An Extended Phase Field Higher-Order Active Contour Model for Networks and its Application to Road Network Extraction from VHR Satellite Images

Ting Peng[1,2], Ian H. Jermyn[1], Véronique Prinet[2], and Josiane Zerubia[1]

[1] Project-Team Ariana, INRIA/I3S, 06902 Sophia Antipolis, France
{tpeng, ijermyn, jzerubia}@sophia.inria.fr
[2] LIAMA & NLPR, CASIA, Chinese Academy of Sciences, Beijing 100190, China
{tpeng, prinet}@nlpr.ia.ac.cn

Abstract. This paper addresses the segmentation from an image of entities that have the form of a 'network', *i.e.* the region in the image corresponding to the entity is composed of branches joining together at junctions, *e.g.* road or vascular networks. We present a new phase field higher-order active contour (HOAC) prior model for network regions, and apply it to the segmentation of road networks from very high resolution satellite images. This is a hard problem for two reasons. First, the images are complex, with much 'noise' in the road region due to cars, road markings, etc., while the background is very varied, containing many features that are locally similar to roads. Second, network regions are complex to model, because they may have arbitrary topology. In particular, we address a severe limitation of a previous model in which network branch width was constrained to be similar to maximum network branch radius of curvature, thereby providing a poor model of networks with straight narrow branches or highly curved, wide branches. To solve this problem, we propose a new HOAC prior energy term, and reformulate it as a nonlocal phase field energy. We analyse the stability of the new model, and find that in addition to solving the above problem by separating the interactions between points on the same and opposite sides of a network branch, the new model permits the modelling of two widths simultaneously. The analysis also fixes some of the model parameters in terms of network width(s). After adding a likelihood energy, we use the model to extract the road network quasi-automatically from pieces of a QuickBird image, and compare the results to other models in the literature. The results demonstrate the superiority of the new model, the importance of strong prior knowledge in general, and of the new term in particular.

1 Introduction

The need to segment network-like structures from images arises in a variety of domains. Examples include the segmentation of road and river networks in remote sensing imagery, and of vascular networks in medical imagery. Extracting automatically the region in the image corresponding to the network is a difficult task, however. Because images often contain confounding elements having similar local properties to the entity of interest, techniques that include no prior knowledge about the region containing the network

cannot succeed. In order to solve the problem, such prior knowledge must be injected somehow, either through the intervention of a user, or by incorporating it into a model. Human users possess very specific prior knowledge about the shape of regions corresponding to networks, and in most applications, this level of knowledge is necessary rather than merely sufficient: generic prior knowledge alone, for example concerning boundary smoothness, is not enough. The need to include more specific prior knowledge raises another, methodological issue, however. The set of network-like regions is complicated: it consists of a large (in principle infinite) number of connected components, corresponding to the different possible topologies of a network (number of connected components in the network, number of loops in each connected component), or equivalently to the set of planar graphs (for 2d data). To this is added a geometric superstructure corresponding to an embedding of the graph in the plane, and to its 'fattening' into a region. The construction of a model that favours regions lying in this set as opposed to those outside it is a non-trivial problem. This paper proposes a new model to address this problem, and applies it to the extraction of road networks from very high resolution satellite imagery.

The incorporation into models of prior knowledge about a region to be segmented from an image has a long history. The earliest and still most widely used models incorporate local knowledge about the boundary, essentially smoothness: active contours [1] are one example, the Ising model another [2, 3]. This degree of prior knowledge is almost never enough to segment an entity of interest automatically, even in relatively simple images. More recent work has focused on models that include more specific prior knowledge [4–8]. This work involves shape priors saying that the region sought must be 'close' to an exemplar region or regions. Although useful for many applications, this type of model is not appropriate when the region sought has arbitrary topology.

To model families of regions such as networks, Rochery *et al.* [9] introduced 'higher-order active contours' (HOACs). HOACs incorporate not only local, differential knowledge about the boundary, but also nonlocal, long-range interactions between tuples of contour points. Via such interactions, they favour regions with particular geometric characteristics without constraining the topology via use of a reference region. For example, the model used in [9], which uses pairwise interactions, favours, for certain ranges of parameter values, network-like regions composed of branches with roughly parallel borders and a constant width that meet at junctions.

The HOAC energy developed in [9] suffers from a serious limitation, however, when it is used to model networks. This is that the interaction between points on the same side of a network branch have the same range of interaction as points on opposite sides. The effect is that a typical maximum curvature of a branch κ is connected to the width of that branch W via $\kappa \sim 1/W$. This is particularly limiting for certain types of networks, *e.g.* road networks in cities, for which $\kappa \ll 1/W$.

