

Extended Phase Field Higher-Order Active Contour Models for Networks

Its Application to Road Network Extraction from VHR Satellite Images

Ting Peng · Ian H. Jermyn · Véronique Prinet ·
Josiane Zerubia

Received: 26 May 2009 / Accepted: 9 October 2009
© Springer Science+Business Media, LLC 2009

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 new phase field higher-order active contour (HOAC) prior models for network regions, and apply them 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 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. We solve this problem by introducing first an additional nonlinear nonlocal HOAC term, and then an addi-

tional linear nonlocal HOAC term to improve the computational speed. Both terms allow separate control of branch width and branch curvature, and furnish better prolongation for the same width, but the linear term has several advantages: it is more efficient, and it is able to model multiple widths simultaneously. To cope with the difficulty of parameter selection for these models, we perform a stability analysis of a long bar with a given width, and hence show how to choose the parameters of the energy functions. After adding a likelihood energy, we use both models to extract the road network quasi-automatically from pieces of a QuickBird image, and compare the results to other models in the literature. The state-of-the-art results obtained demonstrate the superiority of our new models, the importance of strong prior knowledge in general, and of the new terms in particular.

Keywords Active contour · Phase field · Shape prior · Parameter analysis · Remote sensing · Road network extraction

T. Peng (✉) · V. Prinet
LIAMA & NLPR, CASIA, 95 Zhongguancun East Road,
Beijing 100190, China
e-mail: Ting.Peng@loria.fr

V. Prinet
e-mail: prinet@nlpr.ia.ac.cn

T. Peng · I.H. Jermyn · J. Zerubia
Project-Team Ariana, INRIA, 2004 route des Lucioles,
06902 Sophia Antipolis, France

I.H. Jermyn
e-mail: Ian.Jermyn@sophia.inria.fr

J. Zerubia
e-mail: Josiane.Zerubia@sophia.inria.fr

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. Automatically extracting the network region in the image is a difficult task, because images are usually complex, containing much noise and confounding elements having similar local properties to the entity of interest. For this reason, techniques that include no prior knowledge about the

region containing the network cannot succeed. In order to solve this 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. Unfortunately, current methods, based on manual extraction, are time and labour intensive. On the other hand, generic prior knowledge alone, for example concerning boundary smoothness, is not enough. The need to include more specific prior knowledge of a class of shapes raises a difficult methodological issue, however. The set of network-like regions is complicated to model, because they may have arbitrary topology. More concretely, 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 two new models to address this problem, and applies them to the extraction of road networks from very high resolution (VHR) 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 (Kass et al. 1988) are one example, the Ising model (Ising 1925; Geman and Geman 1984) another. This degree of prior knowledge is almost never sufficient to segment an entity of interest automatically, even in relatively simple images. More recent work has focused on the inclusion of more specific prior knowledge in active contour models (Chen et al. 2002; Cremers et al. 2002, 2006; Leventon et al. 2000; Riklin-Raviv et al. 2007; Rousson and Paragios 2007; Srivastava et al. 2003). This work uses shape priors saying that the region sought must be ‘close’ to an exemplar region or regions. Although having prior information about the expected shape of the object can significantly increase the robustness of the segmentation algorithm in 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. (2006) 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 Rochery et al. (2006), 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 Rochery et al. (2006) suffers from a limitation, however. This is that the interactions between points on the same side of a network branch have the same range and strength as the interactions between points on opposite sides. The effect is that typical maximum curvature κ and branch width W will be related approximately by $\kappa \sim 1/W$. This is particularly serious for certain types of networks, e.g. road networks in cities, for which $\kappa \ll 1/W$, and prevents successful extraction.

In this paper, we first construct a new *nonlinear* nonlocal HOAC prior energy for modelling networks that overcomes this limitation, allowing separate control of branch straightness and width. To optimize computational efficiency, we then propose another new *linear* nonlocal HOAC prior energy which achieves a similar effect. Moreover, the latter linear energy permits a broader range of widths to be modelled simultaneously, and can even model two disjoint width ranges. We test both models by applying each of them to the problem of road network extraction from VHR images of Beijing. This is 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, our new energies permit 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. (2005) 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 an active contour model in terms of the parametric curve, and then reformulate it as a phase field model for implementation. In this paper, we follow this procedure by considering each new prior energy using first a parametric contour representation, and then a phase field representation.

For the last few decades, road detection from remotely sensed imagery has been extensively studied, due to the variety and importance of the potential applications of an automatic extraction method. A great number of approaches have been proposed (Mena 2003). Early work, such as that by Merlet and Zerubia (1996) and Geman and Jedynak (1996), used path-finding methods suitable for low resolution images, but could only find networks of restricted topology. Péteri and Ranchin (2003) took advantage of a topologically correct graph of the network in order to extract roads

and junctions using two different types of active contours. Lacoste et al. (2005) used marked point processes to model road networks, but the models were appropriate for medium resolution images; the method has not been applied to VHR images. Amo et al. (2006) proposed a region competition based method for providing large-scale GIS information. Hu et al. (2007) detected roads based on shape classification, and then pruned a road tree using a Bayesian decision process. Bertozzi et al. (2007) modified the Cahn-Hilliard equation to achieve fast inpainting of binary imagery. Dobrosotskaya and Bertozzi (2008) developed an extension of this model. Both of them can be used for interpolating simple roads. However, all these methods are restricted to applications on semi-urban areas and using aerial or SAR images. They are not robust enough to be applied to dense urban areas and optical images.

The paper is organized as follows: Sect. 2 recalls HOAC energies and the phase field framework. In Sect. 3, we introduce two new nonlocal HOAC energies. In Sect. 4, we calculate the conditions for which each new total prior model allows stable bars. In Sect. 5, we define the overall model, including a data term. The application of our models to road extraction from VHR images is illustrated in Sect. 6. We conclude in Sect. 7.

2 Higher-Order Active Contours and Phase Fields

Rochery et al. (2006) proposed a Euclidean-invariant HOAC prior energy for modelling network regions:

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

where ∂R is the boundary of region R ; $\gamma : S^1 \rightarrow \Omega$ is a map representing ∂R , parameterized by t ; $\Omega \subset \mathbb{R}^2$ is the image domain; dots represent differentiation with respect to 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:

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

It is a smoothly decreasing function from 1 at $x = 0$ to 0 for $x \geq 2$. In (1), $L(\partial R)$ acts as a regularizer and encourages smoothness of the boundary, while $A(R)$ controls the expansion of the region. The quadratic HOAC term has two

effects: it controls the curvature of network branches by trying to align tangent vectors, and it controls branch width by creating a repulsive force.

For many reasons (Rochery et al. 2005), the phase field framework provides a more convenient framework for region modelling than do contours. A ‘phase field’ is a function $\phi : \Omega \rightarrow \mathbb{R}$, 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

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

The ‘potential’ U is given by

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

where λ and α are constants. For $\lambda \geq \alpha > 0$, U 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 (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 has been shown in Rochery et al. (2005) that $E_0(\phi_R) \simeq \lambda_C L(\partial R) + \alpha_C A(R)$, i.e. E_0 corresponds to the two linear terms in E_C .

The third, i.e. the HOAC term in E_C , can also be reformulated in terms of an equivalent phase field energy (Rochery et al. 2005). It becomes

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

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

As explained briefly in Sect. 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 on the same side of a network branch ($\dot{\gamma}(t) \cdot \dot{\gamma}(t') > 0$) has the same strength and range as the interaction between two points on opposite sides of a network branch ($\dot{\gamma}(t) \cdot \dot{\gamma}(t') < 0$). The effect is that for a stable network branch, a typical curvature of a branch κ is connected to the width of that branch W approximately by $\kappa \sim 1/W$. In other words, the length/range along which the network branch is expected to be straight is the same as the width of the branch itself. The standard HOAC prior energy E_C (or its phase field counterpart $E_0 + E_S$) thereby provides a poor model of networks with straight narrow branches or highly curved, wide branches. In our application of road extraction in cities, road width gives only an (approximate) upper bound on the radius of curvature of the road: most roads

are much straighter than they are wide. For narrow roads, this is particularly problematic, since the road region is relatively unconstrained due to the small road width.

3 Modelling Networks

To solve these problems, we need to be able to model longer-range or stronger interactions along the road, without changing the interactions across the road. This means that we have to separate the two interaction functions, and hence allow separate control of branch straightness and width, thereby allowing more sophisticated prior knowledge to be included. To achieve this goal, we will construct new nonlocal HOAC prior energies that act in a complementary way to the standard HOAC term. The first new energy term is a nonlinear, nonlocal HOAC energy E_{NL} , which increases the magnitude of the interaction along one side of a network branch. The second is a linear, nonlocal HOAC energy E_L , which provides a longer-range interaction along one side of a network branch. Through the stability analysis of this model, we further demonstrate that the linear nonlocal term permits the modelling of two widths simultaneously. These new models have been previously introduced in Peng et al. (2008a, 2008b). Here we present and analyse them in more detail.

3.1 Nonlinear Nonlocal HOAC Term

In order to separate the two interactions, the interaction function must depend on the tangent/normal vectors at the pairs of points that are its argument. Although the length scale in the interaction function of (1), d , could be made to depend on the inner product between the tangent/normal vectors at the two pixels, it would lead to complicated functional derivatives. Alternatively, we prefer to perform a linear interpolation between two interaction functions. In the contour formulation, our new HOAC prior energy E_{HO} takes the form:

$$E_{HO}(\gamma) = - \iint_{S^1 \times S^1} ds ds' \left\{ f_{+\parallel}(\dot{\gamma}(s) \cdot \dot{\gamma}(s')) \Psi_{+\parallel} - f_{-\parallel}(\dot{\gamma}(s) \cdot \dot{\gamma}(s')) \Psi_{-\parallel} \right\}, \tag{6}$$

where $\gamma : S^1 \rightarrow \Omega$, is an arc length parameterization of the region boundary ∂R ; $\dot{\gamma}(s)$ is the tangent vector to the boundary at s (thus $\dot{\gamma}(s) \cdot \dot{\gamma}(s') \in [-1, 1]$); ‘+ \parallel ’ denotes parallel vectors and ‘- \parallel ’ denotes antiparallel vectors. We define $f_{+\parallel}(x), f_{-\parallel}(x) : [-1, 1] \rightarrow [0, 1]$ by:

$$f_{+\parallel}(x) = (1 + x)/2, \tag{7a}$$

$$f_{-\parallel}(x) = (1 - x)/2. \tag{7b}$$

$\Psi_{+\parallel}$ and $\Psi_{-\parallel}$ are interaction functions similar to that in (2), but have different range or magnitude. They compete with each other: when $\dot{\gamma}(s) \cdot \dot{\gamma}(s') \in [0, 1]$, *i.e.* the two interacting tangent vectors are more parallel, $\Psi_{+\parallel}$ is dominant; while when $\dot{\gamma}(s) \cdot \dot{\gamma}(s') \in [-1, 0]$, *i.e.* the two interacting tangent vectors are more antiparallel, $\Psi_{-\parallel}$ is dominant. Here we decide to adjust only the magnitude of the interaction (although this effectively changes its range also). Thus, we assume that the magnitude of the interaction of parallel vectors is stronger than that of antiparallel vectors, *i.e.* $\Psi_{+\parallel} = a\Psi_{-\parallel}$, where $a > 1$, is a constant. Then, (6) becomes

$$E_{HO}(\gamma) = -\frac{1}{2} \iint_{S^1 \times S^1} ds ds' [(a - 1) + (a + 1)(\dot{\gamma}(s) \cdot \dot{\gamma}(s'))] \Psi_{-\parallel}. \tag{8}$$

