

# A Point Process for Fully Automatic Road Network Detection in Satellite and Aerial Images<sup>1</sup>

P. Cariou\*, X. Descombes\*\*, and E. Zhizhina\*\*\*

\*ISAE - 10 av. Edouard Belin - BP 54032 - 31055 TOULOUSE Cedex 4, France  
email pierre.cariou@supaero.org

\*\*INRIA-SAM, 2004 route des Lucioles, BP93 06902 Sophia Antipolis, cedex, France  
Xavier.Descombes@sophia.inria.fr

\*\*\*Institute for Information Transmission Problems,  
Bolshoy Karetny per. 19, 127994 GPS-4, Moscow, Russia  
ejj@iitp.ru

Поступила в редколлегию 17.09.2010

**Аннотация**—In this paper, we introduce a new model for road detection, based on the multiple birth and death (MBD) optimization algorithm, recently proposed for counting populations of trees or flamingos. We discuss the choice of the energy function guaranteeing accurate detections of both junctions and parallel roads, and preserving the connectivity of the road network despite the occlusions in high resolution images. We also propose here a novel local filter used as a preprocessing step, for detecting roads, and involved in the definition of the energy function.

## 1. INTRODUCTION

We address the problem of road network extraction from satellite or aerial images in order to produce or update maps. To tackle this difficult problem, a large class of methods consists of semi-automatic approaches. In these methods, an operator gives a starting point (a seed) and a direction to initialize a tracking algorithm [13, 4], some endpoints to be linked by a dynamic programming procedure [3, 8] or some control points to initialize a deformable contour model [9] or a dynamic programming algorithm [6]. Fully automatic road network extraction from satellite images, although much more difficult to be solved, is crucial when considering concrete applications. Indeed, the gain in time of an operator is low if one needs to provide seeds or initial checking points. Fully automatic techniques based on local operators lack robustness with respect to noise, especially for high resolution images, showing some occluded road segments in case of tree shadow for example. To reduce this sensitivity to noise, a possibility is to combine several local operators [14], for example to estimate seeds for a semi-automatic algorithm [1]. However, adding prior information in the modeling allows more flexibility and robustness. This information can be embedded in the model through variational [10] or probabilistic frameworks [12]. In this case, the complexity of the models leads to heavy computational issues. A good compromise between complexity, accuracy and generality is still challenging. Recently, a new approach based on a marked point process modeling and called the Quality Candy model has been proposed in [7]. Since roads are associated with linear objects, the grains in the model are considered to be segments, and every segment is determined by a point (the location of the center) and marks (the orientation and the length). The prior knowledge models the high connectivity and the low curvature of a road network using some interactions between neighboring segments. The optimization in this model is performed by an RJMCMC algorithm. The results obtained are satisfactory in terms of network connectivity and are robust with respect to

<sup>1</sup> This work has been conducted within the Associated Team ODESSA supported by INRIA.

occlusions. However, junctions are not taken into account in the prior which leads to misdetections in such areas. Moreover, the optimization using RJMCMC algorithms converges usually slowly. That is why we propose a new model for the road network extraction based on marked point process approach. In our model, we consider a point process (with points as the grains) to speed up the convergence. Connectivity, curvature and junctions of the network are modeled in the prior whereas the data term is computed from a local road detector. Point process models are usually optimized by a RJMCMC sampler embedded in a simulated annealing scheme [5] or a combination of jump kernels and diffusion [11]. In this paper, we consider a recently proposed multiple birth and death (MBD) algorithm, based on the Gibbs field approach [2]. The main advantage of this algorithm with respect to RJMCMC is that each step of the iterative probabilistic scheme concerns the whole configuration and there is no rejection. During the birth step at each iteration, several objects (a random set) are added to the configuration independently of both the energy function and the temperature. During the death step at each iteration, a subconfiguration of the current configuration, and not only one object, can be deleted. Such jumps are much more efficient for scanning the configuration space.

We first describe a local road detector in section 2. The point process, modeling the road network, is detailed in section 3 and the new optimization algorithm in section 4. Results and comparison with other models are shown in section 5 and conclusion are drawn in section 6.