In this paper, we construct a new HOAC prior energy for modelling networks that overcomes this limitation, allowing separate control of branch straightness and width. The new energy also permits a broader range of widths to be modelled simultaneously, and can even model two disjoint width ranges. We test the model by applying it to the problem of road network extraction from very high resolution (VHR) images of Beijing. This represents an extremely challenging problem due to the amount of 'noise'

in the road regions (cars, road markings, shadows, ...) and the degree of variation and detail in the non-road regions. Nevertheless, the new energy permits a quasi-automatic extraction of the road network.

To avoid the complications of expressing regions with arbitrary topology in terms of boundaries and the complexity of the implementation of HOAC terms using standard level-set methods, Rochery *et al.* [10] reformulated HOAC models as equivalent nonlocal phase field models. Phase fields possess many advantages over more traditional methods for region representation and modelling, even in the non-HOAC case, but are particularly advantageous for HOAC energies. It is often convenient to formulate a model in terms of the contour initially, and then reformulate it as a phase field model for implementation; we follow that procedure in this paper.

The paper is organized as follows: section 2 recalls HOAC energies and the phase field framework. In section 3, we introduce our new HOAC energy, and calculate the conditions for which the model allows stable bars. In section 4, we define the overall model, including a data term. The application of the model to road extraction from VHR images is illustrated in section 5. We conclude in section 6.

2 Higher-Order Active Contours and Phase Fields

In [9], Rochery *et al.* proposed an Euclidean-invariant HOAC energy for modelling network regions:

$$E_C(R) = \lambda_C L(R) + \alpha_C A(R) - \frac{\beta_C}{2} \iint_{(\partial R)^2} dt\, dt'\, \dot\gamma(t) \cdot \dot\gamma(t')\, \Psi\Big(\frac{|\Delta\gamma(t,t')|}{d}\Big)\ , \quad (1)$$

where ∂R is the boundary of region R; $\gamma : S^1 \to \Omega$ is a map representing ∂R, parameterized by t; $\Omega \subset \mathbb{R}^2$ is the image domain; dots represent differentiation wrt t; L is boundary length; A is region area; $\Delta\gamma(t,t') = \gamma(t) - \gamma(t')$; and d is a constant that controls the range of the interaction. The long range interaction between t and t' is modulated by Ψ, the interaction function:

$$\Psi(k) = \begin{cases} \frac{1}{2}\left(2 - |k| + \frac{1}{\pi}\sin(\pi|k|)\right) & \text{if } |k| < 2\ , \\ 0 & \text{else .} \end{cases} \quad (2)$$

It is a smoothly decreasing function from 1 at $k = 0$ to 0 for $k \geq 2$.

For many reasons [10], the phase field framework provides a more convenient framework for region modelling than do contours. A 'phase field' is a function $\phi : \Omega \to \mathbb{R}$, which defines a region $R \in \Omega$ via a threshold z: $R = \zeta_z(\phi) = \{x \in \Omega : \phi(x) > z\}$. The basic phase field energy term E_0 is

$$E_0(\phi) = \int_\Omega dx \left\{ \frac{1}{2} \nabla\phi(x) \cdot \nabla\phi(x) + V(\phi(x)) \right\} . \quad (3)$$

where the 'potential' V is given by

$$V(y) = \lambda(\frac{1}{4}y^4 - \frac{1}{2}y^2) + \alpha(y - \frac{1}{3}y^3)\ , \quad (4)$$

where λ and α are constants. For $\lambda \geqslant \alpha > 0$, V has two minima, at $y = -1$ and $y = 1$, and a maximum at $y = \alpha/\lambda$. Define $\phi_R = \arg\min_{\phi:\, \zeta_z(\phi)=R} E_0(\phi)$. If we ignore the gradient term in equation (3), and set $z = \alpha/\lambda$, we clearly find that $\phi_R(x) = 1$ for $x \in R$ and $\phi_R(x) = -1$ for $x \in \bar{R} = \Omega \setminus R$. Adding the gradient term results in a smooth transition from 1 to -1 over an interface region R_C around the boundary ∂R. Note that to a very good approximation $\nabla\phi$ is non-zero only in R_C. It can be shown [10] that $E_0(\phi_R) \simeq \lambda_C L(R) + \alpha_C A(R)$.

The third, HOAC term in E_C can also be reformulated in terms of an equivalent phase field energy [10]. It becomes

$$E_S(\phi) = -\frac{\beta}{2} \iint_{\Omega^2} dx\, dx'\, \nabla\phi(x) \cdot \nabla\phi(x')\, \Psi\Big(\frac{|x - x'|}{d}\Big)\,. \tag{5}$$

The sum $E_0 + E_S$ is then equivalent to E_C in equation (1) [10].