In order to implement $E_{HO}(\gamma)$ in the phase field framework, it needs to be reformulated as a function of the phase field ϕ , instead of the arc length parameterization γ used in (8). Since the constant length of $\dot{\gamma}(s)$ corresponds to the fixed change in ϕ across the interface, we replace tangent vectors by normal vectors, and then normal vectors by $\nabla\phi$. Subsequently, the range of interactions is extended from the region boundary ∂R to the whole of the image domain Ω . Due to the fact that $\nabla\phi(x)$ is approximately equal to zero everywhere outside the narrow interface R_C in Ω , the boundary indicator function

$$S(\phi) = (\nabla\phi(x) \cdot \nabla\phi(x))(\nabla\phi(x') \cdot \nabla\phi(x')) \simeq \begin{cases} 1 & \forall x, x' \in R_C, \\ 0 & \text{otherwise,} \end{cases} \tag{9}$$

is inserted into the first term of (8). Thus we have

$$E_{HO}(\phi) = -\frac{1}{2} \iint_{\Omega^2} dx dx' [(a - 1)S(\phi) + (a + 1)(\nabla\phi(x) \cdot \nabla\phi(x'))] \Psi\left(\frac{|x - x'|}{d}\right). \tag{10}$$

When $a = 1$, this reduces to the standard phase field HOAC term E_S (up to a factor of $\beta/2$). Therefore, we define our new additional energy term E_{NL} by

$$E_{NL}(\phi) = -\frac{\beta_2}{4} \iint_{\Omega^2} dx dx' (\nabla\phi(x) \cdot \nabla\phi(x)) \times (\nabla\phi(x') \cdot \nabla\phi(x')) \Psi\left(\frac{|x - x'|}{d}\right), \tag{11}$$

where $\beta_2 > 0$ is a constant. (Note that d is the same as that in E_S .) Thus in this case, we have a term that is quartic in $\nabla\phi$.

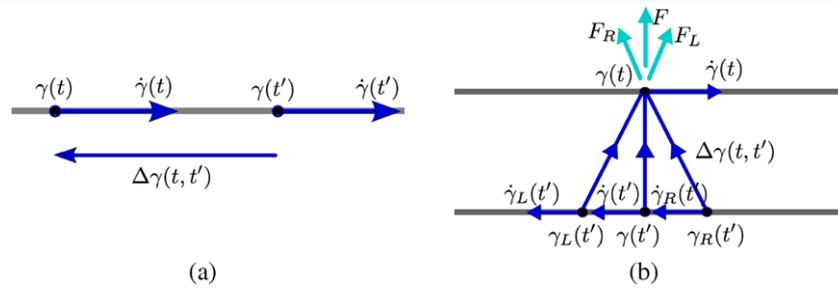


Fig. 1 The effects of E_L in (14) (blue arrow: vector; cyan arrow: interaction force; black dot: interacting point). (a) When two tangent vectors are nearly aligned or anti-aligned with $\Delta\gamma$, the energy E_L favours their alignment; (b) when at least one of the two tangent

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, as denoted by F

The functional derivative of E_{NL} is

$$\begin{aligned} & \frac{\delta E_{NL}(\phi)}{\delta \phi(x)} \\ &= \beta_2 \int_{\Omega} dx' \left\{ \nabla^2 \phi(x) \Psi\left(\frac{|x-x'|}{d}\right) (\nabla \phi(x') \cdot \nabla \phi(x')) \right. \\ & \quad \left. + (\nabla \phi(x') \cdot \nabla \phi(x')) \left(\nabla \phi(x) \cdot \nabla \Psi\left(\frac{|x-x'|}{d}\right) \right) \right\}. \end{aligned} \tag{12}$$

Since this functional derivative of E_{NL} contains a term nonlinear in $\nabla\phi$, due to $S(\phi)$ being $O(\phi^4)$, we refer to it as the *nonlinear nonlocal* term.

Whether the two tangent/normal vectors at a pair of interacting points are parallel or antiparallel, the effect of E_{NL} is always to encourage two points inside the range of the interaction to attract each other. Thus E_{NL} weakens the repulsive effect of E_S between opposite sides, so that E_S along a network branch can be strengthened without changing the width. Consequently, the interaction between pairs of points on the same side of a network branch is stronger than that between pairs of points on opposite sides of a network branch.

3.2 Linear Nonlocal HOAC Term

In the previous subsection, we proposed a nonlinear nonlocal HOAC prior term E_{NL} to overcome the limitation of the standard HOAC prior term. The term E_{NL} reinforces the interaction along the bar branch and weakens the interaction across the bar branch, but both interactions still depend on the same interaction function, hence the effects in these two directions cannot be tuned in a completely independent way. In this section, as an alternative to E_{NL} , we construct a new, Euclidean invariant linear nonlocal HOAC prior term E_L that works in rather a different way.

As we have seen in (1), one general class of quadratic HOAC terms can be written as

$$E_{HO}(\gamma) = - \iint_{(\partial R)^2} dt dt' \dot{\gamma}(t) \cdot \mathbf{G}_C(\gamma(t), \gamma(t')) \cdot \dot{\gamma}(t'), \tag{13}$$

where \mathbf{G}_C is a map from Ω^2 to 2×2 matrices. Imposing Euclidean invariance on this term leads to several possibilities. One is $G_C(\gamma(t), \gamma(t')) = \Psi(|\Delta\gamma|/d) \delta$, where δ is the unit matrix, and $\Delta\gamma = \gamma(t) - \gamma(t')$. Another is $G_C(\gamma(t), \gamma(t')) = \Psi(|\Delta\gamma|/d) \Delta\gamma \Delta\gamma^T$. The former leads to the standard HOAC term in E_C (1). The latter leads to our new linear nonlocal HOAC prior energy, E_L :

$$\begin{aligned} E_L(\gamma) = & - \iint_{(\partial R)^2} dt dt' \left\{ (\dot{\gamma}(t) \cdot \Delta\gamma(t, t')) \right. \\ & \left. \times (\dot{\gamma}(t') \cdot \Delta\gamma(t, t')) \Psi\left(\frac{|\Delta\gamma(t, t')|}{d_2}\right) \right\}, \end{aligned} \tag{14}$$

where we use the same Ψ (2) as in E_C , but with a different range d_2 .

E_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_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 the dot products is small. In this configuration, changing the distance between the two points in the argument to Ψ does not change the energy much, and thus 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_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 again shown in Fig. 1(b). From the above, $\gamma(t')$ exerts no force on $\gamma(t)$, but for both $\gamma_L(t')$ and $\gamma_R(t')$, the product of the dot products is negative. The energy E_L can decrease when the value of Ψ becomes less positive, i.e. both $\gamma_L(t')$ and $\gamma_R(t')$ 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_L in the model is too large, this repulsion may begin to dominate the bar width.

Making the following change of variables in the linear nonlocal HOAC term E_L in (14):

$$\dot{\gamma}(t) = (\cos \theta, \sin \theta), \quad \dot{\gamma}(t') = (\cos \theta', \sin \theta'),$$

$$\Delta\gamma(t, t') = |\Delta\gamma(t, t')|(\cos \eta, \sin \eta),$$

E_L can be rewritten (up to a multiplicative factor) as

$$E_L(\gamma) = - \iint_{(\partial R)^2} dt dt' |\Delta\gamma(t, t')|^2 (\cos(\bar{\theta} - \bar{\theta}') + \cos(\bar{\theta} + \bar{\theta}')) \Psi\left(\frac{|\Delta\gamma(t, t')|}{d_2}\right), \quad (15)$$

where $\bar{\theta} = \theta - \eta$ and $\bar{\theta}' = \theta' - \eta$ are the angles made by $\dot{\gamma}$ and $\dot{\gamma}'$ with $\Delta\gamma$. Thus E_L is a function not only of $\bar{\theta} - \bar{\theta}'$, like the standard HOAC quadratic term in (1), but also of $\bar{\theta} + \bar{\theta}'$. We will continue to use the form of E_L given in (14), however, as this is more convenient for analysis and implementation.

We now reformulate $E_L(\gamma)$ in the phase field framework. We rotate tangent vectors to normal vectors, and replace the latter by $\nabla\phi$. Since $\nabla\phi$ 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. By introducing a weight parameter β_3 , we define the linear nonlocal HOAC phase field term $E_L(\phi)$ as

$$E_L(\phi) = -\frac{\beta_3}{2} \iint_{\Omega^2} dx dx' (\nabla\phi(x) \times (x - x')) \times (\nabla\phi(x') \times (x - x')) \Psi\left(\frac{|x - x'|}{d_2}\right), \quad (16)$$

where \times is the 2D vectorial antisymmetric product. The functional derivative of E_L is

$$\frac{\delta E_L(\phi)}{\delta\phi(x)} = \beta_3 \int_{\Omega} dx' \nabla \cdot (\epsilon(x - x')(x - x')^T \epsilon^T) \cdot \nabla\phi(x') \Psi\left(\frac{|x - x'|}{d_2}\right), \quad (17)$$

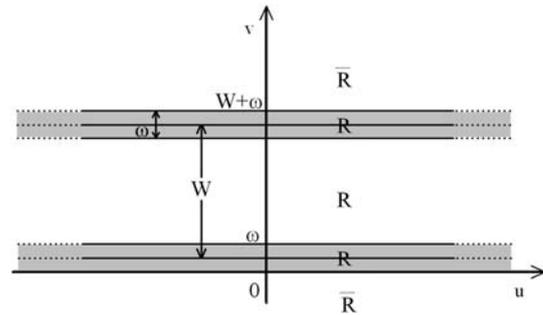


Fig. 2 A bar of length $L \rightarrow \infty$ and width W

where ϵ rotates the tangent vectors to the inward normal vectors. Since this functional derivative is linear in ϕ , we call E_L the *linear nonlocal term*.

4 Stability Analysis

We now add the new nonlinear term E_{NL} or the new linear term to the standard model $E_0 + E_S$. The sum of the three energies $E_{P,NL} = E_0 + E_S + E_{NL}$ constitutes the nonlinear nonlocal HOAC total prior model and $E_{P,L} = E_0 + E_S + E_L$ constitutes the linear nonlocal HOAC total prior model. For both models, there are a number of parameters to tune, $(\alpha, \lambda, \beta, \beta_2, d)$ for the former and $(\alpha, \lambda, \beta, \beta_3, d, d_2)$ for the latter. Unfortunately, not all parameter values allow stable network structures. The behaviour of the prior energy depends on the parameter settings, and can vary significantly. If we wish to model networks with this energy, it is therefore very important to analyse the stability of a network structure and to deduce the resulting constraints on the parameters.

