Incorporating Generic and Specific Prior Knowledge in a Multi-Scale Phase Field Model for Road Extraction from VHR Images

Ting Peng, Ian H. Jermyn, *Member, IEEE,* Véronique Prinet, and Josiane Zerubia, *Fellow, IEEE,*

***Abstract*—This paper addresses the problem of updating road digital maps in dense urban areas by extracting the main road network from very high resolution (VHR) satellite images. Building on the work of Rochery *et al.* (2005), we represent the road region as a 'phase field'. In order to overcome the difficulties due to the complexity of the information observed in VHR images, we first propose a multi-scale statistical data model. It enables the integration of segmentation results from coarse resolution, which furnishes a simplified representation of the data, and fine resolution, which provides accurate details. Moreover, an outdated GIS digital map is introduced into the model, serving as the *specific* prior knowledge of the road network. This new penalty term balances the effect of the *generic* prior knowledge describing the geometric shape of road networks (i.e. elongated and of low-curvature) and carried by a 'phase field HOAC' term. Promising results on QuickBird panchromatic images and comparisons with several other methods demonstrate the effectiveness of our approach.**

***Index Terms*—Dense urban areas, Geographical Information Systems (GIS), multi-scale analysis, road network extraction, variational models, Very High Resolution (VHR) image.**

I. Introduction

KEEPING the road network information contained in Geographical Information Systems (GIS) up to date is crucial for many applications, for example urban planning, vehicle navigation, and environmental monitoring. The high rate of urban growth, especially in many developing countries, means that this has become an increasingly important research topic in remote sensing. Fig. 1 shows two pairs of QuickBird panchromatic images of Beijing, retrieved respectively in the year 2002 and the year 2006, showing the great changes in the past few years. Very high resolution (VHR) optical satellite images (*e.g.* QuickBird and Ikonos, and Pléiades in the near future), with sub-metric resolutions, provide new opportunities for the extraction of information from remote sensing data: qualitatively new categories of information are available, and the accuracy of previously extracted categories of information can be quantitatively improved. Higher resolution brings with it new challenges however. Details invisible in lower resolution images, *e.g.* cars, road markings, shadows, and other linear but non-road features, can easily disrupt the recognition process, and demand more sophisticated modelling, both of the image and of the road network. For the former, the existence of phenomena at multiple scales suggests a multi-scale approach, while for the latter, the incorporation into the models of our prior knowledge of the geometry of the road network becomes critical.

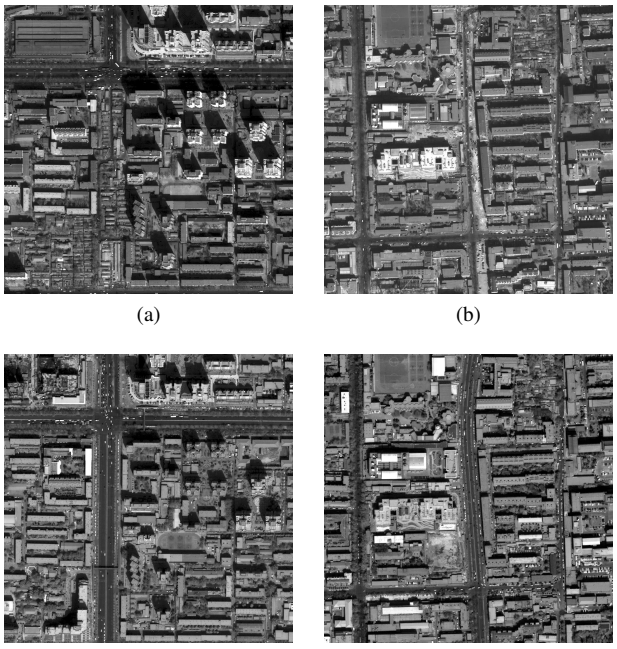


Fig. 1. Two pairs of QuickBird panchromatic images 0.61 m/pixel (both size: 1000×1000) of Beijing, showing the great changes in the past few years. Top: year 2002; bottom: year 2006.

In this paper, we address the issue of road network updating from VHR panchromatic images in dense urban areas. Specifically, we will show how to make use of an outdated GIS digital map and a recently acquired QuickBird image to generate an up-to-date road network of the observed region. We model the region corresponding to the road network using a 'phase field' [1], which offers a number of important advantages over other region modelling frameworks. Building

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

T. Peng is with Project-Team Ariana, INRIA/I3S, 06902 Sophia Antipolis, France, and also with LIAMA & NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing 100080, China (email: tpeng@sophia.inria.fr). Corresponding author. T: +33 492387595. F: +33 492387643.

I. H. Jermyn and J. Zerubia are with Project-Team Ariana, INRIA/I3S, 06902 Sophia Antipolis, France (email: firstname.lastname@sophia.inria.fr).