3 Modelling Networks

As explained briefly in section 1, E_C (or equivalently $E_0 + E_S$) suffers from a significant limitation when it comes to modelling networks. Apart from a sign change, the interaction between two points with parallel tangent vectors is the same, and in particular has the same range, as that between anti-parallel tangent vectors. The former interaction controls the curvature of network branches by trying to align tangent vectors, while the latter controls branch width by creating a repulsive force. Hence we expect that for a stable network branch, typical maximum curvature κ and branch width W will be related approximately by $\kappa \sim 1/W$. Thus, E_C does not model well networks with straight narrow branches or highly curved, wide branches. To overcome these limitations, we will set up a new, Euclidean invariant nonlocal energy term E_L that will act in a complementary way to the HOAC term in E_C. We will also find conditions that ensure that a long bar of a given width is a stable configuration of the new model. This enables the fixing of one of the parameters of the energy in terms of the others, and places constraints on the rest.

3.1 Linear Nonlocal HOAC Term

One general class of quadratic HOAC terms can be written as

$$E_{C,HO}(R) = -\iint_{(\partial R)^2} dt\, dt'\, \dot\gamma(t) \cdot \mathbf{G}_C(\gamma(t), \gamma(t')) \cdot \dot\gamma(t')\,, \tag{6}$$

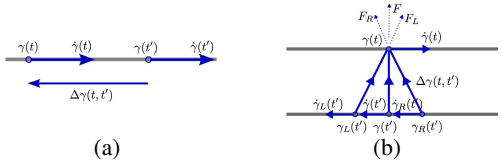
where $\mathbf{G}_C$ is a map from Ω^2 to 2×2 matrices. Imposing Euclidean invariance, and choosing $G_C(\gamma(t), \gamma(t')) = \Psi(|\Delta\gamma|)\delta$, where δ is the unit matrix, leads to the HOAC term in E_C. Choosing $G_C(\gamma(t), \gamma(t')) = \Psi(|\Delta\gamma|)\Delta\gamma\Delta\gamma^T$ leads to

$$E_{C,L}(R) = -\iint_{(\partial R)^2} dt\, dt' \Big[\dot\gamma(t) \cdot \Delta\gamma(t,t')\Big]\Big[\dot\gamma(t') \cdot \Delta\gamma(t,t')\Big] \Psi\Big(\frac{|\Delta\gamma(t,t')|}{d_2}\Big)\,. \tag{7}$$

where we use the same Ψ as in E_C, but with a different range d_2.

$E_{C,L}$ compares each tangent vector to the vector $\Delta\gamma(t,t')$ joining the two interacting points. When two points have tangent vectors that are both nearly aligned or anti-aligned with $\Delta\gamma$, the product of the dot products is positive. The energy $E_{C,L}$ can decrease further by further aligning these tangent vectors with $\Delta\gamma$ and hence with each other. This situation corresponds to two points on the same side of a network branch, as shown in Fig. 1(a). The energy thus favours straight lines, within a range controlled by d_2. On the other hand, when at least one of the two tangent vectors is nearly orthogonal to $\Delta\gamma$, the product of dot products is small. This means that changing the distance between the two points in the argument to Ψ does not change the energy much, and thus that the force between two such points is small. This situation corresponds to two points on opposite sides of a network branch, as shown in Fig. 1(b).

As a result, when $E_{C,L}$ is added to E_C, the width of the network branches is controlled largely by the parameter d of E_C, while the distance over which the branch will be straight is controlled largely by d_2, if $d_2 > d$. For thin, straight bars, we will indeed fix $d_2 > d$. The exception to this rule is also shown in Fig. 1(b). From the above, $\gamma(t')$ exerts no force on $\gamma(t)$, but $\gamma_L(t')$ and $\gamma_R(t')$ both repel $\gamma(t)$, as shown by the force arrows F_L and F_R in the figure. The tangential parts of F_L and F_R cancel, and there is an overall normal repulsion F. If the weight of $E_{C,L}$ in the model is too large, this repulsion may begin to dominate the bar width, which we want to avoid.