## 2. A LOCAL ROAD DETECTOR

We propose in this section a local road detector (LRD). It is a two steps transformation of the input image  $\mathcal{I}$  resulting in the filtered image  $\mathcal{U} = \{u_p\}$  (so-called LRD image). The LRD image  $\mathcal{U}$  is obtained by computing a local contrast term from  $\mathcal{I}$  weighted by coefficients depending the local orientation consistency of the gradient (see formula (4) below). The image  $\mathcal{U}$  is considered as the data for the proposed point process modeling road networks.

The input image  $\mathcal{I}$  is a real-valued function (the grey level function) defined on a bounded subset  $\Omega$  of  $\mathbb{Z}^2$

$$\mathcal{I} : \Omega \subset \mathbb{Z}^2 \rightarrow [0, 255]. \quad (1)$$

We denote by  $v = \mathcal{I}/255$  the normalized grey level function, so that  $v_p \in [0, 1]$ ,  $p \in \Omega$ .

Locally a road has a contrast with both sides. We propose to estimate this property by considering the following filter:

$$\forall p \in \Omega, \tilde{n}_p = \max_{q \in O_1(p)} [\min(v_p - v_{p+q}; v_p - v_{p-q})], \quad (2)$$

where  $O_1(p)$  is a fixed neighborhood of pixel  $p$  ( $O_1(p) = \{q = (q_i, q_j) \in \Omega : \max(|p_i - q_i|, |p_j - q_j|) < l\}$ ). The size  $l$  of this neighborhood corresponds to the maximum width of roads to be detected and is therefore linked to the data resolution and the type of roads. In the experiments shown in this paper, we have taken  $l = 2$ , which means that we consider roads having a width lower or equal to 5 pixels. It is clear that  $\tilde{n}_p \in [-1, 1]$ , and 1 represents strongly bright points which have at least two symmetrical dark points from the both sides. Since we assume that roads are bright lines on a dark background, we will consider the following contrast term

$$n_p = \max\{\tilde{n}_p, 0\}, \quad (3)$$

and  $\{n_p\}$  is called the contrast image.

To reduce the noise induced by this simple local filter (including for example, village areas with isolated high contrast points), we regularize the output by an adaptative smoothing. We first

estimate locally the perpendicular orientation of the road,  $\theta_p$ , given by the eigen vector corresponding to the leading eigen value (the eigen value with maximal modulus) of the difference Hessian  $\nabla_d \mathcal{I}$ .

The contrast image is finally filtered as follows:

$$u_p = 1 - 2n_p \frac{\sum_{q \sim p} \frac{n_q}{r_{pq}^2} \phi(\theta_q - \theta_p)}{\sum_{q \sim p} \frac{n_q}{r_{pq}^2}}, \quad u_p \in [-1, 1], \quad (4)$$

and the resulting image  $\{u_p\}$  is called the LRD image. Here  $r_{pq}$  is the distance between pixels  $p$  and  $q$ ,  $q \sim p$  means that  $q \in O_2(p) \{q = (q_i, q_j) \in \Omega : \max(|p_i - q_i|, |p_j - q_j|) < w\}$ .  $w$  is a smoothing parameter which is equal to 8 for all the results shown in this paper.  $\phi$  is defined as a  $\pi$ -periodic as follows:

$$\phi(\theta_q - \theta_p) = \max \left( 0, 1 - 2 \frac{|\theta_p - \theta_q| \bmod(\pi)}{\theta_0} \right), \quad (5)$$

where  $\theta_0$  is a parameter depending on the road network curvature. It can takes different values for mountain road or in urban areas. It is equal to  $\pi/2$  for the results presented in this paper. The output of the LRD,  $u_p$ , takes low values (close to -1), if most of contrasted neighbors  $n_p$  have spacially homogeneous orientation.

The result of this local road detector is shown on an aerial image on Fig. 2. The contrast filter gives a noisy result (cf Fig. 2 top right). The estimated orientation is shown on pixels having a strictly positive value on the contrast filter (see Fig. 2 bottom left), the segment size being proportional to the corresponding eigen value. Finally, Fig. 2 bottom right shows the proposed local road detector, which is considered as the data (the LRD image) for the point process model of the road network.



**Fig.1.** Aerial image © IGN (top left), contrast filter (top right), local orientation (bottom left) and LRD (bottom right).

Although the LRD image gives an authentic view on the road network, it has evident defects: gaps in the network and junctions, residual noise. To improve results we propose here a stochastic algorithm based on point processes approach.