V. Prinet is with LIAMA & NLPR, Institute of Automation, Chinese Academy of Sciences, Beijing 100080, China (email: prinet@nlpr.ia.ac.cn).

upon [1], our contribution is twofold: first, we propose a multi-scale framework for road extraction; it is based on a wavelet decomposition of the image and enables an accurate extraction of the road region; second, we introduce a *specific* prior term into the energy functional; we will show how this *specific* prior term can be derived from an existing GIS digital map.

An exhaustive review of the work on road extraction, multi-scale approaches, and active contour models is out of the scope of this paper. Here we briefly review those we believe to be the most relevant. In [2], Mayer *et al.* first extract lines at coarse resolution, which in turn are used to initialize ribbon snakes at fine resolution. Péteri and Ranchin [3] take advantage of a topologically correct graph of the network in order to extract roads and junctions using two different types of active contours. Fortier *et al.* [4] initialize an active contour model from an existing GIS map and junctions previously detected from the image. Agouris *et al.* [5] compute uncertainty factors (positional uncertainty at each contour point and global shape uncertainty) to control the active contour. However, all these methods are mostly restricted to applications on rural or semi-urban areas and using aerial images. They are not robust enough to be applied to dense urban areas.

The rest of the paper is organised as follows. In section II, we introduce the essentials of the basic phase field model, and then describe our multi-scale data model. In section III, we introduce a new GIS prior energy. In section IV, we detail the optimization algorithm. In section V, we illustrate experimental results on QuickBird panchromatic images and perform validation and comparison with other techniques. We conclude in section VI.

II. The Model: Prior and Data Energies

In this section, we first recall the general Bayesian formulation for image segmentation and show how to convert the maximization of a posterior probability into the minimization of the corresponding energy. We will then detail the terms of the energy function of our model.

Given an image $I : \Omega \to \mathbb{R}$, and given the prior knowledge K we may have, our goal is to find the region R in the image domain Ω that corresponds to the main road network. Using Bayesian decomposition, the problem of segmenting the image I can then be tackled by maximizing the posterior probability

$$P(R|I,K) = \frac{P(I|R,K)P(R|K)}{P(I|K)} \ , \tag{1}$$

with respect to the estimated region R. The denominator in (1) is independent on R and thus does not intervene during the maximization process. Equivalently, one can minimize a total energy functional defined by

$$E(R;I) = \theta E_P(R) + E_D(I,R) \ , \tag{2}$$

where θ is a constant that balances the contribution of the prior energy E_P and the data energy E_D.

To model the region of interest R, we use a *phase field* function ϕ, much used in physics and first introduced to image processing by [1]. Later we will use the symbol ϕ to denote the region R.

A. Prior Energy

We use the same prior energy E_P as [1]. It is composed of two terms: a basic phase field energy $E_{P,0}$ and a higher-order active contour phase field energy $E_{P,NL}$.

Conventional active contours [2], [6]–[9] are defined by linear functionals, so they can incorporate only weak prior knowledge of region geometry. In contrast, HOACs [10] are defined by polynomial functionals. Via long-range interactions between points in the region boundary ∂R, HOACs allow the inclusion of complex prior geometrical constraints. For this reason, HOACs are more robust to noise than conventional active contours, and permit a generic initialization that renders them more automatic.

Phase fields [1] model the region R using a level set function ϕ defined over the entire image. They have several advantages over parametric active contours or standard level sets: a linear representation space; ease of implementation; and a neutral initialization. In addition, they allow greater topological freedom, which is critical when the topology of the region is not known *a priori*. In the present application of road network extraction, dealing with the topological complexity of roads is arguably one of the most difficult aspects; phase fields handle it 'naturally'. *Phase field HOACs* are phase field models that also include the long-range interactions characteristic of HOACs.

To model R with the phase field function, the basic phase field energy $E_{P,0}$ is given by

$$E_{P,0}(\phi) = \int_\Omega dx \left\{ \frac{1}{2}\nabla\phi(x)\cdot\nabla\phi(x) + W(\phi(x)) \right\} . \tag{3}$$

$\nabla\phi(x)$ is gradient vector at pixel x. The potential W is

$$W(z) = \lambda(\frac{1}{4}z^4 - \frac{1}{2}z^2) + \alpha(z - \frac{1}{3}z^3) \ , \tag{4}$$

where λ and α are constants. For $\lambda \geqslant \alpha > 0$, W has two minima, at $z = -1$ and $z = 1$, and a maximum at $z = \alpha/\lambda$.