(a) (b)

Fig. 1. The effect of $E_{C,L}$: (a) when two tangent vectors are nearly aligned or anti-aligned with $\Delta\gamma$, the energy $E_{C,L}$ favours their alignment; (b) when at least one of the two tangent vectors is nearly orthogonal to $\Delta\gamma$, there is only a very small force between the two points, but contributions from many points can add up to a significant repulsion

We now reformulate $E_{C,L}(R)$ in the phase field framework [10]. We rotate tangent vectors to normal vectors, and replace the latter by $\nabla\phi$. Since $\nabla\phi_R$ is very small outside R_C, the domains of integration can be extended from ∂R to Ω without significantly changing the energy, except for a multiplicative factor. The new, linear nonlocal HOAC phase field term $E_L(\phi)$ becomes (we introduce a weight parameter β_2)

$$E_L(\phi) = -\frac{\beta_2}{2}\iint_{\Omega^2} dx\,dx' \left[\nabla\phi(x)\times(x-x')\right]\left[\nabla\phi(x')\times(x-x')\right]\Psi\Big(\frac{|x-x'|}{d_2}\Big)\ , \quad (8)$$

where $\times$ is the 2D antisymmetric product.

3.2 Stability Analysis

The sum of the three energies we have introduced so far, $E_P = E_0 + E_S + E_L$ will constitute the prior energy for our model. The behaviour of E_P depends on the six parameters $(\alpha, \lambda, \beta, \beta_2, d, d_2)$, and can vary significantly. If we wish to model networks with this energy, it is therefore important to ensure that a network branch is a stable configuration. An important side-effect is that some of the (rather abstract) model parameters are effectively replaced by 'physical' quantities, such as bar and interface width, which we can reasonably fix from numerical or application considerations.

Since network branches are locally like straight bars, we can to a good approximation analyse the stability of a long (because we want to ignore boundary effects) straight bar, of length L and width $W << L$. Ideally, we should minimize E_P under the constraint that $\zeta_z(\phi) = R_{\text{bar}}$, and then expand around that point to test stability, but this is very difficult. Instead, we take a simple *ansatz* for $\phi_{R_{\text{bar}}}$, and study its stability in a low-dimensional subspace of function space; the results may be justified *a posteriori* by numerical experiments. In [10] a similar procedure was followed, the results comparing favourably to those obtained by more sophisticated 'matched asymptotics'. The *ansatz* is as follows. The phase field is given by $\phi(x) = 1$ for $x \in R \setminus R_C$; $\phi(x) = -1$ for $x \in \bar{R} \setminus R_C$, while in R_C, ϕ changes linearly from 1 to -1. The energy E_P evaluated on this *ansatz*, per unit length of bar, which we denote e_P, is given by

$$e_P(w, W) = \frac{4}{3}\alpha W + \frac{4}{15}\lambda w + \frac{4}{w} + \frac{4\beta}{d}\int_W^{2d} d\eta\, \sqrt{\eta^2 - W^2}\,\big(1 - \cos\frac{\pi\eta}{d}\big) + 4\beta_2 \int_W^{2d_2} d\eta\, \eta\sqrt{\eta^2 - W^2}\Big(2 - \frac{\eta}{d_2} + \frac{1}{\pi}\sin\frac{\pi\eta}{d_2}\Big)\,.$$

where w is the width of the interface region R_C. The energy e_P is now minimized with respect to w and W by setting its first derivatives to zero, while ensuring that the second derivatives are positive. For w this is trivial, and leads to $\lambda = 15/w^2$, and thus $\lambda \sim 1$ for reasonable interface widths. For W, the calculation is lengthy and will not be detailed here. Note that stability in fact depends only on the three scaled parameters $\hat{\beta} = \beta/\alpha$, $\hat{\beta}_2 = \beta_2 d^2/\alpha$ and $\hat{d}_2 = d_2/d$, and on the scaled width $\hat{W} = W/d$. The main results are then as follows. If $\hat{d}_2$ is less than a threshold D_2, at most one minimum can be found. If $\hat{d}_2 > D_2$, there are three cases, depending on the values of $\hat{\beta}$, $\hat{\beta}_2$, and $\hat{d}_2$: e_P has no local minimum; e_P has one local minimum, with either $\hat{W} \simeq 1$ (*i.e.* $W \simeq d$) or $\hat{W} \simeq \hat{d}_2$ (*i.e.* $W \simeq d_2$); or e_P has two local minima, at $\hat{W} \simeq 1$ and $\hat{W} \simeq \hat{d}_2$. (This behaviour is an example of a swallowtail catastrophe.) The two regimes are further illustrated in Figs. 2 and 3.

The variety of behaviour is important for applications. As well as being able to model networks with branches of more or less fixed width, but with greater 'stiffness' than provided by the model in [9], the new energy can model two widths at the same time. At certain 'critical points' in parameter space, essentially where pairs of minima merge, it can also model a large range of widths, all of which are approximately stable.