### 3. A POINT PROCESS FOR MODELING ROAD NETWORKS

Let  $V \in \mathbb{R}^2$  be the bounded continuous domain supporting  $\Omega$ . We define the set of point configurations in  $V$ , denoted  $\Gamma_V$ , as follows:

$$\Gamma_V = \bigcup_{n=0}^{\infty} \Gamma_V^{(n)}, \quad (6)$$

where  $\Gamma_V^{(n)} = \{\gamma = \{x_1, \dots, x_n\}, x_i \in V\}$  is the set of configurations with exactly  $n$  points. In addition we suppose that configurations  $\gamma \in \Gamma_V$  meet the following condition: there exists at most one point from  $\gamma$  at each pixel  $p \in \Omega$ .

The pixel in  $\Omega$  containing point  $x \in V$  from the configuration  $\gamma$  is denoted by  $p(x)$

The energy function is a sum of three terms: the data term (8), the interaction term for connectivity (11) and a prior (14) for managing curvature and junctions:

$$U(\gamma) = U_d(\gamma) + U_i(\gamma) + U_p(\gamma), \quad (7)$$

the weighting factors between these three terms being included in their definition.

The data driven term has the following form:

$$U_d(\gamma) = \sum_{x \in \gamma} \left( U_d^{(1)}(x) + U_d^{(2)}(x) \right), \quad (8)$$

where:

$$U_d^{(1)}(x) = U_{of} + u_{p(x)}, \quad (9)$$

and:

$$U_d^{(2)}(x) = \begin{cases} 0 & \text{if } x \text{ has no neighbors from } \gamma \\ U_{of} + \min_{y \in \gamma: y \sim_2 x} \left( \frac{1}{r_{xy} \sum_{s \in [p(x), p(y)]} u_s} \right) & \text{otherwise} \end{cases} \quad (10)$$

where  $\sim_2$  is a neighborhood relation such that  $x \sim_2 y \Leftrightarrow d(x, y) \leq D$ ,  $D$  being a parameter which defines the density of points along the roads ( $D$  is set to 10 pixels in our experiments). The first data term attracts or penalizes points depending on the LRD value taken at their location. The first data term has a general form, whereas the second data term is specific for a problem under consideration. The second data term considers the different neighbors from the configuration of a given point and favors it if at least one path between this point and one of its neighbors corresponds to "good" values of the LRD image. We consider the min operator to avoid penalizing neighboring points lying on junctions or close parallel roads.

We now consider a connectivity condition generated by pair potential depending on the LRD image and two additional parameters  $U_+$  and  $U_-$ :

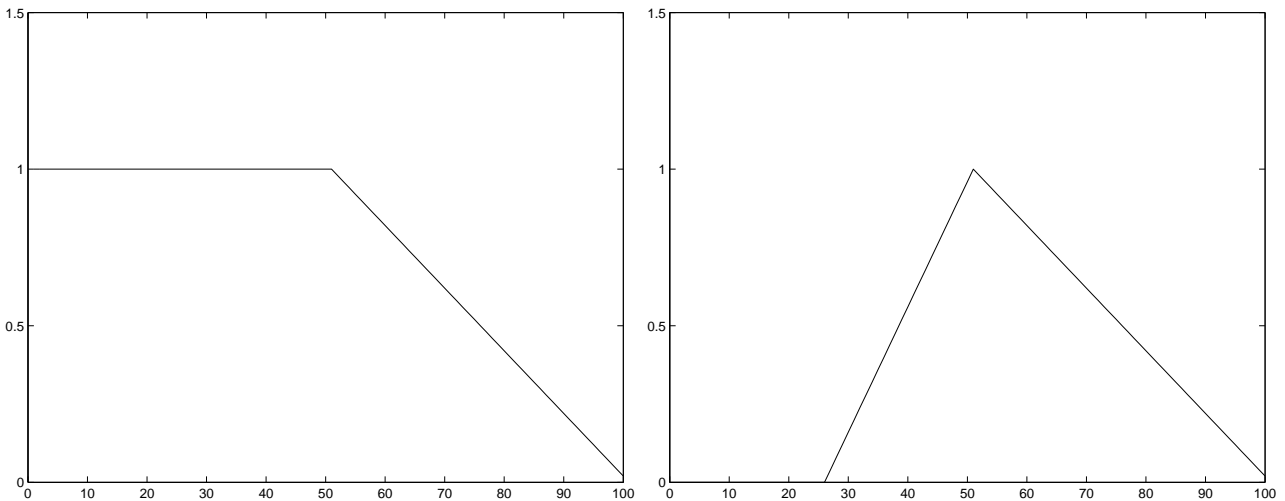
$$U_i(\gamma) = \sum_{x, y \in \gamma: x \sim_2 y} U_i(x - y). \quad (11)$$