By definition, $R = \{x \in \Omega : \phi(x) > \alpha/\lambda\}$, but in addition the potential W effectively constrains $\phi(x) \simeq 1$ for $x \in R$ and $\phi(x) \simeq -1$ for $x \in \bar{R} = \Omega \setminus R$. As a result, the quantities $\phi_\pm = (1 \pm \phi)/2$ are approximately equal to the characteristic functions of R and $\bar{R}$. The local derivative product $\nabla\phi(x)\cdot\nabla\phi(x)$ penalizes large gradients and ensures that ϕ makes a smooth transition from -1 to 1 across the boundary ∂R. [1] has shown that $E_{P,0}$ is equivalent to a linear active contour model with an energy $\lambda_C L(\partial R) + \alpha_C A(R)$, where L is region boundary length and A region area, and λ_C and α_C are constants. Therefore, $E_{P,0}$ ensures stability, boundary smoothness, and the properties of the functions $\phi_\pm$.

We introduce sophisticated geometric constraints into the model via a higher-order energy term $E_{P,NL}$. $E_{P,NL}$ describes long-range interactions between the gradient vectors of ϕ at pairs of points separated by many pixels. It is defined as

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

where d controls the range of the interaction. The interaction function Ψ, is given by

$$\Psi(x) = \begin{cases} \frac{1}{2}\left(2 - |x| + \frac{1}{\pi}\sin(\pi|x|)\right) & \text{if } |x| < 2\ , \\ 0 & \text{else .} \end{cases} \quad (6)$$

$E_{P,NL}$ has two effects: it prevents pairs of points with anti-parallel normal vectors from coming too close; and it encourages pairs of points with parallel normal vectors to attract each other, and thus the growth of armlike structures. Consequently, the effect is to assign low energy to (and hence favour) regions composed of long arms of a certain width and with roughly parallel sides that join together at junctions.

B. Data Energy

The data energy E_D takes into account the following radiometric properties of the dense urban areas, which discriminate roads from the background:

- Roads are mainly built from the same materials (concrete, asphalt) and thus tend to have somewhat homogeneous spectral properties. In contrast, the background (*i.e.* non-road regions) has no single photometric characteristic.
- The surfaces of main roads are not entirely uniform due to the presence of noise, such as zebra crossings, overbridges, vehicles, shadows, etc. Nevertheless, they still show much less variability than the background.

E_D is the negative logarithm of $P(I|R,K)$ in (1). We assume that this factorizes as $P(I_R|R,K)P(I_{\bar{R}}|R,K)$, where subscripts indicate 'restricted to'. We use the same parameterized model for I_R and $I_{\bar{R}}$, the choice of model being based on a study of the image statistics. We model both the one point statistics of the image intensity, *i.e.* the histogram, and the two-point statistics, which we characterize by the variance $V(x)$ of the image in a small window around each pixel. Because of the factorization, the data energy is the sum of two pieces, one referring to I_R and one to $I_{\bar{R}}$:

$$E_{D,\text{SIG}}(I,\phi) = -\int_\Omega \Big\{ \big[\ln P_+(I(x)) + \theta_\text{v} \ln Q_+(V(x))\big]\phi_+(x) + \big[\ln P_-(I(x)) + \theta_\text{v} \ln Q_-(V(x))\big]\phi_-(x) \Big\}\ , \quad (7)$$

where θ_v is the weight of the two-point statistics. P_+ and P_- are two-component Gaussian mixture models, modelling the image intensities, while Q_+ and Q_- are Gamma distributions, modelling the variances:

$$P_\pm(I) = a_\pm N(I;\mu_{1\pm},\sigma^2_{1\pm}) + (1-a_\pm)N(I;\mu_{2\pm},\sigma^2_{2\pm})\ , \quad (8a)$$

$$Q_\pm(V) = \frac{V^{b_\pm}}{d_\pm} e^{-\frac{V}{c_\pm}}\ , \quad (8b)$$

where $+$ denotes the road and $-$ denotes the background, $a_\pm \in [0,1]$, and N is the normal distribution.

The complexity of VHR images in dense urban areas, however, compels us to introduce a multi-resolution approach. The motivations are as follows. First, as noted in section I, VHR images contain objects, *e.g.* roads, buildings, at many different scales. In order to capture this complicated behaviour, it makes sense to analyse an image at several resolutions simultaneously. Second, the same object observed at high or low resolutions presents different characteristics. In particular, at low resolutions, the background can be viewed as noise, while the larger roads are still clearly distinguished as homogeneous areas. Road segmentation is thereby facilitated, but is also less precise. In contrast, higher resolutions can provide a more precise location and width for the road. The use of several resolutions thus allows the combination of coarse data, in which details in the image that can disrupt the recognition process have been eliminated, with fine data to increase precision.

The multi-scale framework we use was first introduced in [11]. The Haar wavelet scaling coefficients at different scales (levels) [12] provide us with a multi-scale representation of the original data. Our multi-resolution data energy is a sum of energies computed at four different levels:

$$E_{D,\text{MUL}}(I,\phi) = \sum_s E_{D,\text{SIG},s}(I_s,\phi)\ , \quad (9)$$

where $I_s, s \in \{0,1,2,3\}$, are the scaling coefficients at level s of a wavelet transform and $E_{D,\text{SIG},s}$ is the data energy at a single level s. In practice, since the size of the image varies with a factor 2 from level s to level $s+1$, we up-sample all I_s to the finest resolution.