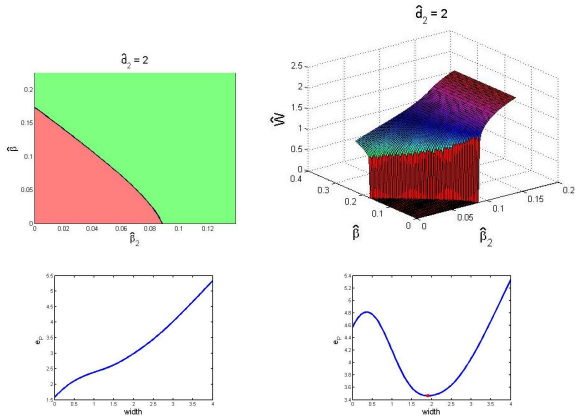


Fig. 2. Top-left: different regions in the $\hat{\beta}_2 - \hat{\beta}$ plane for $\hat{d}_2 = 2 < D_2$. e_P has either no local minimum (red) or one local minimum (green). Top-right: the associated stable bar width. Bottom-left: e_P with no local minimum ($\hat{\beta} = 0.05, \hat{\beta}_2 = 0.04$). Bottom-right: e_P with one local minimum ($\hat{\beta} = 0.2, \hat{\beta}_2 = 0.1$)

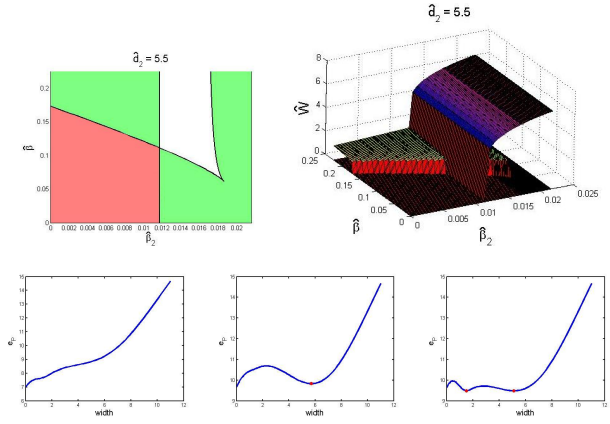


Fig. 3. Top-left: different regions in the $\hat{\beta}_2 - \hat{\beta}$ plane for $\hat{d}_2 = 5.5 > D_2$. e_P has either no local minimum (red), one local minimum (green), or two local minima (white). Top-right: the associated stable bar width(s). Bottom-left: e_P with no local minimum ($\hat{\beta} = 0.1, \hat{\beta}_2 = 0.01$). Bottom-middle: e_P with one local minimum ($\hat{\beta} = 0.05, \hat{\beta}_2 = 0.015$). Bottom-right: e_P with two local minima ($\hat{\beta} = 0.2, \hat{\beta}_2 = 0.013$)

4 Overall Model for Linear Network Extraction

In addition to the prior energy E_P, we also need a likelihood energy linking the region R (which in our case corresponds to the road network) to the data, in our case a VHR optical satellite image. We will also specify some of the implementation details.

4.1 Total Energy

The total energy is the sum of the prior energy E_P and the likelihood energy E_D:

$$E(\phi; I) = E_D(I, \phi) + \theta E_P(\phi) \ , \tag{9}$$

where $I : \Omega \rightarrow \mathbb{R}$ is the image, and $\theta \in \mathbb{R}^+$ balances the contributions of the two terms. E_D is given by

$$E_D(I, \phi) = -\int_\Omega dx \left\{ \phi_+(x) \ln P_+(I(x)) + \phi_-(x) \ln P_-(I(x)) \right\} . \tag{10}$$

$P_\pm(I)$ are models of the histograms of the image intensity, inside $(+)$ and outside $(-)$ the road region. They are both mixtures of Gaussians whose parameters are learned *a priori*, in a supervised way. The quantities $\phi_\pm = (1 \pm \phi)/2$ are, by construction, approximately equal to the characteristic functions of R and $\bar{R}$. The likelihood energy is quite weak, in the sense that maximum likelihood classification produces very poor results (see Fig. 4(d)), mainly due to the 'noise' in the road region and the great variations in the background. No image model with independent pixels can do much better than this, which is why a powerful prior model is needed.