By considering the total prior energy of a long, straight bar of a given width, we establish constraints on the parameters that ensure that a long network branch of the desired width is a stable configuration of the energy functional. 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. Note that, in addition, to guarantee the Turing stability of the model, Rochery et al. (2005) introduced another constraint, which we will not detail here.

4.1 Definition of a Bar

Since network branches are locally like straight bars, we can, to a good approximation, analyse the stability of a long straight bar, of length L and width $W \ll L \rightarrow \infty$. This allows us to ignore boundary effects. Denote the width of the interface by w . Such a bar is shown in Fig. 2.

Ideally, we should minimize the prior energy under the constraint that $\zeta_z(\phi) = R_{\text{bar}}$ (where R_{bar} denotes the bar region), 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 Rochery et al. (2005), a similar procedure was followed, the results comparing favourably to those obtained by more sophisticated ‘matched asymptotics’. The *ansatz* is defined 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 .

Now we evaluate both prior energies on this *ansatz*, per unit length of bar. The constraint that the network branch is a local energy extremum requires the first derivatives with respect to w and W to be zero, while the constraint that it is a local minimum requires that the second derivatives must be positive semi-definite. These requirements provide the stability conditions of the model.

4.2 Nonlinear Nonlocal HOAC Term

We evaluate the energy $E_{P,NL}$ on the *ansatz*, per unit length of bar, which is denoted by $e_{P,NL}$. Up to an additive constant, it is given by

$$e_{P,NL}(w, \hat{W}) = \frac{4}{3}\alpha\hat{W}d + \frac{4}{15}\lambda w + \frac{4}{w} - \frac{16\beta_2d}{w^2} + 4d\left(\beta - \frac{2\beta_2}{w^2}\right) \int_{\hat{W}}^2 d\eta \sqrt{\eta^2 - \hat{W}^2} \times (1 - \cos(\pi\eta)), \tag{18}$$

where the scaled width $\hat{W} = W/d$. Its first derivatives are set to zero:

$$\frac{\partial e_{P,NL}}{\partial \hat{W}} = \frac{4}{3}\alpha d - 4\hat{W}d\left(\beta - \frac{2\beta_2}{w^2}\right) \times \int_{\hat{W}}^2 d\eta \frac{1}{\sqrt{\eta^2 - \hat{W}^2}} (1 - \cos(\pi\eta)) = 0, \tag{19a}$$

$$\frac{\partial e_{P,NL}}{\partial w} = \frac{4}{15}\lambda - \frac{4}{w^2} + \frac{16\beta_2d}{w^3} \times \int_{\hat{W}}^2 d\eta \frac{1}{\sqrt{\eta^2 - \hat{W}^2}} (1 - \cos(\pi\eta)) = 0. \tag{19b}$$

This is a complicated system of two-variable equations. To simplify the problem, we fix arbitrarily the value of w beforehand ($w = 2 \sim 4$), and then solve (19a) to get a sub-optimal solution of \hat{W} . In this way, there is no need to calculate the second derivative of $e_{P,NL}$ with respect to w , but we still need to ensure that the second derivative of $e_{P,NL}$

with respect to \hat{W} is non-negative. It is given by

$$\frac{\partial^2 e_{P,NL}}{\partial \hat{W}^2} = 4d\left(\beta - \frac{2\beta_2}{w^2}\right) \int_{\hat{W}}^2 d\eta \frac{1 + \pi^2(\eta^2 - \hat{W}^2)}{\sqrt{\eta^2 - \hat{W}^2}} \cos(\pi\eta) - 4d\left(\beta - \frac{2\beta_2}{w^2}\right) \ln\left(\frac{2 + \sqrt{4 - \hat{W}^2}}{\hat{W}}\right) \geq 0. \tag{20}$$

The above analysis shows that stability of the nonlinear nonlocal HOAC total prior model is related to the scaled control parameters $\hat{\beta} = \beta/\alpha$ and $\hat{\beta}_2 = \beta_2/\alpha$, and to the scaled width $\hat{W} = W/d$. Therefore, we obtain the parameter constraints:

$$\hat{\beta} - \frac{2\hat{\beta}_2}{w^2} = \frac{1}{3\hat{W}I_1(\hat{W})}, \tag{21a}$$

$$I_2(\hat{W}) \geq \ln\left(\frac{2 + \sqrt{4 - \hat{W}^2}}{\hat{W}}\right), \tag{21b}$$

where

$$I_1(\hat{W}) = \int_{\hat{W}}^2 d\eta \frac{1}{\sqrt{\eta^2 - \hat{W}^2}} (1 - \cos(\pi\eta)), \tag{22a}$$

$$I_2(\hat{W}) = \int_{\hat{W}}^2 d\eta \frac{1 + \pi^2(\eta^2 - \hat{W}^2)}{\sqrt{\eta^2 - \hat{W}^2}} \cos(\pi\eta). \tag{22b}$$

Now we can draw the diagram of the relationship between $\hat{\beta}$, $\hat{\beta}_2$, and \hat{W} , which takes the form of a 3D surface. When w is set to 2, an example of such a surface is illustrated in Fig. 3. The left-hand part of the surface represents the extremum as a local maximum, *i.e.* the second derivative is negative (this is an unstable configuration); the right-hand part of the surface represents the extremum as a local minimum, *i.e.* the second derivative is positive; and the borderline between these two surfaces, where $\hat{W} = 0.8798$, indicates the situation when two extrema of $e_{P,NL}$ become an inflection point (*i.e.* $\partial e_{P,NL}/\partial \hat{W} = 0$ and $\partial^2 e_{P,NL}/\partial \hat{W}^2 = 0$). Clearly, once \hat{W} is given, $\hat{\beta}$ and $\hat{\beta}_2$ have to be chosen in the right-hand part of this surface. For each given width \hat{W} , there is a curve that indicates all the possible pairs of scaled parameters $\hat{\beta}$ and $\hat{\beta}_2$. Note that, due to the scaling of $\hat{\beta}$ and $\hat{\beta}_2$, the dependence on α has been eliminated.

4.3 Linear Nonlocal HOAC Term

Similarly, the linear nonlocal HOAC total prior energy per unit length of bar, $e_{P,L}$, is

$$e_{P,L}(w, \hat{W}) = \frac{4}{3}\alpha\hat{W}d + \frac{4}{15}\lambda w + \frac{4}{w}$$

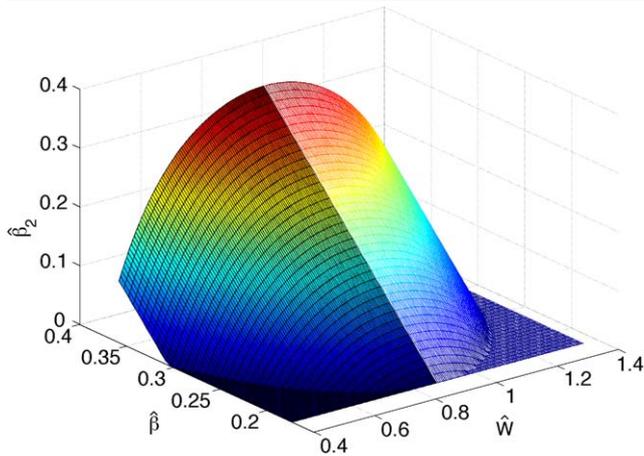


Fig. 3 Diagram of the scaled control parameters $\hat{\beta}$ and $\hat{\beta}_2$, and the scaled width \hat{W} , when $w = 2$. The *left* part of the surface represents the parameter space for which $e_{P,NL}$ has one local maximum; in the *right* part of the surface, $e_{P,NL}$ has one local minimum; and the borderline between them, where $\hat{W} = 0.8798$, represents the situation when two extrema become an inflection point

$$\begin{aligned}
 &+ 4\beta d \int_{\hat{W}}^2 d\eta \sqrt{\eta^2 - \hat{W}^2} (1 - \cos(\pi\eta)) \\
 &+ 4\beta_3 d^3 \int_{\hat{W}}^{2\hat{d}_2} d\eta \eta \sqrt{\eta^2 - \hat{W}^2} \\
 &\times \left(2 - \frac{\eta}{\hat{d}_2} + \frac{1}{\pi} \sin\left(\frac{\pi\eta}{\hat{d}_2}\right) \right). \tag{23}
 \end{aligned}$$

As for $e_{P,NL}$, we find the minimum of $e_{P,L}$ by setting its first derivatives, with respect to w and \hat{W} , to zero, while ensuring that its second derivative, with respect to \hat{W} , is non-negative. For the derivatives, we find

$$\begin{aligned}
 \frac{\partial e_{P,L}}{\partial \hat{W}} &= \frac{4}{3}\alpha d - 4\beta \hat{W} d \int_{\hat{W}}^2 d\eta \frac{1}{\sqrt{\eta^2 - \hat{W}^2}} (1 - \cos(\pi\eta)) \\
 &\quad - \frac{4\beta_3 \hat{W} d^3}{\hat{d}_2} \int_{\hat{W}}^{2\hat{d}_2} d\eta \sqrt{\eta^2 - \hat{W}^2} \left(1 - \cos\left(\frac{\pi\eta}{\hat{d}_2}\right) \right), \tag{24a}
 \end{aligned}$$

$$\frac{\partial e_{P,L}}{\partial w} = \frac{4}{15}\lambda - \frac{4}{w^2}, \tag{24b}$$

$$\begin{aligned}
 \frac{\partial^2 e_{P,L}}{\partial \hat{W}^2} &= -4\beta d \ln\left(\frac{2 + \sqrt{4 - \hat{W}^2}}{\hat{W}}\right) \\
 &\quad + 4\beta d \int_{\hat{W}}^2 d\eta \frac{1 + \pi^2(\eta^2 - \hat{W}^2)}{\sqrt{\eta^2 - \hat{W}^2}} \cos(\pi\eta)
 \end{aligned}$$

$$\begin{aligned}
 &- \frac{4\beta_3 d^3}{\hat{d}_2} \int_{\hat{W}}^{2\hat{d}_2} d\eta \frac{\eta^2 - 2\hat{W}^2}{\sqrt{\eta^2 - \hat{W}^2}} \\
 &\times \left(1 - \cos\left(\frac{\pi\eta}{\hat{d}_2}\right) \right). \tag{24c}
 \end{aligned}$$