III. GIS Prior Energy

The prior energy E_P proposed in section II.1 is *generic*: it incorporates constraints on the form of the road network region that are true of any road network. To improve further the results at the original resolution, we introduced a *specific* prior energy [13]. It says that the region sought must be 'close' to an exemplar region, *e.g.* a GIS map of the road network R_0 at an earlier date, which can also be described by its minimum energy phase field function ϕ_{R_0}. $\phi_{R_0\pm} = (1 \pm \phi_{R_0})/2$ thus denote the characteristic functions of R_0 and $\bar{R}_0$. $E_{P,GIS}$ takes the form

$$E_{P,GIS}(\phi,\phi_{R_0}) = \int_\Omega dx\ \big[\omega_+\phi_{R_0+}(x) + \omega_-\phi_{R_0-}(x)\big]\big[\phi(x) - \phi_{R_0}(x)\big]^2\ . \quad (10)$$

The two terms correspond to the two components of the symmetric area difference between R and R_0: $x \in R \cap \bar{R}_0$ and $x \in \bar{R} \cap R_0$ respectively. These are separated so that they can be weighted differently by the parameters ω_+ and ω_-. Since this term takes into account the exterior of R_0 (*i.e.* $\bar{R}_0$), it counteracts the background 'noise' appearing in the data.

IV. Implementation

The parameters of the Gaussian mixture and Gamma distributions in E_D are learned with supervised learning. The known region R_0 is used to create samples of road and non-road in the image. Note that the samples may contain errors, since R_0 does not correspond exactly to the current road

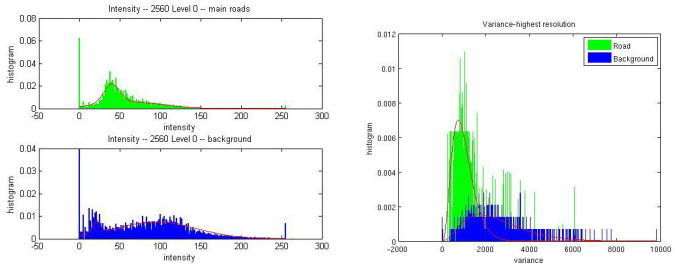
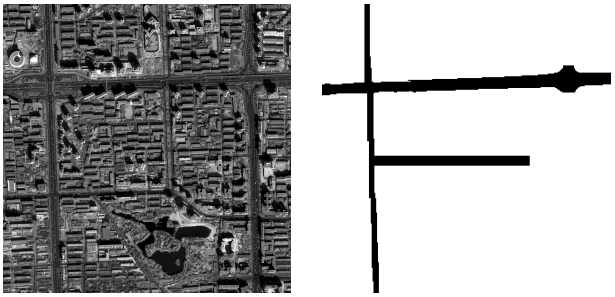


Fig. 2. Left: histograms of the pixel intensity I on-road (top) and off-road (bottom); right: histograms of the variances V on-road (green/light grey) and off-road (blue/dark grey), and of the models fitted to them (solid lines).



(a) (b)

Fig. 3. A QuickBird image (size: 2560×2560) and deliberately 'damaged' ground truth, to simulate an earlier GIS map.

network (see Fig. 3). Examples of histograms and the models fitted to them are shown in Fig. 2.

To minimize the total energy E, we performed a gradient descent with the neutral initialization [1]: the initial value of ϕ is set equal to α/λ everywhere in Ω, which corresponds to the local maximum of the potential W. During the algorithm, no re-initialization or *ad hoc* regularization is required. Parameter values and weights are for the moment set by hand, but they are constrained by Turing stability of the model [1], and further by a condition that ensures that a long bar of the desired road width is a stable configuration of the energy. In the case of the single-scale model energy $E_1 = \theta(E_{P,0} + E_{P,NL}) + E_{D,\text{SIG}}$, the evolution equation is

$$\frac{\partial \phi}{\partial t} = \theta\Big\{\nabla^2\phi - \lambda(\phi^3 - \phi) - \alpha(1-\phi^2) - \beta\nabla^2\Psi * \phi\Big\} + \frac{1}{2}\Big\{\big[\ln P_+ + \theta_{\text{v}} \ln Q_+\big] - \big[\ln P_- + \theta_{\text{v}} \ln Q_-\big]\Big\}, \quad (11)$$