4.2 Optimization and Parameter Settings

To minimize E, we perform gradient descent with the neutral initialisation: the initial value of ϕ is set equal to the threshold $z = \alpha/\lambda$ everywhere in Ω [10]. The algorithm is thus quasi-automatic. The functional derivatives of the HOAC terms $\delta E_S/\delta\phi$ and $\delta E_L/\delta\phi$ involve convolutions: they are calculated in the Fourier domain, as are all derivatives. The evolution equation is

$$\begin{aligned} \frac{\partial\phi(x)}{\partial t} = \frac{1}{2} \ln \frac{P_+}{P_-} + \theta \Big\{ & \nabla^2\phi(x) - \lambda(\phi^3(x) - \phi(x)) - \alpha(1 - \phi^2(x)) \\ & + \beta \mathcal{F}^{-1}\Big\{ k^2 d\hat{\Psi}(kd)\hat{\phi}(k) \Big\} + \beta_2 \mathcal{F}^{-1}\Big\{ \Big(k \cdot \mathcal{F}\{\epsilon x x^T \epsilon^T \Psi(|x|/d_2)\} \cdot k \Big) \hat{\phi}(k) \Big\} \Big\} , \end{aligned} \tag{11}$$

where $\mathcal{F}$ and $\mathcal{F}^{-1}$ denote the Fourier and the inverse Fourier transforms respectively, and a hat $\hat{}$ indicates the Fourier transform of a variable. ϵ rotates the tangent vectors to the inward normal vectors. The time evolution of ϕ uses the forward Euler method. The parameters of the prior energy, *i.e.* θ, α, λ, β, β_2, d, and d_2, are constrained by the stability analysis of section 3.2. This enables a choice of λ, β, β_2, d, and d_2 based on the width(s) to be modelled: only α and θ remain.

5 Experimental Results

As input data I, we use a number of images, with average size 1500×1500 pixels, extracted from a QuickBird optical panchromatic image of Beijing. The scenes are characteristic of dense urban regions. Fig. 4(a) illustrates one of these images. Our aim is to

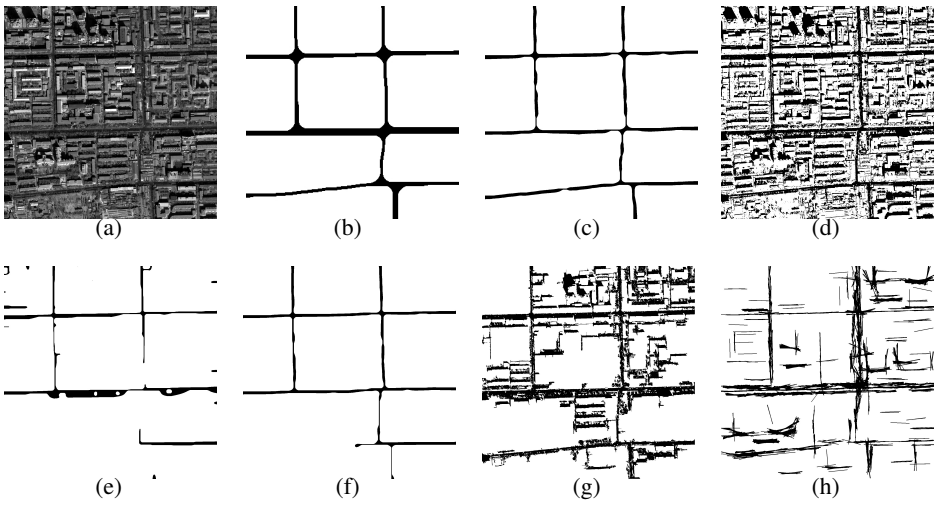


Fig. 4. Experiments and comparisons. 4(a): a QuickBird image, 0.61m; 4(b)-4(c): results obtained using the new model, E, at $1/4$ resolution and at full resolution. 4(d)-4(h): results obtained using: MLE; the model with $\beta = \beta_2 = 0$ (equivalent to a standard active contour); the model in [9] $\beta_2 = 0$; the methods in [11] and [12]

extract, completely and accurately, the road network from an image. In order to evaluate the performance of our new model, we compare it quantitatively to ground truth and to other methods from the literature. We also analyse the effect of the different terms in our energy.

We will focus on two particular cases of road extraction: extraction of a network consisting of roads of roughly the same width; and extraction of networks containing roads of two different widths. In the former case, we choose the parameters so that e_P has one local minimum. The resulting model can extract roads whose widths are close to the minimizing value. In the latter case, we choose the parameters so that e_P has two local minima. Again a small range of widths around each minimum is possible.

5.1 Extraction of Roads of Similar Widths

We apply our model to both the full-resolution and reduced resolution images. We fixed the parameters as described in section 3.2. For all experiments, the parameters $(\theta, \alpha, \lambda, \beta, \beta_2, d, d_2)$ were $(200, 0.15, 4, 0.02, 2 \times 10^{-4}, 4, 12)$ and $(200, 0.15, 4, 0.02, 1.25 \times 10^{-5}, 16, 48)$ at $1/4$ and full resolution respectively. Note that apart from the obvious scaling of d and d_2, and a change in β_2, the other parameters are the same for the two resolutions.