It follows that stability is related to the three scaled control parameters $\hat{\beta} = \beta/\alpha$, $\hat{\beta}_3 = \beta_3 d^2/\alpha$ and $\hat{d}_2 = d_2/d$, and also to the scaled width $\hat{W} = W/d$. The parameter constraints are then as follows:

$$1 - 3\hat{\beta}\hat{W}I_1(\hat{W}) - \frac{3\hat{\beta}_3\hat{W}}{\hat{d}_2}I_3(\hat{W}) = 0, \tag{25a}$$

$$\hat{\beta}I_2(\hat{W}) - \frac{\hat{\beta}_3}{\hat{d}_2}I_4(\hat{W}) \geq \hat{\beta} \ln\left(\frac{2 + \sqrt{4 - \hat{W}^2}}{\hat{W}}\right), \tag{25b}$$

$$\lambda = \frac{15}{w^2}, \tag{25c}$$

where $I_1(\hat{W})$ and $I_2(\hat{W})$ have been defined in (22); and $I_3(\hat{W})$ and $I_4(\hat{W})$ take the form:

$$I_3(\hat{W}) = \int_{\hat{W}}^{2\hat{d}_2} d\eta \sqrt{\eta^2 - \hat{W}^2} \left(1 - \cos\left(\frac{\pi\eta}{\hat{d}_2}\right) \right), \tag{26a}$$

$$I_4(\hat{W}) = \int_{\hat{W}}^{2\hat{d}_2} d\eta \frac{\eta^2 - 2\hat{W}^2}{\sqrt{\eta^2 - \hat{W}^2}} \left(1 - \cos\left(\frac{\pi\eta}{\hat{d}_2}\right) \right). \tag{26b}$$

For w , the constraint is trivial, leading to $\lambda = 15/w^2$. The stable width \hat{W} , *i.e.* the value of \hat{W} where a local minimum of $e_{P,L}$ is found, depends on the parameters $\hat{\beta}$, $\hat{\beta}_3$ and \hat{d}_2 . The term $\hat{W}I_1(\hat{W})$ in (25a) is a simple curve with one maximum inside the interval $[0, 2)$, so when $\hat{\beta}_3 = 0$, for a given $\hat{\beta}$, the number of solutions of this equation is at most two: one corresponds to the local minimum of the energy, the other the local maximum. The additional term $\hat{W}I_3(\hat{W})/\hat{d}_2$ is also a simple curve with one maximum but in the interval $[0, 2\hat{d}_2)$. As a result, when \hat{d}_2 is small, the two curves are mostly overlapped, and (25a) possesses the same properties as when $\hat{\beta}_3 = 0$: the number of possible solutions is zero or two; while when \hat{d}_2 is large enough so that one curve has non-zero values inside a longer interval than the other, the first constraint changes significantly: the number of possible solutions satisfying $\partial E_{P,L}/\partial \hat{W} = 0$, is zero, two, or four.

Consequently, there is a singular point of \hat{d}_2 , where the maximum number of solutions jumps from two to four, and accordingly the stable width(s) that the prior energy can model jump(s) from one to two. However, since (25a) involves complicated integrals and the piecewise interaction function, it is non-trivial to obtain analytically the number of its solutions. Therefore, we carry out numerical experiments to find an approximate value for this singular point.

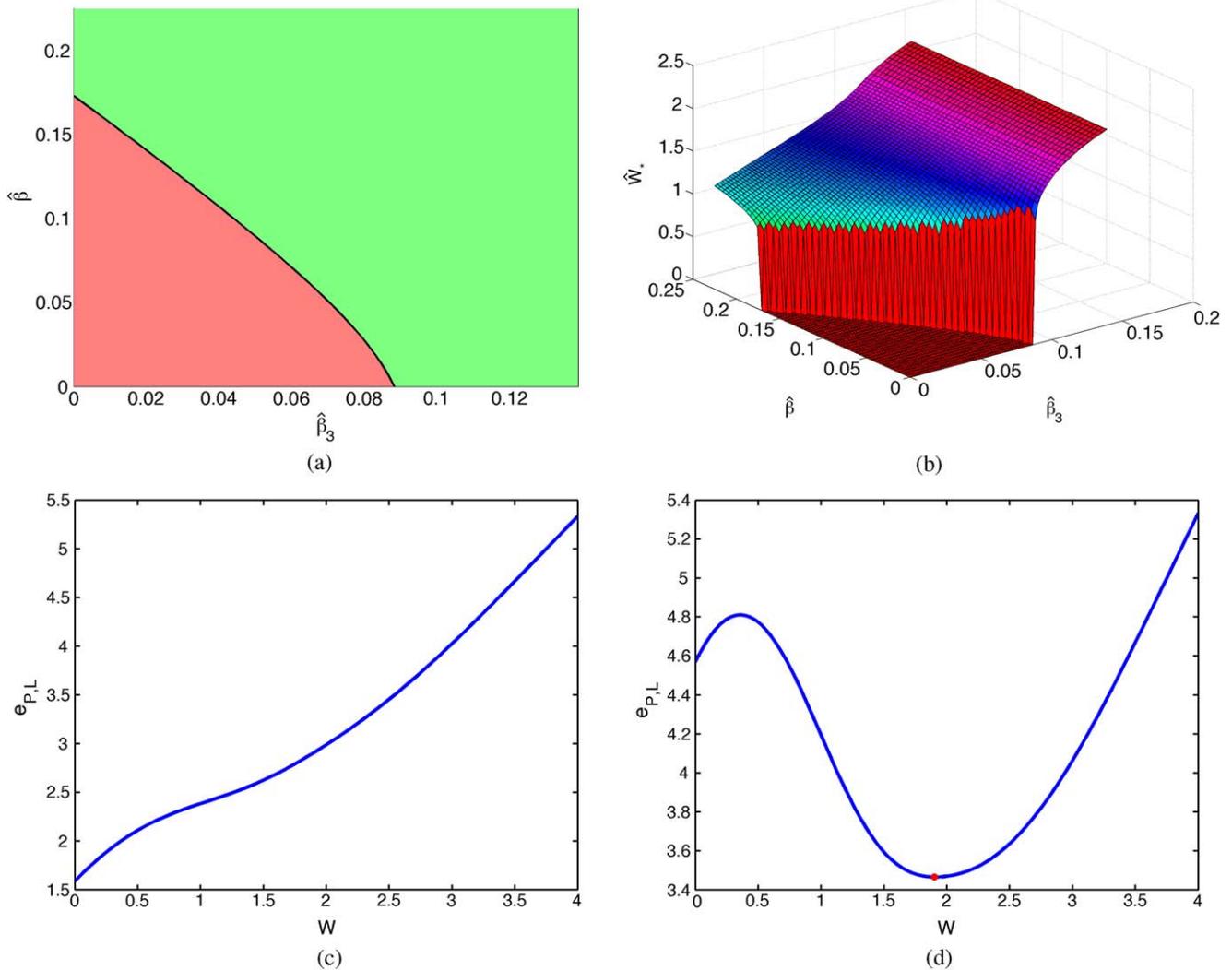


Fig. 4 $\hat{d}_2 = 2 < \hat{D}_2$. (a) Different regions in the $\hat{\beta}_3 - \hat{\beta}$ plane. $e_{P,L}$ has either no local minimum (red) or one local minimum (green). (b) The associated stable bar width \hat{W}_* . (c) $e_{P,L}$ with no local minimum ($\hat{\beta} = 0.05, \hat{\beta}_3 = 0.04$). (d) $e_{P,L}$ with one local minimum ($\hat{\beta} = 0.2, \hat{\beta}_3 = 0.1$)

We find that such a singular point of \hat{d}_2 (named \hat{D}_2) indeed exists, with $\hat{D}_2 \simeq 2.7$. If \hat{d}_2 is less than \hat{D}_2 , at most one local minimum can be found. If $\hat{d}_2 > \hat{D}_2$, there are three cases, depending on the values of $\hat{\beta}, \hat{\beta}_3$ and \hat{d}_2 : $e_{P,L}$ has no local minimum; $e_{P,L}$ 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 $e_{P,L}$ has two local minima, at $\hat{W} \simeq 1$ and $\hat{W} \simeq \hat{d}_2$.

The two regimes are illustrated in Figs. 4 and 5. For both regimes, the associated stable bar width, as well as the possible behaviours of the energy, are shown in the same figure. Figure 4 gives an example for $\hat{d}_2 = 2 < \hat{D}_2$. In the $\hat{\beta}_3 - \hat{\beta}$ plane, the red region and the green region correspond to the situations when $e_{P,L}$ has no local minimum, and one local minimum respectively. The sole separation curve between the two regions corresponds to the situation when the two extrema of $e_{P,L}$ become an inflection point. This curve demonstrates a qualitative change in the behaviour of $e_{P,L}$.

The case for $\hat{d}_2 > \hat{D}_2$ (here $\hat{d}_2 = 5.5$) is shown in Fig. 5. To make things clearer, we label the three curves by numbers; and we name the widths where the two possible local minima are found \hat{W}_{MIN_1} and \hat{W}_{MIN_2} , and the widths where the two possible local maxima are found \hat{W}_{MAX_1} and \hat{W}_{MAX_2} , by the ascending order of their values, so $\hat{W}_{MAX_1} \leq \hat{W}_{MIN_1} \leq \hat{W}_{MAX_2} \leq \hat{W}_{MIN_2}$. (In the case of equality, the extrema of the energy merge as one or more inflection points.) We examine the situation at each curve and at each intersection point between two curves:

- Curve 1: \hat{W}_{MAX_1} and \hat{W}_{MIN_1} merge together and become an inflection point. See Fig. 6(c).
- Curve 2: \hat{W}_{MAX_2} and \hat{W}_{MIN_2} merge together and become an inflection point. See Fig. 6(e).
- Curve 3: \hat{W}_{MIN_1} and \hat{W}_{MAX_2} merge together and become an inflection point. See Fig. 6(g).