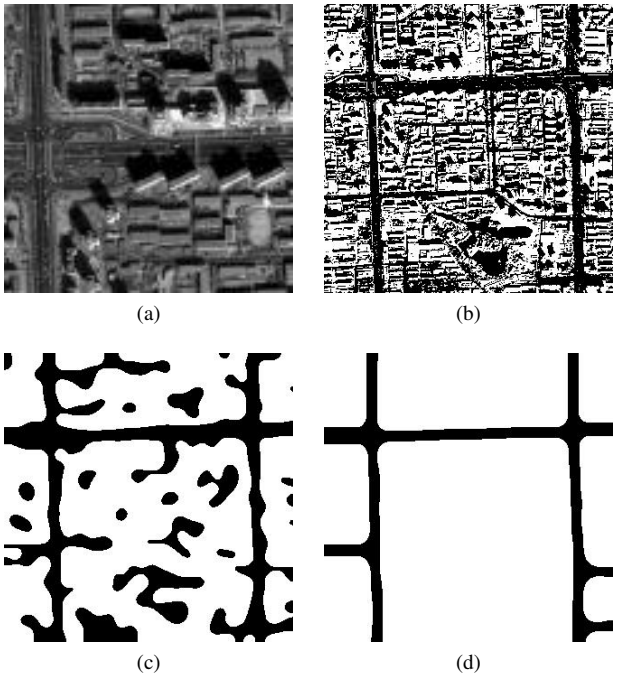
where $*$ indicates convolution. The equation for multiple scales involves adding a copy of the last line for each scale. When the GIS prior is taken into account, the term $-2(\phi - \phi_{R_0})\big[\omega\phi_{R_0+} + \bar{\omega}\phi_{R_0-}\big]$ should be added to (11).

V. Experimental Results

The input data I was a QuickBird panchromatic image (see Fig. 3a). The associated GIS map was obtained a few years earlier than the satellite image, and thus represents a slightly different road network. For the purpose of testing the robustness of our model, we introduced artificial errors that increase the difficulty of the problem, as shown in Fig. 3b. Note that R_0 has some roads added, some roads narrowed and some roads missing. We will first present the results of applying the single-scale model to the images at different resolutions and of applying the multi-scale model to the original image. Then, we will show that the GIS prior energy can significantly increase the robustness of the method. More results and comparisons with several other approaches are presented at the end.

A. Results Using the Single-Scale Model without GIS Prior

First we apply the single-scale model E_1 to the scaling coefficients of the original image (Fig. 3a) at different levels of the wavelet decomposition.

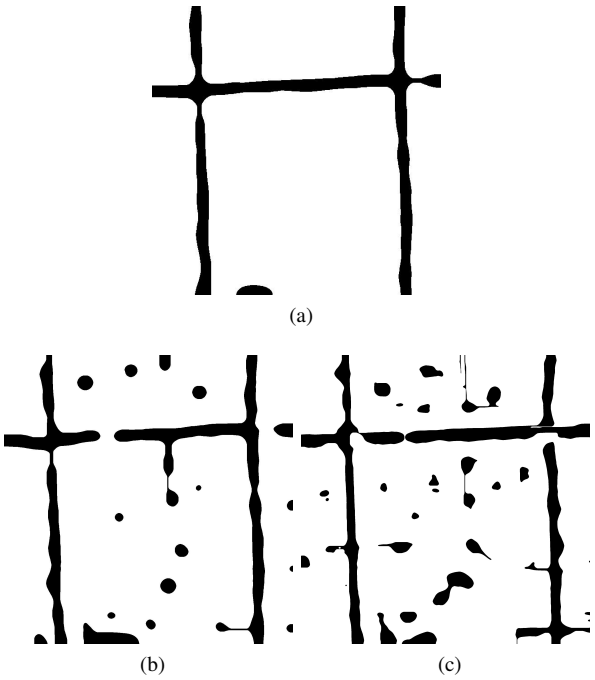


(a) (b) (c) (d)

Fig. 4. Experiments at level 3 (size: 320×320, road width $\simeq 12$ pixels). 4a: zoom to illustrate the complexity remaining even at level 3; 4b-4d: the thresholded phase field function at iterations 1 and 400, and at the final iteration 20000, using the single-scale model E_1.

We start from level 3 (equivalent to a compression factor of 2^3). Fig. 4a shows a zoom on part of the input image at this level, illustrating that even after three levels of smoothing and down-sampling, the data is still rather complex. Figs. 4b-4d show the thresholded phase field function at iterations 1 and 400 of gradient descent, and at the final iteration 20000, using the model E_1. The segmentation is very successful: the main road networks are retrieved nearly completely.

The level 3 image is already quite complex, and we observe experimentally that if we try to use the same model at finer resolutions, using the images at levels 2, 1, or 0, the details of the scene in the image make road extraction more difficult (see Fig. 5). The erroneous detections in the background result, on the one hand, from regions of poor contrast between the main roads and the buildings or areas of vegetation, and on the other hand, from the smaller roads, which have statistical properties



(a) (b) (c)

Fig. 5. Experiments at finer resolutions using the single-scale model E_1. 5a: result at level 2 (size: 640×640, road width $\simeq 24$ pixels); 5b: result at level 1 (size: 1280×1280, road width $\simeq 48$ pixels); 5c: result at full resolution, level 0 (size: 2560×2560, road width $\simeq 96$ pixels).

similar to the main roads. The shadows of high buildings, cars, road markings and bridges lead to jagged borders or gaps along the roads. The former indicates a lack in the single level image model, while the latter seems more likely to be due to a weakness in the prior model, which therefore needs to be improved in order to enforce the road geometry more effectively.

