \int \int dp \ dp' \ \vec{\mathbf{t}} \cdot \vec{\mathbf{t}}' \ (\nabla I \cdot \nabla I') \ \Psi(R(p, p')) \ ,
\end{aligned} \tag{20}
$$

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 (3), while the quadratic term takes the general form (9).

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 4. First, it favours configurations in which pairs of points whose tangent vectors are parallel and that are not too distant 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 (19) is minimized using gradient descent implemented via level sets as described in section 3. The resulting descent equation is

$$\hat{\mathbf{n}} \cdot \frac{\partial C}{\partial t} = -\kappa - \alpha - \lambda \nabla^2 I + 2\lambda \int dp' \ (\nabla I' \cdot \nabla \nabla I \cdot \hat{\mathbf{n}}') \ \Psi(R(p, p'))$$

Figure 5. Gradient descent on the two SPOT satellite images in the first column.

$$+ \ 2 \int dp' \ (\hat{\mathbf{R}} \cdot \hat{\mathbf{n}}')(\beta + \lambda \nabla I \cdot \nabla I') \ \Psi'(R(p,p')) \ . \tag{21}$$

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 5. 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 5. In both images, the roads are perfectly extracted.

Figure 6 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 occasion-
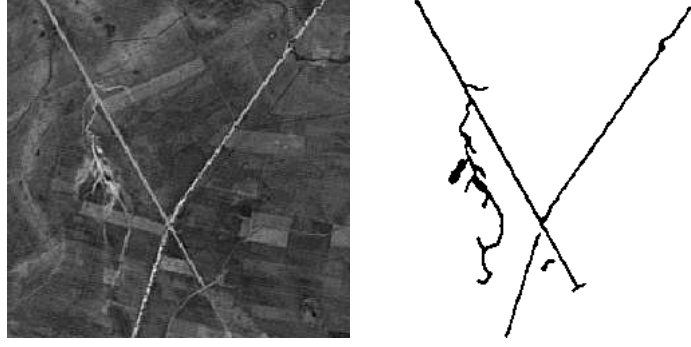
Figure 6. Result on a larger piece of the SPOT image.

ally 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 (3) and (4):

$$E(C) \; = \; -\int_R \star G = -\int_{\partial R} A_G = -\int_{\mathrm{dom}\,C} C^* A_G \;\; , \tag{22}$$

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

$$\frac{\delta E}{\delta C(p)} = G(C(p))\hat{\mathbf{n}}(p) \;\; . \tag{23}$$

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 $[-1, 1]$:

$$Q(x) \; = \; \begin{cases} 1 & \text{if } x < s_1 \;\; , \\ 1 - 2\frac{x-s_1}{s_2-s_1} & \text{if } s_1 \le x \le s_2 \;\; , \\ -1 & \text{if } x > s_2 \;\; , \end{cases} \tag{24}$$
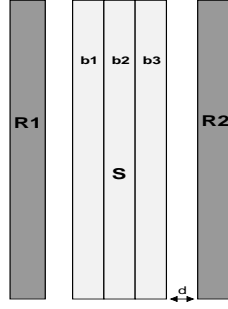
Figure 7. Mask for Student tests.

where s_1 and s_2 are two thresholds, chosen empirically.

4.4.2. *Hypothesis tests*

? (**?**) 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 7.

The Student t-test computes

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

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

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

where

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

The function G is then defined by

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

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.
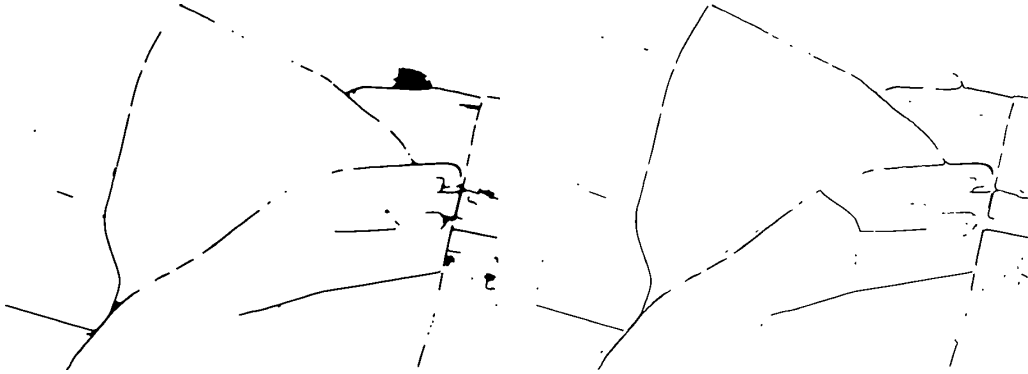
Figure 8. Aerial image. (©IGN.)

Figure 9. Results of extraction with the two functions G.

4.5. EXPERIMENTAL RESULTS WITH THE SECOND MODEL

We add this new energy (22) to the model (19), and test the model on the high-resolution aerial image shown in figure 8. The image presents several difficulties because of high gradients that do 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 9.

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 8, 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.