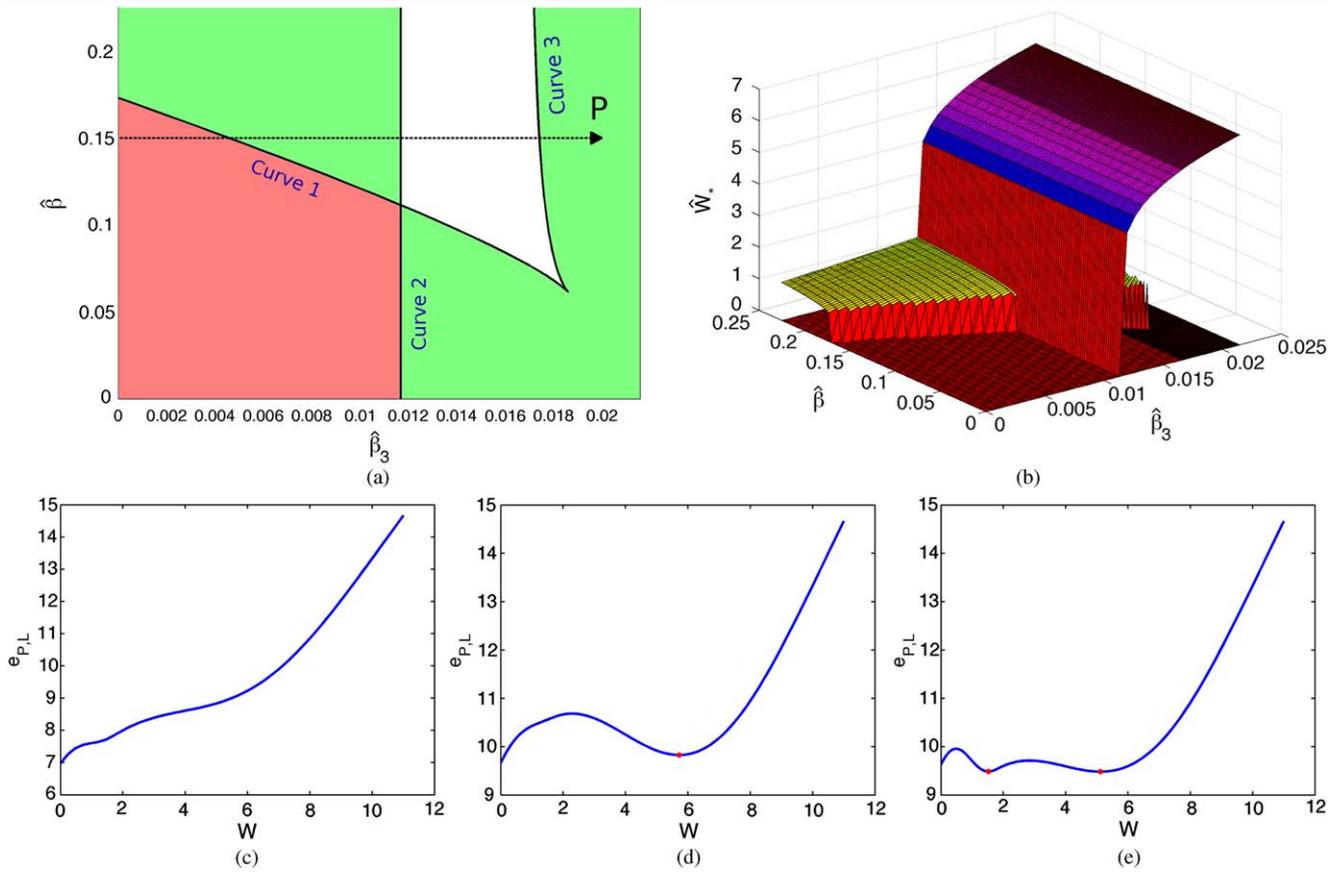


Fig. 5 $\hat{d}_2 = 5.5 > \hat{D}_2$. (a) Different regions in the $\hat{\beta}_3$ - $\hat{\beta}$ plane. $e_{P,L}$ has either no local minimum (red), one local minimum (green), or two local minima (white). (b) The associated stable bar width(s) \hat{W}_* . (c) $e_{P,L}$ with no local minimum ($\hat{\beta} = 0.1, \hat{\beta}_3 = 0.01$). (d) $e_{P,L}$ with one local minimum ($\hat{\beta} = 0.05, \hat{\beta}_3 = 0.015$). (e) $e_{P,L}$ with two local minima ($\hat{\beta} = 0.2, \hat{\beta}_3 = 0.013$)

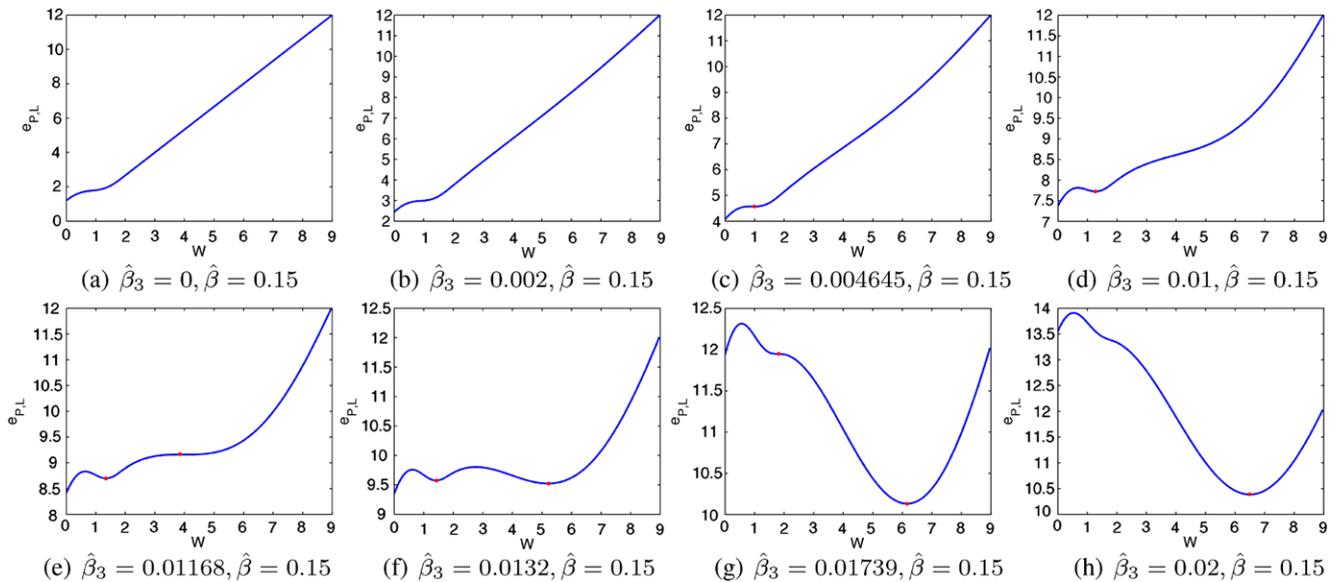


Fig. 6 Evolution of $e_{P,L}$ as the parameter setting is moved from left to right along the dotted path P shown in Fig. 5 ($\hat{d}_2 = 5.5$). Red dot: local minimum

- Intersection point between curve 1 and curve 2: \hat{W}_{MAX_1} merges with \hat{W}_{MIN_1} and also \hat{W}_{MAX_2} merges with \hat{W}_{MIN_2} . Four extrema of the energy merge as two inflection points. See Fig. 7(a).
- Corner point between curve 1 and curve 3: \hat{W}_{MAX_1} , \hat{W}_{MIN_1} and \hat{W}_{MAX_2} merge together. The energy has a very wide maximum (the inflection point of the three extrema) and a local minimum. See Fig. 7(b).
- Corner point between curve 2 and curve 3: the point actually goes off to infinity, which arises from the finite range of the interaction function Ψ . It corresponds to the situation where \hat{W}_{MIN_1} , \hat{W}_{MAX_2} and \hat{W}_{MIN_2} merge together. The energy should have a local maximum and a very wide local minimum (the inflection point of the three extrema). This is exactly what we need for modelling network branches with widely varying widths. We cannot really use this ‘critical point’ with the current interaction function, but the possibility of changing the interaction function in order to benefit from this behaviour is certainly a point for future study.

Thus, each curve is a set of points where two extrema of the energy merge as an inflection point, and where a qualitative change in behaviour happens; each intersection point combines the properties of the intersecting curves.

Let us see how the local minimum/minima of the energy $e_{P,L}$ evolve(s) through the three different states, *i.e.* no local minimum, one local minimum and two local minima, if the parameter setting is moved from left to right along the dotted path P shown in Fig. 5. The sequence of energies $e_{P,L}$ is plotted in Fig. 6. The state starts with no local minimum (Figs. 6(a)–6(b)). When the path P meets curve 1, a first local minimum appears, and the state jumps from no local minimum to one local minimum. Initially, the first local minimum is an inflection point (Fig. 6(c)). The first local minimum gradually becomes deeper, but the state stays in one local minimum (Fig. 6(d)). Eventually, when the path P meets curve 2, a second local minimum, *i.e.* an inflection point, appears (Fig. 6(e)), and the state jumps from one local minimum to two local minima (Fig. 6(f)). When the parameter setting continues to the right and meets curve 3, the first local minimum and the second local maximum merge together, and become an inflection point (Fig. 6(g)). In the end, the first local minimum disappears, and the state jumps back to one local minimum again (Fig. 6(h)). The red dots in Fig. 6 denote the local minima. They are shown in Fig. 5.

The parameters governing energy stability constitute an M -dimensional parameter space C ; the width of the bar is a 1-dimensional state space X ; and the energy e_P is a smooth function on X parameterized by C . In the linear nonlocal HOAC prior model, the dimension M of the parameter space is 3 (there are three parameters $\hat{\beta}$, $\hat{\beta}_3$ and \hat{d}_2 in space C), and so the possible stable states of the 1-dimensional state space should form a *swallowtail catastrophe* (Thom 1975). The

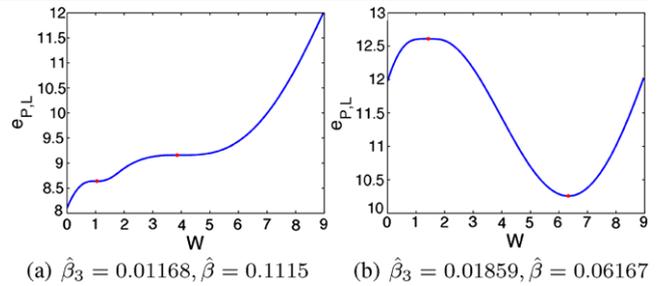


Fig. 7 Graphs of $e_{P,L}$ when the parameter setting is chosen in the two intersection points of curves shown in Fig. 5 ($\hat{d}_2 = 5.5$). Red dot: local minimum

numerical study illustrated by the diagrams in Figs. 4 and 5 confirms this statement.

The variety of possible behaviours 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 Rochery et al. (2006), the new linear 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 potentially model a large range of widths.

5 Overall Model for Linear Network Extraction

In addition to the prior energy E_P ($E_{P,NL}$ or $E_{P,L}$), we also need a likelihood energy term linking the region R (which in our case corresponds to the road network) to the data, in our case a VHR optical satellite image. In this section, we will introduce this term, E_D , and specify some implementation details.