B. Results Using the Multi-Scale Model without GIS Prior

In an attempt to overcome the above problems at finer resolutions, we apply the multi-scale model $E_2 = \theta(E_{P,0} + E_{P,NL}) + E_{D,\text{MUL}}$. The result is shown in Fig. 6. The result is not perfect, but the use of a multi-scale model improves the result obtained from the original image at a single level (see Fig. 5c). However, there are still some false detections in the background and the road borders are rather inaccurate due to geometric noise along the boundaries of the road. The result indicates that a simple sum of data energies at several different scales, while helpful, is not sufficient to solve the problem completely. It suggests that the model should include stronger prior knowledge. We show, in the next subsection, the great improvement at full resolution resulting from the incorporation of the GIS prior.

C. Results Using the Single-Scale GIS Model

In this subsection we apply a model making use of the single scale data term and the GIS prior term $E_3 = \theta(E_{P,0}+E_{P,NL}+E_{P,GIS})+E_{D,\text{SIG}}$ at full resolution. The result is illustrated in Fig. 7a. The addition of $E_{P,GIS}$ greatly improves the result, when compared to Figs. 5c and 6. Its main effect is to eliminate

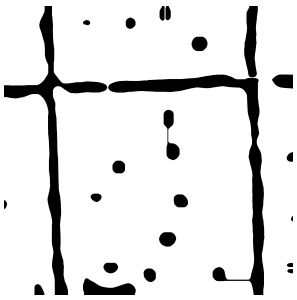
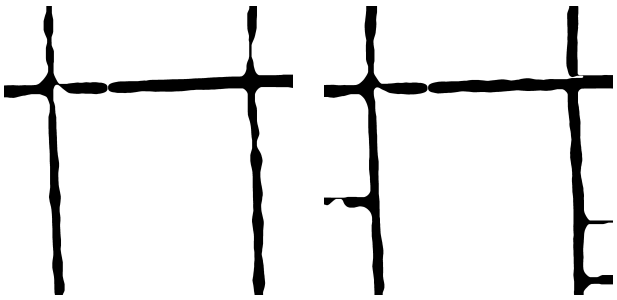


Fig. 6. Experiment at full resolution (size: 2560×2560, road width $\simeq 96$ pixels) using the multi-scale model E_2.



(a) (b)

Fig. 7. Experiments at full resolution (size: 2560×2560, road width $\simeq 96$ pixels) using the single-scale GIS model E_3. Left: result obtained with 'damaged' ground truth (Fig. 3b); right: result obtained using the result obtained with E_1 at level 3 (Fig. 4d) as a replacement for the GIS information.

false positives in the background, while preserving the correct segmentation of the roads themselves. To obtain this result, ω_+ must be small, since the mistakes that may exist in the old map, should not affect the process, while ω_- is somewhat bigger, because a strong constraint is needed to overcome the 'noise' in the background.

Fig. 7b shows the result we obtain when we use as R_0, not the GIS map, but the result obtained at reduced resolution, *i.e.* level 3 (see Fig. 4d). This shows that, in principle, we can free ourselves from the need to have a GIS map available.

In Fig. 8, the top row shows the real outdated GIS maps of main road networks before the year 2002 covering the same zones as images in Fig. 1. We can see that significant changes exist between these maps and the QuickBird images in the year 2006 (Fig. 1). The results obtained with the single-scale GIS model E_3 from images in the year 2006 are shown in the bottom row of Fig. 8. All these experiments prove that for the main roads, at full resolution, the single-scale GIS model is able to keep the unchanged roads, to correct the mistakes, and to extract new roads. However, it is still not capable of retrieving the smaller roads very accurately (see the rightmost vertical road in Fig. 8d).

D. Evaluation and Comparison

To evaluate the performance of our models, we compare our results with four other methods: a simple maximum likelihood estimation (MLE); a level set approach with global shape constraint by Bailloeul [14]; a classification, tracking, and