The results obtained using the model E (equation (9)), at $1/4$ resolution and at full resolution, are shown in Fig. 4. The complete road network is retrieved successfully, at both resolutions. Although the segmentation at $1/4$ resolution appears geometrically

Table 1. Quantitative criteria at full resolution (except first row) for Fig. 4(a) (T = True, F = False, P = Positive, N = Negative)

Measure / Method	Completeness TP/(TP+FN)	Correctness TP/(TP+FP)	Quality TP/(TP+FP+FN)
Our model E (with E_L) at $1/4$ resolution, (*e.g.* Fig. 4(b))	0.9688	0.8519	0.8292
Our model E (with E_L) (*e.g.* Fig. 4(c))	0.8756	0.9693	0.8520
MLE (*e.g.* Fig. 4(d))	0.9356	0.2073	0.2044
$\theta E_0 + E_D$ (*e.g.* Fig. 4(e))	0.6047	0.8249	0.5359
$\theta(E_0 + E_S) + E_D$ (*e.g.* Fig. 4(f))	0.6946	0.9889	0.6892
Wang [11] (*e.g.* Fig. 4(g))	0.9350	0.3463	0.3381
Yu [12] (*e.g.* Fig. 4(h))	0.6050	0.3695	0.2977

smoother, the extraction result is actually more accurate at full resolution. Accuracy at $1/4$ resolution is limited both directly, by the low resolution of the phase field, and indirectly, because each scaling coefficient in the data at level 2 is the average of 16 pixels at full resolution: coefficients near the road border therefore include both road and background contributions, and the road boundary is thereby blurred.

To evaluate the performance of the new model, we now compare our result with other methods. In order to illustrate the effects of different terms in the model, we computed results using maximum likelihood estimation (MLE, *i.e.* $\theta = 0$); a standard, non-higher-order active contour, ($\beta = \beta_2 = 0$); and the model in [9] ($\beta_2 = 0$). The results are shown in Figs. 4(d)–4(f). MLE is clearly incapable of distinguishing the roads from the background, while the models without E_S and/or E_L are not able to recover the complete road network (although that with E_S does better than the standard active contour, which has only local prior knowledge). In addition, we apply two other methods, proposed in [11] and [12], and compare them to ours. The approach in [11] is a classification, tracking, and morphology algorithm; [12] introduced a fast but rough segmentation technique based on 'straight line density'. Without much prior geometric knowledge, they extract many incorrect areas that happen to have similar statistical properties to roads. Moreover, the accuracy of the delineation of the road boundary is poor. Some quantitative evaluations based on standard criteria [13], are shown in Table 1. Ground truth for this evaluation was segmented by hand. The 'quality' is the most important measure because it considers both completeness and correctness. Our complete model outperforms all others. Fig. 5 presents more results using E.

5.2 Extraction of Roads of Different Widths

Images containing roads of different widths are processed after choosing parameter values for which $e_{P,L}$ has two local minima. Fig. 6(a) shows an input image containing two roads: their widths are approximately 20 pixels and 80 pixels. The results obtained using our complete model E and the model in [9] (with $\beta_2 = 0$) are illustrated in Figs. 6(b) and 6(c) respectively. The parameter values used in this experiment were $(25, 0.15, 5, 0.02, 1.228 \times 10^{-4}, 4, 22)$. The estimated stable widths for these parame-

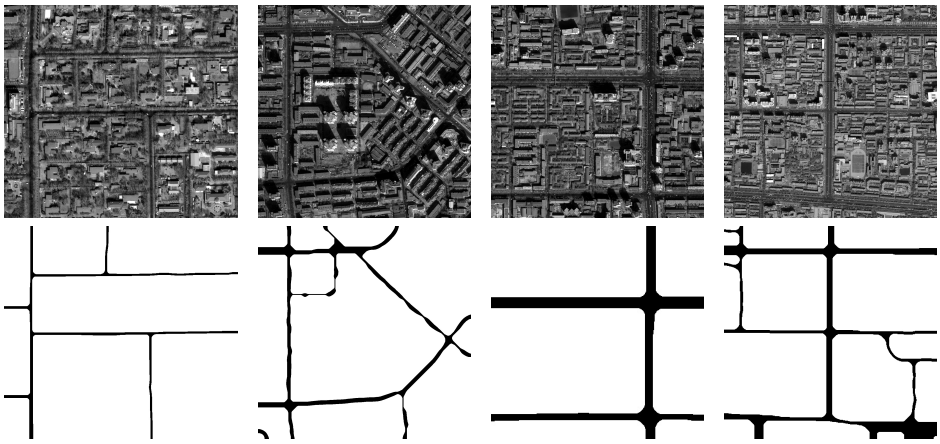
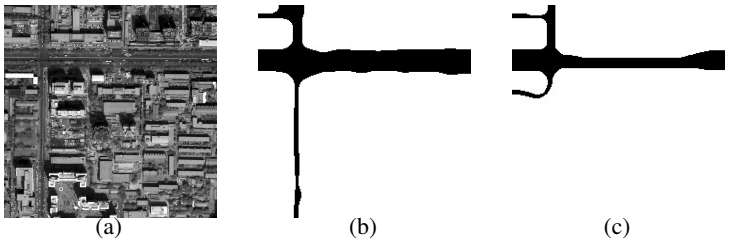