5.1 Overall Energy

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

$$E(\phi; I) = E_D(I, \phi) + \theta E_P(\phi), \tag{27}$$

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

$$E_D(I, \phi) = - \int_{\Omega} dx \{ \phi_+(x) \ln P_+(I(x)) + \phi_-(x) \ln P_-(I(x)) \}. \tag{28}$$

$P_{\pm}(I)$ 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

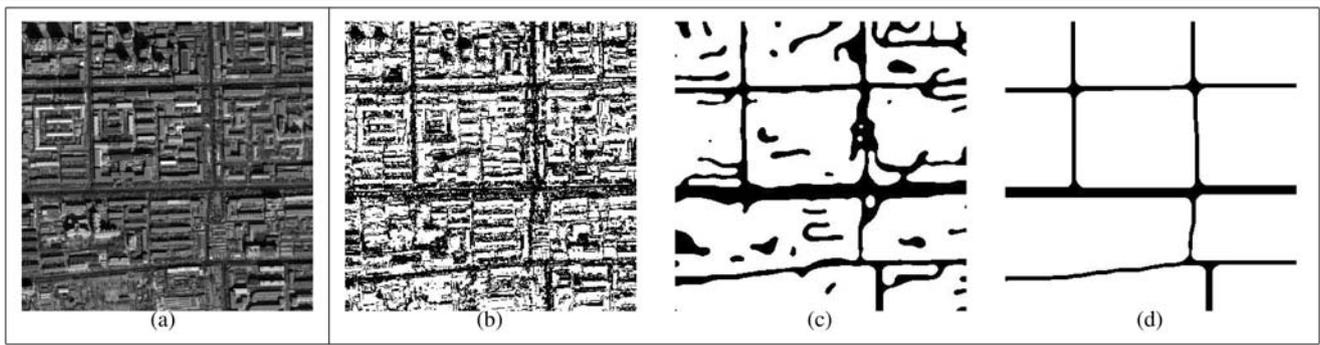


Fig. 8 Data, and experiment at 1/4 resolution. (a) Image data (size = 350×350 , road width = $3 \sim 5$ pixels). (b)–(d) Results obtained using the new nonlinear model (with E_{NL}) at iterations 1, 1,500 and 27,000

weak, in the sense that maximum likelihood classification produces very poor results (see Figs. 9(b) and 11(d)), mainly due to the ‘noise’ in the road region and the great variations in the background. No image model that only takes into account the radiometry of independent pixels can do much better than this, which is why a powerful geometric prior model is needed.

5.2 Optimization and Parameter Settings

To minimize E , we perform gradient descent with the neutral initialization: the initial value of ϕ is set equal to the threshold $z = \alpha/\lambda$ everywhere in Ω (Rochery et al. 2005). The algorithm is thus quasi-automatic. The functional derivatives of the HOAC terms $\delta E_S/\delta\phi$, $\delta E_{NL}/\delta\phi$ and $\delta E_L/\delta\phi$ involve convolutions: they are calculated in the Fourier domain, as are all derivatives. The evolution equation for E with E_{NL} is

$$\begin{aligned} \frac{\partial\phi(x)}{\partial t} = & \frac{1}{2} \ln \frac{P_+}{P_-} + \theta \left\{ \nabla^2\phi(x) - \lambda(\phi^3(x) - \phi(x)) \right. \\ & - \alpha(1 - \phi^2(x)) + \beta \mathcal{F}^{-1} \{ k^2 d \hat{\Psi}(kd) \hat{\phi}(k) \} \\ & - \beta_2 \nabla^2\phi(x) \mathcal{F}^{-1} \{ d \hat{\Psi}(kd) \mathcal{F} \{ \nabla\phi(x) \cdot \nabla\phi(x) \} \} \\ & - \beta_2 \nabla\phi(x) \cdot \nabla \{ \mathcal{F}^{-1} \{ d \hat{\Psi}(kd) \\ & \left. \times \mathcal{F} \{ \nabla\phi(x) \cdot \nabla\phi(x) \} \} \right\}, \end{aligned} \quad (29)$$

where \mathcal{F} and \mathcal{F}^{-1} denote the Fourier and the inverse Fourier transform respectively, and a hat $\hat{\cdot}$ indicates the Fourier transform of a function. The evolution equation for E with E_L involves replacing the last two lines of (29) by

$$\beta_3 \mathcal{F}^{-1} \left\{ k \cdot \mathcal{F} \left\{ \Psi \left(\frac{|x|}{d_2} \right) \epsilon_x x^T \epsilon^T \right\} \cdot k \hat{\phi}(k) \right\}. \quad (30)$$

The time evolution of ϕ uses the forward Euler method. The parameters of the two prior energies $E_{P,NL}$ and $E_{P,L}$ are constrained by the stability analysis of Sect. 4.

6 Experimental Results

As input data I , we use a number of images, with average size 1200×1200 pixels, extracted from a QuickBird optical panchromatic image of Beijing. The scenes are characteristic of dense urban regions. Our aim is to extract, completely and accurately, the road network from an image. We also analyse the effect of the different terms in our energies. In order to evaluate the performance of our new nonlinear and linear models, we compare them quantitatively to ground truth and to other methods from the literature. Ground truth was created by slightly correcting by hand an available GIS map from a few years earlier of the road network in the zone shown in the image. Note that this ground truth *defines* what we mean by the phrase ‘road network’.

6.1 Nonlinear Nonlocal Overall Model

In this subsection, we demonstrate the behaviour of our new model $E = \theta(E_0 + E_S + E_{NL}) + E_D$ via experiments at reduced resolutions.

Figure 8(a) shows one of the input images at 1/4 resolution. The parameters $(\theta, \alpha, \lambda, \beta, \beta_2, d)$ are $(100, 0.12, 3.8, 0.0375, 0.0338, 4)$. The results obtained using the energy with the new nonlinear nonlocal term E_{NL} at iterations 1, 1,500 and 27,000 are illustrated in Figs. 8(b)–8(d). The result obtained using the standard HOAC model without E_{NL} (*i.e.* $\beta_2 = 0$) is shown in Fig. 9(a). We see that adding E_{NL} enables the recovery of the main and secondary road network, whereas the model without E_{NL} misses a secondary road. In order to illustrate the effects of other terms in the model, we compute results using maximum likelihood estimation (MLE, *i.e.* $\theta = 0$) and a standard, non-higher-order active contour (*i.e.* $\beta = \beta_2 = 0$) (see Figs. 9(b) and 9(c)). The ground truth is presented in Fig. 9(d). The MLE result shows that local image information alone is not sufficient to distinguish the roads from the background, while the standard active contour result shows the importance of the geo-

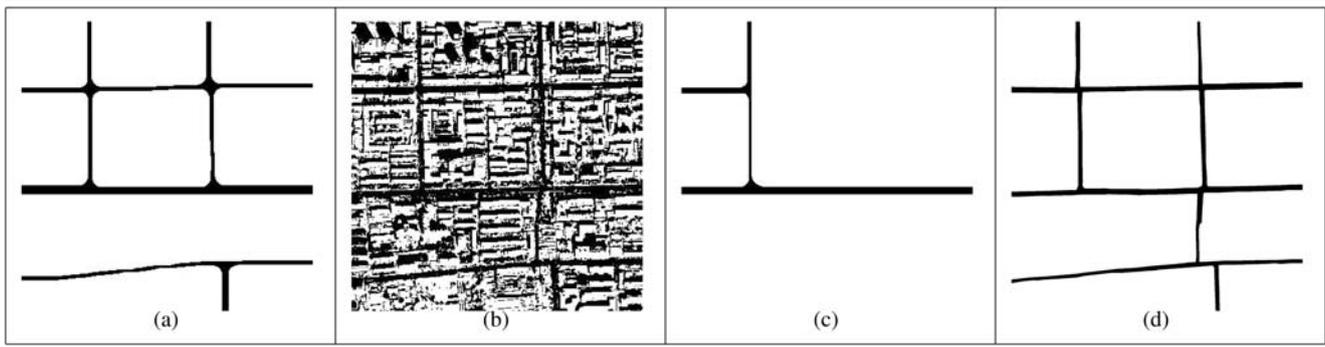


Fig. 9 Experiments at 1/4 resolution and ground truth. (a)–(c) Results obtained respectively using the standard energy without E_{NL} , MLE, and a standard, non-higher-order active contour model (with neither E_S nor E_{NL}). (d) Ground truth segmented manually

Table 1 Quantitative evaluation criteria for the different methods as tested on Fig. 8(a) at 1/4 resolution (T = True, F = False, P = Positive, N = Negative). The completeness is the percentage of ground truth road network that is extracted; the correctness is the percentage of

extracted road network that is correct; and the quality is the most important measure of the ‘goodness’ of the result, because it takes into account the completeness and the correctness

Method	Measure	Completeness TP/(TP + FN)	Correctness TP/(TP + FP)	Quality TP/(TP + FP + FN)
New model E with E_{NL} (Fig. 8(d))		0.9524	0.8591	0.8237
$\theta(E_0 + E_S) + E_D$ (Fig. 9(a))		0.8832	0.8659	0.7769
$\theta E_0 + E_D$ (Fig. 9(c))		0.4282	0.8314	0.3940
MLE ($\theta = 0$) (Fig. 9(b))		0.9734	0.1831	0.1822

metric knowledge introduced by HOACs. Quantitative evaluations based on standard criteria (Heipke et al. 1997) are shown in Table 1.

Figure 10 presents more results at reduced resolutions. The first column shows the input image data, which is either at 1/4 or 1/2 resolution. The two columns on the right show the corresponding results obtained with and without the new nonlinear, nonlocal term E_{NL} . The importance of E_{NL} is clear: it facilitates greatly the retrieval of secondary roads. However the nonlinear nature of the energy means that it is prohibitively expensive computationally at full resolution. In the next subsection, we show results at both reduced and full resolution using the linear nonlocal model.

6.2 Linear Nonlocal Overall Model

As in the previous subsection, we study primarily the extraction of a network consisting of roads of roughly the same width, but in Sect. 6.2.2, we briefly consider the extraction of networks containing roads of two very different widths. In the former case, we choose the parameters so that $e_{P,L}$ 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 $e_{P,L}$ has two local minima. Again a small range of widths around each minimum is possible.