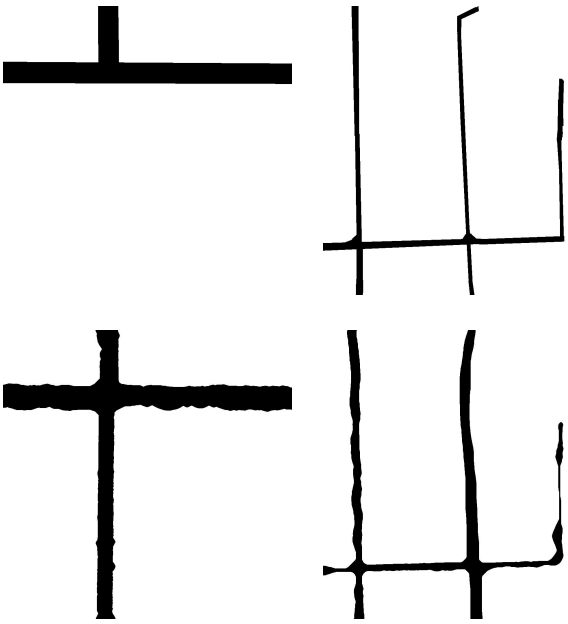


Fig. 8. More experiments on the two pairs of images in Fig. 1 at full resolution using the single-scale GIS model E_3. Top: the real GIS maps of main road networks before the year 2002; bottom: the updated main road networks from QuickBird images in the year 2006.

morphology algorithm by Wang [15]; and a fast but rough segmentation technique based on "straight line density" by Yu [16]. Except for those involving the multi-scale model, all results are obtained from full resolution images.

Fig. 9a shows the ground truth used to calculate quantitative criteria. Figs. 9c-9f illustrate the results obtained using the four methods mentioned above. MLE is obviously insufficient to distinguish the roads from the background. The 'flexible active contour' method of Bailloeul (initially dedicated to building extraction) fails because it is not able to eliminate road sections that exist in the map but not in the image. On the other hand, the methods of Yu and Wang are able to detect the main road network and smaller roads, but, for both, the accuracy obtained in the delineation of the road boundary is poor, and the results show a great deal of noise. In addition, in order to illustrate the importance of the generic prior term of our model, Fig. 9b shows the result obtained if $E_{P,NL}$ is omitted, leaving a model equivalent to a standard (*i.e.* not higher-order) active contour $\theta E_{P,0} + E_{D,\text{SIG}}$. The importance of the geometric prior knowledge carried by the prior term is clear.

Some quantitative evaluation measures [17] are shown in Table I. 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. Note that for each method and each measure, the average value of three experimental images (Figs. 1c, 1d and 3a) is calculated.

TABLE I
QUALITY MEASURES OF THE DIFFERENT METHODS TESTED AT FULL RESOLUTION (T = TRUE, F = FALSE, P = POSITIVE, N = NEGATIVE)

Method \ Measure	Completeness TP/(TP+FN)	Correctness TP/(TP+FP)	Quality TP/(TP+FP+FN)
Our model E_1	0.8159	0.6758	0.5893
Our model E_2	0.7859	0.7301	0.6141
Our model E_3	0.7920	0.8914	0.7198
$\theta E_{P,0} + E_{D,\text{SIG}}$	0.6358	0.8743	0.5810
MLE	0.7567	0.2775	0.2545
Bailloeul	0.5529	0.8318	0.4990
Wang	0.8918	0.6180	0.5717
Yu	0.7743	0.7196	0.5893

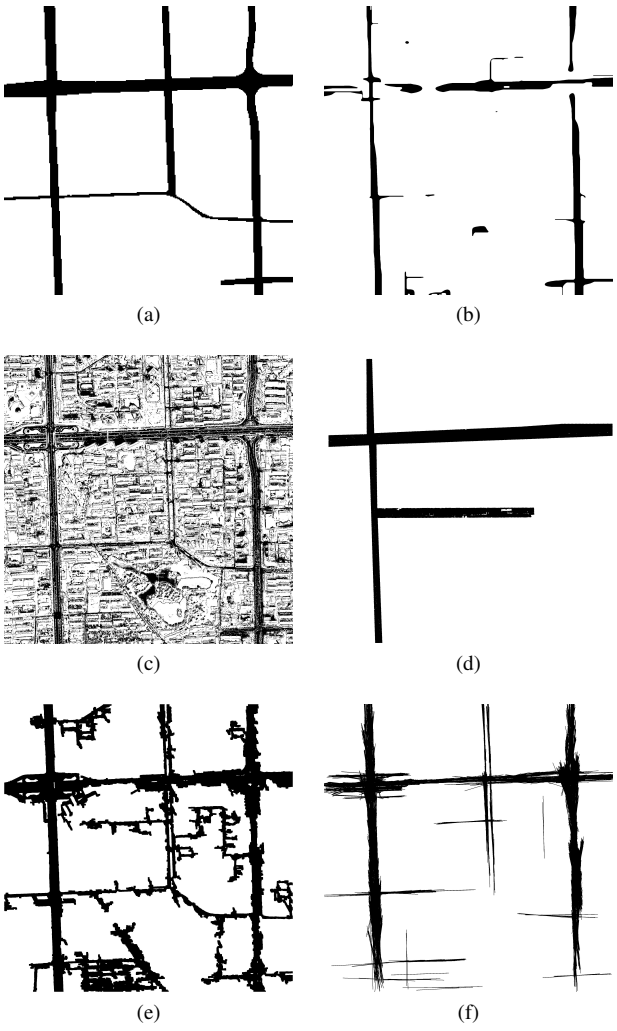


Fig. 9. Ground truth and comparison with methods taken from the literature, at full resolution. 9a: ground truth used to calculate the quantitative measures; 9b: result obtained when the higher-order term $E_{P,NL}$ is dropped, leaving a model equivalent to a standard active contour; 9c-9f: results obtained respectively using MLE, the approaches of Bailloeul [14], Wang [15], and Yu [16].