Fig. 5. More results using the model E on pieces of a QuickBird image



(a) (b) (c)

Fig. 6. Extraction of a road network containing two different widths, at $1/4$ resolution. Left to right: image data; results using: the new model E; the model in [9] ($\beta_2 = 0$)

ter values are 5.28 and 20.68, corresponding to the road widths at $1/4$ resolution, *i.e.* 5 pixels and 20 pixels. This comparison shows clearly that adding E_L enables the detection of roads with both widths, while the model without E_L finds only an incomplete network.

In practice, the results are not very sensitive to the precise choice of parameter values, provided they lie in the correct subset of the $\hat{\beta} - \hat{\beta}_2 - \hat{d}_2$ diagram.

6 Conclusions

We have proposed a new HOAC term for modelling bar shape and embedded it in the phase field framework. Based on a stability analysis of a bar with a desired width, we established constraints linking the parameters of the energy function. We explored the possible behaviours of the resulting prior energy E_P as a function of the parameter settings, and showed that as well as separating the interactions between points on the

same and opposite sides of a network branch, the new model permits the modelling of two widths simultaneously. The analysis also fixes some of the model parameters in terms of network width(s). Experiments on road network extraction from VHR satellite images demonstrate the superiority of the new model to others in the literature. Our current work is focused on constructing a prior energy E_P that has a very flat local minimum in a wide range, instead of two sharp local minima. This might be a better solution for the extraction of roads with multiple widths.

ACKNOWLEDGEMENTS

This work was partially supported by European Union Network of Excellence MUSCLE (FP6-507752) and by INRIA Associated Team "SHAPES". The work of the first author is supported by an MAE/Thales Alenia Space/LIAMA grant.

References

1. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. Int. J. Comput. Vis. **1** (1988) 321–331
2. Ising, E.: Beitrag zur theorie des ferromagnetismus. **31** (1925) 253–258
3. Geman, S., Geman, D.: Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. IEEE Trans. Pattern Anal. Mach. Intell. **6** (1984) 721–741
4. Chen, Y., Tagare, H., Thiruvenkadam, S., Huang, F., Wilson, D., Gopinath, K., Briggs, R., Geiser, E.: Using prior shapes in geometric active contours in a variational framework. Int. J. Comput. Vis. **50** (2002) 315–328
5. Cremers, D., Tischhäuser, F., Weickert, J., Schnörr, C.: Diffusion snakes: Introducing statistical shape knowledge into the Mumford-Shah functional. Int. J. Comput. Vis. **50** (2002) 295–313
6. Leventon, M.E., Grimson, W.E.L., Faugeras, O.: Statistical shape influence in geodesic active contours. In: Proc. IEEE CVPR. Volume 1., Hilton Head Island, South Carolina, USA (2000)
7. Rousson, M., Paragios, N.: Shape priors for level set representations. In: Proc. ECCV. Volume 2., Copenhague, Danemark (2002)
8. Srivastava, A., Joshi, S., Mio, W., Liu, X.: Statistical shape analysis: Clustering, learning, and testing. IEEE Trans. Pattern Anal. Mach. Intell. **27** (2003) 590–602
9. Rochery, M., Jermyn, I.H., Zerubia, J.: Higher-order active contours. Int. J. Comput. Vis. **69** (2006) 27–42
10. Rochery, M., Jermyn, I.H., Zerubia, J.: Phase field models and higher-order active contours. In: Proc. IEEE ICCV, Beijing, China (2005)
11. Wang, R., Zhang, Y.: Extraction of urban road network using Quickbird pan-sharpened multispectral and panchromatic imagery by performing edge-aided post-classification. In: Proc. ISPRS, Quebec City, Canada (2003)
12. Yu, Z., Prinet, V., Pan, C., Chen, P.: A novel two-steps strategy for automatic GIS-image registration. In: Proc. ICIP, Singapore (2004)
13. Heipke, C., Mayr, H., Wiedemann, C., Jamet, O.: Evaluation of automatic road extraction. Int. Arch. Photogram. Rem. Sens. **XXXII** (1997) 47–56