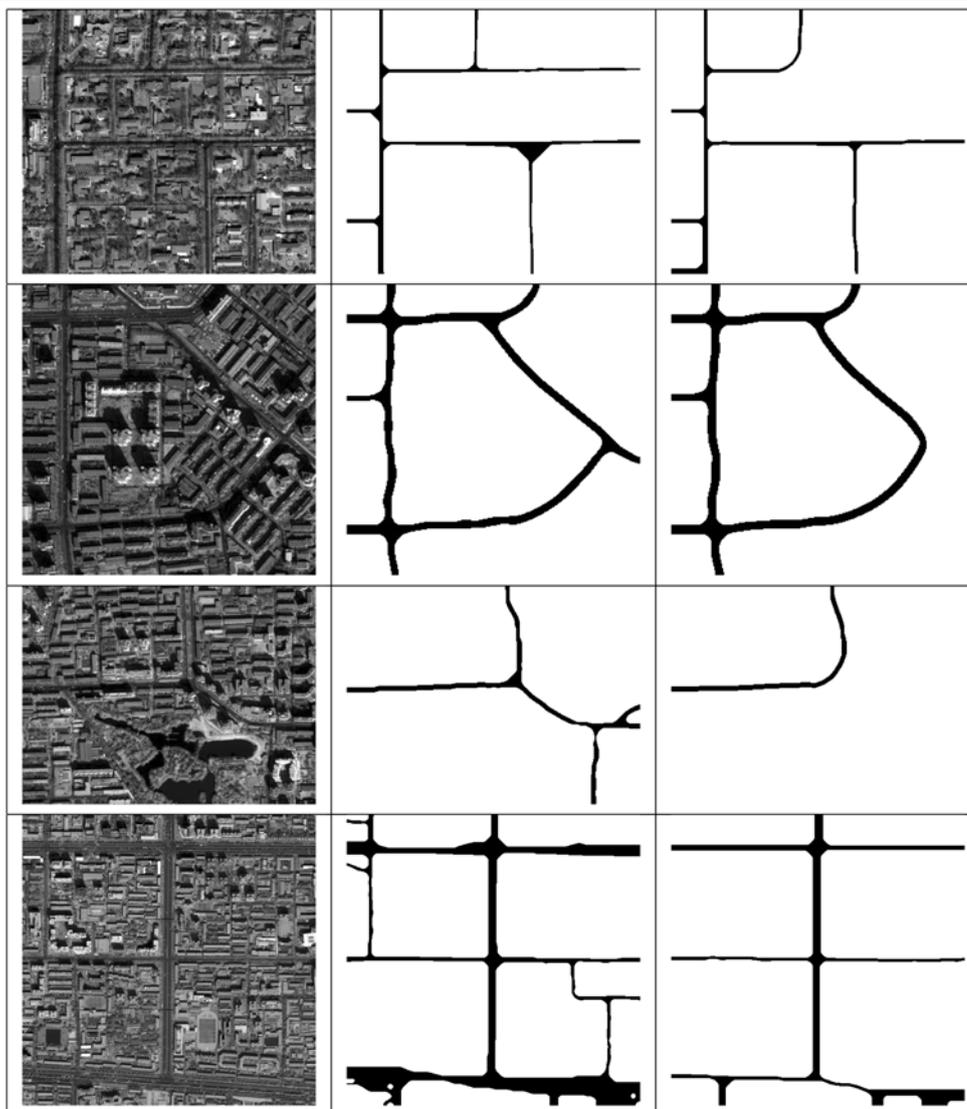
6.2.1 Extraction of Roads of Similar Width

We apply our new linear model $E = \theta(E_0 + E_S + E_L) + E_D$ to both full-resolution and reduced resolution images. We fix the parameters as described in Sect. 4.3. For all experiments, the parameters $(\theta, \alpha, \lambda, \beta, \beta_3, d, d_2)$ are $(200, 0.15, 4, 0.02, 2 \times 10^{-4}, 4, 12)$ 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_3 , the other parameters are the same for the two resolutions.

The results obtained using the new linear model E (with E_L), at 1/4 resolution and at full resolution, are shown in Figs. 11(b) and 11(c). The complete road network is retrieved successfully at both resolutions. Although the segmentation at 1/4 resolution appears geometrically 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 linear model, we now compare our result to other methods at full resolution: with MLE (*i.e.* $\theta = 0$); with a standard, non-higher-order active

Fig. 10 More experiments at reduced resolutions. *First column:* input images, *first row:* 1/2 resolution, size = 400×440 , road width = $2 \sim 4$ pixels; *second row:* 1/4 resolution, size = 300×300 , road width = $3 \sim 5$ pixels; *third row:* 1/4 resolution, size = 300×400 , road width = $3 \sim 6$ pixels; *last row:* 1/4 resolution, size = 512×512 , road width = $3 \sim 15$ pixels. *Two rightmost columns:* corresponding results obtained using the new nonlinear model (with E_{NL}) and the standard model (without E_{NL})



contour (i.e. $\beta = \beta_3 = 0$); and with the standard model without E_L (i.e. $\beta_3 = 0$). The results are shown in Figs. 11(d)–11(f). Again MLE discriminates poorly between the roads and the background, while the models without E_L and/or E_S 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, as described in Wang and Zhang (2003) and Yu et al. (2004), and compare them to ours (see Fig. 12). Wang and Zhang (2003) presented a classification, tracking, and morphology algorithm to extract urban road networks from QuickBird images; Yu et al. (2004) proposed a fast but rough segmentation technique based on ‘straight line density’ to extract urban road networks. However, without much prior geometric knowledge, both of them extract many incorrect areas that happen to have statistical properties similar to roads. Moreover, the accuracy of the delin-

ation of the road boundary is poor. Some quantitative evaluations based on standard criteria (Heipke et al. 1997), are shown in Table 2. The ‘quality’ is the most important measure because it considers both completeness and correctness. Our complete model outperforms all others.

Figure 13 presents more results using the new linear model (with E_L), at full resolution. We also apply the same model to river extraction, as shown in Fig. 14.

6.2.2 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. Figure 15(a) shows an input image containing two roads: their widths are approximately 20 pixels and 80 pixels. The results obtained using the new linear model (with E_L) and the standard model (without E_L), at 1/4 resolution,

Fig. 11 Experiments using the new linear model (with E_L), and analysis of the effects of the different terms in the energy.

(a) Image data (size: 1400×1400); (b) result obtained using the new linear model (with E_L), at 1/4 resolution; (c) result obtained using the new linear model (with E_L), at full resolution; and results obtained at full resolution using (d) MLE; (e) the model with $\beta = \beta_3 = 0$ (equivalent to a standard active contour); (f) the standard model without E_L (i.e. $\beta_3 = 0$)

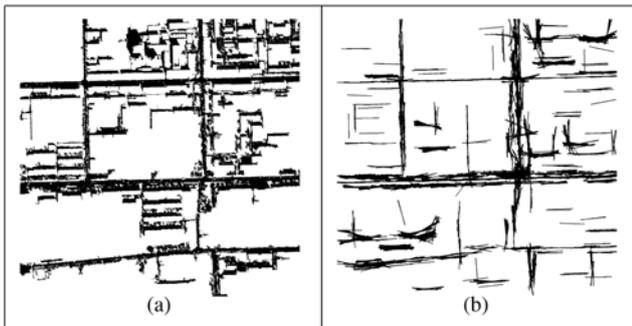
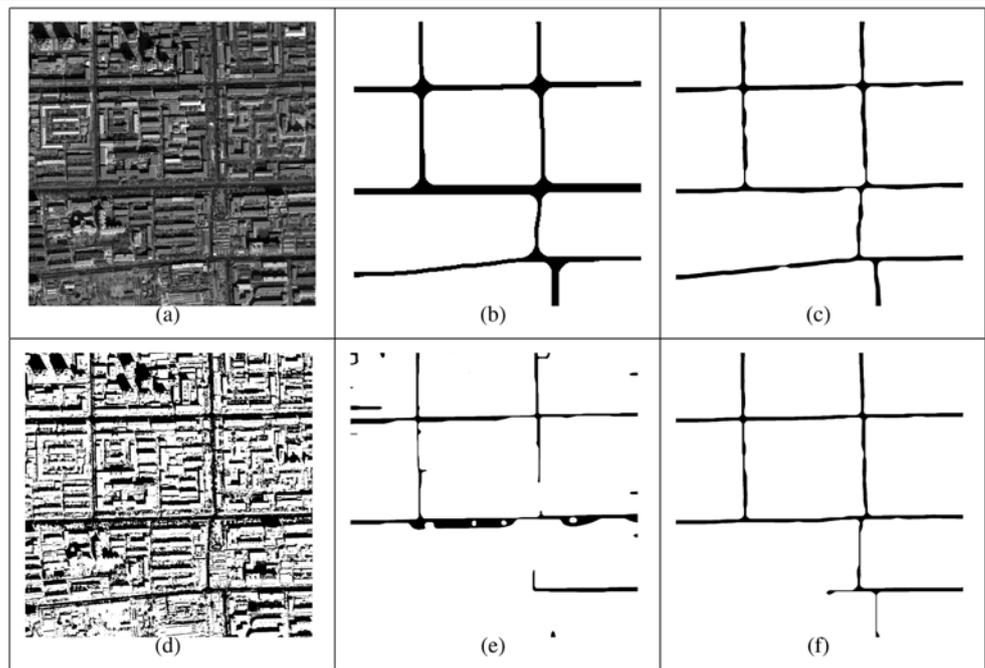


Fig. 12 Comparisons at full resolution. (a, b) Results obtained with Wang and Zhang (2003) and Yu et al. (2004)

are illustrated in Figs. 15(b) and 15(c) respectively. The parameter values ($\theta, \alpha, \lambda, \beta, \beta_3, d, d_2$) used in this experiment are $(25, 0.15, 5, 0.02, 1.228 \times 10^{-4}, 4, 22)$. The estimated stable widths for these parameter 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 standard model without E_L finds only an incomplete network.

6.3 Discussion

Computational cost The computational costs of the two proposed algorithms depend on two factors: the number of iterations and the number of operations at each iteration. The number of iterations depends greatly on the complexity of the image, and on the values of the parameters, even when the final result does not. On the other hand, the number of

operations at each iteration is related to the functional derivative of the energy. As already mentioned, we compute the functional derivative in the Fourier domain. At each iterative step, for the nonlinear model, we need to compute two forward Fourier transforms and two inverse Fourier transforms; while for the linear model, only one forward and one inverse transform are required, the rest of the operations being the same. The computation time for the result in Fig. 8, obtained with E_{NL} , is around 80 minutes, while that for the result in Fig. 11(b), obtained with E_L , is around 60 minutes. These computation times are long of course, but to our knowledge no other method even approaches the quality of the results obtained.

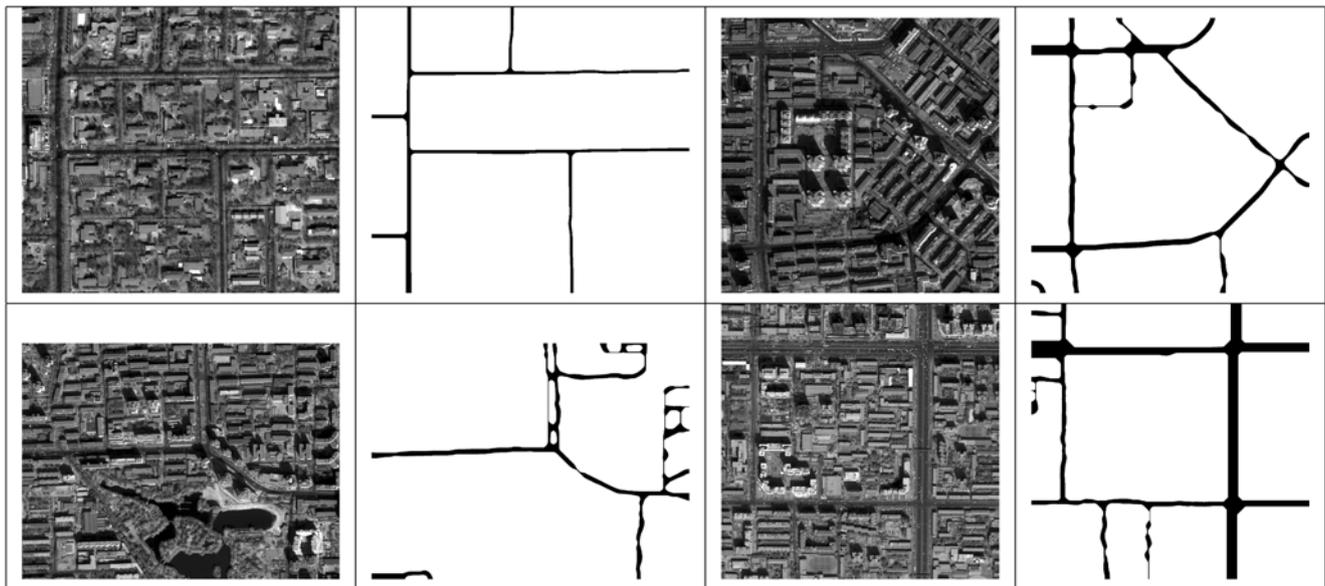
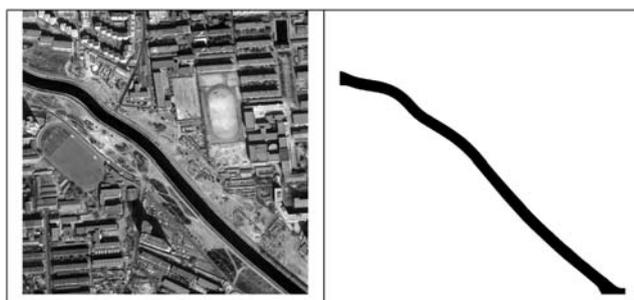
Parameter choice We note that in practice, the results obtained are not sensitive to the precise choice of parameter values, provided they lie in the correct subset of the $\hat{\beta} - \hat{\beta}_3 - \hat{d}_2$ diagram.