VI. CONCLUSION

We have presented two models for the updating of road maps in dense urban areas by extracting the main road network

from VHR QuickBird panchromatic images. To adapt the original phase field HOAC model [1], which was developed for road extraction from medium resolution images, to VHR images, we first proposed a new multi-resolution data energy. Although the result at full resolution is better than that obtained with the single-scale model, the multi-scale approach needed further improvements in order to eliminate false detections and improve the accuracy of road border delineation. Consequently, we introduced *specific* prior knowledge in the form of an outdate GIS map, to complement the *generic* prior knowledge encoded by HOACs. Our results indicate that, when working at full resolution, the combination of *generic* and *specific* prior knowledge is essential, due to the great complexity of VHR images. Our model gives better results than several other methods in the literature, and our on-going work focuses on the extraction of the smaller roads.

ACKNOWLEDGMENT

The authors would like to thank the Beijing Institute of Surveying and Mapping for providing the GIS data.

REFERENCES

[1] M. Rochery, I. H. Jermyn, and J. Zerubia, "Phase field models and higher-order active contours," in *Proc. ICCV*, Beijing, China, Oct. 2005.
[2] H. Mayer, I. Laptev, and A. Baumgartner, "Multi-scale and snakes for automatic road extraction," in *Proc. ECCV*, vol. 1, Freiburg, Germany, July 1998.
[3] R. Péteri and T. Ranchin, "Detection and extraction of road networks from high resolution satellite images," in *Proc. ICIP*, Barcelona, Spain, Sept. 2003.
[4] M. F. A. Fortier, D. Ziou, C. Armenakis, and S. Wang, "Automated correction and updating of road databases from high-resolution imagery," *Can. J. Rem. Sens.*, vol. 27, no. 1, pp. 76–89, 2001.
[5] P. Agouris, A. Stefanidis, and S. Gyftakis, "Differential snakes for change detection in road segments," *Photogramm. Eng. Rem. Sens.*, vol. 67, no. 12, pp. 1391–1399, 2001.
[6] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, 2001.
[7] Y. Chen, H. Tagare, S. Thiruvenkadam, F. Huang, D. Wilson, K. Gopinath, R. Briggs, and E. Geiser, "Using prior shapes in geometric active contours in a variational framework," *Int. J. Comput. Vis.*, vol. 50, no. 3, pp. 315–328, 2002.
[8] D. Cremers, F. Tischhäuser, J. Weickert, and C. Schnörr, "Diffusion snakes: Introducing statistical shape knowledge into the Mumford-Shah functional," *Int. J. Comput. Vis.*, vol. 50, no. 3, pp. 295–313, 2002.
[9] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.*, vol. 1, no. 4, pp. 321–331, 1988.
[10] M. Rochery, I. H. Jermyn, and J. Zerubia, "Higher-order active contours," *Int. J. Comput. Vis.*, vol. 69, no. 1, pp. 27–42, 2006.
[11] T. Peng, I. H. Jermyn, V. Prinet, J. Zerubia, and B. Hu, "Urban road extraction from VHR images using a multiscale approach and a phase field model of network geometry," in *Proc. URBAN*, Paris, France, Apr. 2007.
[12] S. Mallat, *A wavelet tour of signal processing*, 2^{nd} ed. Academic Press, 1999.
[13] T. Peng, I. H. Jermyn, V. Prinet, J. Zerubia, and B. Hu, "A phase field model incorporating generic and specific prior knowledge applied to road network extraction from VHR satellite images," in *Proc. BMVC*, Warwick, England, Sept. 2007.
[14] T. Bailloeul, "Active contours and prior knowledge for change analysis: Application to digital urban building map updating from optical high resolution remote sensing images," Ph.D. dissertation, CASIA and INPT, Oct. 2005. [Online]. Available: http://kepler.ia.ac.cn
[15] R. Wang and Y. Zhang, "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, Oct. 2003.
[16] Z. Yu, V. Prinet, C. Pan, and P. Chen, "A novel two-steps strategy for automatic GIS-image registration," in *Proc. ICIP*, Singapore, Oct. 2004.
[17] C. Heipke, H. Mayr, C. Wiedemann, and O. Jamet, "Evaluation of automatic road extraction," *Int. Arch. Photogram. Rem. Sens.*, vol. XXXII, pp. 47–56, 1997.