7 Conclusion

Using the phase field HOAC framework, we have presented, in this paper, two novel nonlocal HOAC energies for modelling network-shaped regions. Both of them enable the generation of longer, straighter branches with better prolongation, but by different means. The first, nonlinear energy causes pairs of points inside the range of the interaction to attract each other. In conjunction with the standard HOAC geometric term, it allows the interaction between points on the same side of a network branch to be stronger than the

Table 2 Quantitative criteria for the experiments shown in Fig. 11(a) at full resolution (except first row) (T = True, F = False, P = Positive, N = Negative). See Table 1 for an explanation of completeness, correctness and quality

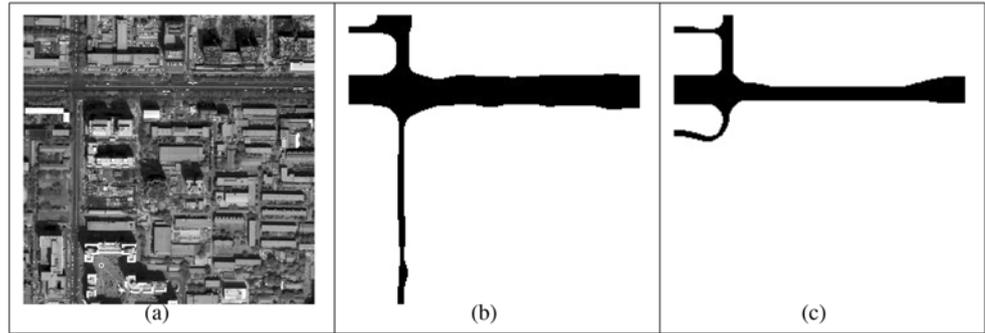
Method	Measure	Completeness	Correctness	Quality
		TP/(TP + FN)	TP/(TP + FP)	TP/(TP + FP + FN)
New model E (with E_L) at 1/4 resolution (Fig. 11(b))		0.9688	0.8519	0.8292
New model E (with E_L) at full resolution (Fig. 11(c))		0.8756	0.9693	0.8520
MLE (Fig. 11(d))		0.9356	0.2073	0.2044
$\theta E_0 + E_D$ (Fig. 11(e))		0.6047	0.8249	0.5359
$\theta(E_0 + E_S) + E_D$ (Fig. 11(f))		0.6946	0.9889	0.6892
Wang (Wang and Zhang 2003) (Fig. 12(a))		0.9350	0.3463	0.3381
Yu (Yu et al. 2004) (Fig. 12(b))		0.6050	0.3695	0.2977

**Fig. 13** More results using the new linear model (with E_L) on pieces of a QuickBird image at full resolution. Image size: *top left*: 800×880 ; *top right*: 1200×1200 ; *bottom left*: 1200×1600 ; *bottom right*: 1400×1400 **Fig. 14** Result of river extraction on a QuickBird panchromatic image (size: 1024×1024)

interaction between points on opposite sides of a network branch. The second linear energy, which includes a longer-range interaction, controls the prolongation of the network branch, but has little effect on the branch width. The standard HOAC geometric term still controls the branch width

with a (relatively) short-range interaction. Based on a stability analysis of a bar with a desired width, we established constraints on the parameters of the energy function. We explored the possible behaviours of each of the two resulting prior energies as a function of the parameter settings. We showed that as well as separating the interactions between points on the same and opposite sides of a network branch, the new linear model permits the modelling of two widths simultaneously. The analysis also fixes some of the model parameters in terms of network width(s). Moreover, the linear nonlocal term is more efficient from a computational point of view, and can thus be applied to images at full resolution. As a consequence, narrower network branches can be extracted than is possible at reduced resolution, and in general, extraction accuracy is improved. Experiments also demonstrate the superiority of our new models 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

Fig. 15 Extraction of a road network containing two different widths, at 1/4 resolution. (a–c) Image data; results using: the new linear model E ; the standard model ($\beta_3 = 0$)



wide range, instead of two sharp local minima. This might be a better solution for the extraction of roads with multiple widths. We note finally that although in this paper we have focused on reducing curvature, one could use the flexibility of the new models to allow more curvature than the model of Rochery et al. (2006), while still allowing a wide range of branch widths.

Acknowledgements This work was partially supported by European Union Network of Excellence MUSCLE (FP6-507752) and by INRIA Associated Team ‘SHAPES’. This work was also supported by the 863 program of the Ministry of Science and Technology of China. The work of the first author was supported by an MAE/Thales Alenia Space/LIAMA grant during her PhD.

References

- Amo, M., Martínez, F., & Torre, M. (2006). Road extraction from aerial images using a region competition algorithm. *IEEE Transactions on Image Processing*, 15(5), 1192–1201.
- Bertozzi, A., Esedoglu, S., & Gillette, A. (2007). Inpainting of binary images using the Cahn-Hilliard equation. *IEEE Transactions on Image Processing*, 16(1), 285–291.
- Chen, Y., Tagare, H., Thiruvankadam, S., Huang, F., Wilson, D., Gopinath, K., Briggs, R., & Geiser, E. (2002). Using prior shapes in geometric active contours in a variational framework. *International Journal of Computer Vision*, 50(3), 315–328.
- Cremers, D., Tschhäuser, F., Weickert, J., & Schnörr, C. (2002). Diffusion snakes: Introducing statistical shape knowledge into the Mumford-Shah functional. *International Journal of Computer Vision*, 50(3), 295–313.
- Cremers, D., Osher, S., & Soatto, S. (2006). Kernel density estimation and intrinsic alignment for shape priors in level set segmentation. *International Journal of Computer Vision*, 69(3), 335–351.
- Dobrosotskaya, J. A., & Bertozzi, A. L. (2008). A wavelet-Laplace variational technique for image deconvolution and inpainting. *IEEE Transactions on Image Processing*, 17(5), 657–663.
- Geman, S., & Geman, D. (1984). Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6), 721–741.
- Geman, D., & Jedynek, B. (1996). An active testing model for tracking roads from satellite images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(1), 1–14.
- Heipke, C., Mayr, H., Wiedemann, C., & Jamet, O. (1997). Evaluation of automatic road extraction. *International Archives of Photogrammetry and Remote Sensing*, XXXII, 47–56.
- Hu, J., Razdan, A., Femiani, J. C., Cui, M., & Wonka, P. (2007). Road network extraction and intersection detection from aerial images by tracking road footprints. *IEEE Transactions on Geoscience and Remote Sensing*, 45(12), 4144–4157.
- Ising, E. (1925). Beitrag zur theorie des ferromagnetismus. *Zeitschrift für Physik*, 31, 253–258.
- Kass, M., Witkin, A., & Terzopoulos, D. (1988). Snakes: Active contour models. *International Journal of Computer Vision*, 1(4), 321–331.
- Lacoste, C., Descombes, X., & Zerubia, J. (2005). Point processes for unsupervised line network extraction in remote sensing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(10), 1568–1579.
- Leventon, M. E., Grimson, W. E. L., & Faugeras, O. (2000). Statistical shape influence in geodesic active contours. In *Proc. IEEE conference on computer vision and pattern recognition*, Hilton Head Island, South Carolina, USA.
- Mena, J. B. (2003). State of the art on automatic road extraction for GIS update: A novel classification. *Pattern Recognition Letters*, 24(16), 3037–3058.
- Merlet, N., & Zerubia, J. (1996). New prospects in line detection by dynamic programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(4), 426–431.
- Peng, T., Jermyn, I. H., Prinet, V., & Zerubia, J. (2008a). An extended phase field higher-order active contour model for networks and its application to road network extraction from very high resolution satellite image. In *Proc. European conference on computer vision*, Marseille, France.
- Peng, T., Jermyn, I. H., Prinet, V., & Zerubia, J. (2008b). Extraction of main and secondary roads in VHR images using a higher-order phase field model. In *Proc. XXI ISPRS congress, commission III, Part A*, Beijing, China.
- Péteri, R., & Ranchin, T. (2003). Detection and extraction of road networks from high resolution satellite images. In *Proc. international conference on image processing*, Barcelona, Spain.
- Riklin-Raviv, T., Kiryati, N., & Sochen, N. (2007). Prior-based segmentation and shape registration in the presence of perspective distortion. *International Journal of Computer Vision*, 72(3), 309–328.
- Rochery, M., Jermyn, I. H., & Zerubia, J. (2005). Phase field models and higher-order active contours. In *Proc. IEEE international conference on computer vision*, Beijing, China.
- Rochery, M., Jermyn, I. H., & Zerubia, J. (2006). Higher-order active contours. *International Journal of Computer Vision*, 69(1), 27–42.
- Rousson, M., & Paragios, N. (2007). Prior knowledge, level set representations & visual grouping. *International Journal of Computer Vision*, 76(3), 1573–1405.

- Srivastava, A., Joshi, S., Mio, W., & Liu, X. (2003). Statistical shape analysis: Clustering, learning, and testing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(4), 590–602.
- Thom, R. (1975). *Structural stability and morphogenesis*. Elmsford: Benjamin/Addison-Wesley.
- Wang, R., & Zhang, Y. (2003). Extraction of urban road network using Quickbird pan-sharpened multispectral and panchromatic imagery by performing edge-aided post-classification. In *Proc. international society for photogrammetry and remote sensing (ISPRS)*, Quebec City, Canada.
- Yu, Z., Prinnet, V., Pan, C., & Chen, P. (2004). A novel two-steps strategy for automatic GIS-image registration. In *Proc. international conference on image processing*, Singapore.