$U_i$  is a repulsive energy on short distances  $|x - y|$  to prevent accumulation of points in the configuration and an attractive energy on fitting distances  $|x - y|$ , and  $U_i$  vanishes on large enough distances:

$$U_i(x - y) = \begin{cases} \lambda U_+, & \text{if } d \leq \frac{1}{4}, \\ \lambda(2 - 4d)U_+ + (-1 + 4d)U_-, & \text{if } \frac{1}{4} < d < \frac{1}{2}, \\ (2 - 2d)U_-, & \text{if } \frac{1}{2} \leq d \leq 1 \end{cases} \quad (12)$$

where  $d = \frac{|x - y|}{D}$ . The repulsive property aims at reducing the number of points describing a road (when  $\lambda$  takes values close to 1), but should not penalize close roads or hidden parts of the road (the case when  $\lambda$  close to 0). On figure 3 left,  $\lambda$  is equal to 1 and the energy is positive for close point within the same road, whereas on figure 3 right,  $\lambda$  is equal to 0, and the energy is zero for small distances between points so that close points belonging to two different roads are not penalized. Therefore, we introduce a modulating factor  $\lambda$  depending on the LRD values in the potential to improve detection of junctions and parallel roads:

$$\lambda = \lambda(\{u_p\}; x, y) = 1 + \frac{u_{p(x)} + u_{p(y)} - 2u_m}{4}, \quad m = p\left(\frac{x + y}{2}\right). \quad (13)$$



**Fig.2.** Energy function  $U_i(|x - y|)$  for  $\lambda = 1$  (left) and  $\lambda = 0$  (right).

To model curvature and junctions in road networks we introduce a priori term depending on the number of neighbors of a given point as follows:

$$U_p(\gamma) = \sum_{x \in \gamma} E(n(x)), \quad (14)$$

where  $n(x)$  is the number of neighbors of point  $x$ :  $n(x) = |\{y \in \gamma \setminus \{x\} : d(x, y) \leq D\}|$ , and:

$$E(0) = E(1) = E_1 \tag{15}$$

$$E(2) = \begin{cases} E_2 \left(1 - \frac{\theta}{\theta_0}\right)^2, & \text{if } \theta \leq \theta_0; \\ -E_2 \frac{\theta - \theta_0}{\pi - \theta_0}, & \text{if } \theta > \theta_0, \end{cases} \tag{16}$$

$$E(3) = \begin{cases} E_2 \left(1 - \frac{\theta}{\frac{2\theta_0}{3}}\right)^2, & \text{if } \theta \leq \frac{2\theta_0}{3}; \\ -E_2 \frac{\theta - \frac{2\theta_0}{3}}{\pi - \frac{2\theta_0}{3}}, & \text{if } \theta > \frac{2\theta_0}{3}, \end{cases} \tag{17}$$

$$E(n) = E_2 \times (n - 3) \text{ if } n > 3. \tag{18}$$

Here  $\theta = \theta(x, \gamma)$  is the minimal angle between vectors started at  $x$  and going to its neighbors.  $\theta_0$  is a parameter depending on the area under study (american city or sinuous hilly zone).

#### 4. OPTIMIZATION PROCEDURE

We optimize the model using the multiple birth and death algorithm proposed in [2]. The algorithm finds configurations minimizing the energy function  $U(\gamma)$  (7). It is a Markov chain on the configuration space  $\Gamma_V$ , constructed as an approximation (in time) of a continuous time stochastic birth and death dynamics in continuum embedded into simulated annealing procedure. At each iteration, we consider transitions from the current configuration  $\gamma$  to any configuration of the form  $(\gamma \setminus \gamma_1) \cup \omega$  with  $\gamma_1 \subseteq \gamma$ . The corresponding transition probability is written as follows:

$$Pr(\gamma \rightarrow (\gamma \setminus \gamma_1) \cup \omega) = \frac{1}{k(V, \delta)} \delta^{|\omega|} \prod_{x \in \gamma_1} \frac{a_\beta(x)\delta}{1 + a_\beta(x)\delta} \prod_{x \in \gamma \setminus \gamma_1} \frac{1}{1 + a_\beta(x)\delta} \tag{19}$$

with  $a_\beta(x) = \exp\{\beta(U(\gamma) - U(\gamma \setminus x))\}$ , the energy function  $U(\gamma)$  is defined by (7) and  $k(V, \delta)$  is the normalizing factor. The convergence of the Markov chain (19) to a measure concentrated on the set of global minimizers of  $U(\gamma)$  under a proper decreasing scheme of parameters  $\delta$  and  $1/\beta$  has been proved in [2]. One can see that the transition probability for the death of a particle depends on the relative energy of the particle in the configuration, the transition probability of the birth is spatially homogeneous and independent on both the energy and the temperature. The algorithm is written as follows:

- **Main program:** initialize the inverse temperature parameter  $\beta(\beta_0 = 1)$  and the discretization step  $\delta(\delta_0 = 250)$  and alternate birth and death steps
  - **Birth step:** Add a realization of the Poisson process with intensity  $\delta$
  - **Sorting step:** once the birth step is finished, compute the data term  $U_d^{(1)}(p(x_i)), x_i \in \gamma$  of the current configuration objects  $x_i$ . Then, sort them, in the decreasing order, according to their data energy.
  - **Death step:** for each object  $x$  in the current configuration  $\gamma$ , taken in the above order, compute the death rate as follows:

$$d(x) = \frac{\delta a_\beta(x)}{1 + \delta a_\beta(x)} \tag{20}$$

then the point  $x$  is removed from the configuration with probability  $d(x)$ .

- **Convergence test:** if the process has not converged, decrease the temperature and the discretization step by given factors and go back to the birth step. The convergence is obtained when all the objects added during the birth step, and only these ones, have been killed during the death step.

Note that the death step is performed after having sorted the objects. We first propose to eliminate points corresponding to a low response of the associated local filter, which helps to escape from local minima of the energy. This heuristic does not change the convergence properties but allow a smarter scanning of the configuration space.

### 5. RESULTS

In this section, we present some results on different types of data: a synthetic image (see Fig. 5), an aerial image (see Fig. 6) with 50cm resolution and a satellite image (see Fig. 6) with 10m resolution. The presented results consists in the black points (point configuration maximizing the density of the point process), whereas the gray lines represent non-zero interactions between points.

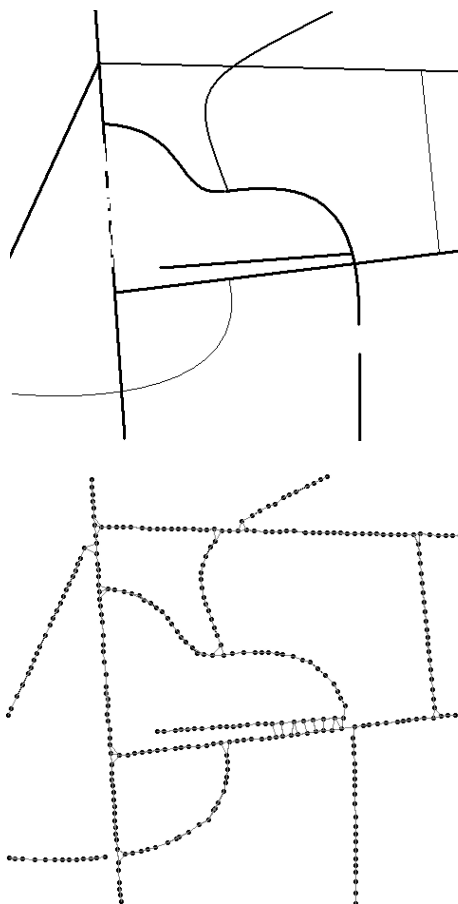
		$E_0 > 1 - U_{of}$
		$2E_1 + 3E_2 + 2U_- > 2 + 2U_{of}$
		$3E_1 + 4E_2 + 3U_- > 4 + 4U_{of}$
		$E_2 + U_- < 2 + 2U_{of}$

**Table 1.** Examples of local configurations: the energy associated with the configuration in the left column should be higher than in the middle column. The right column gives the corresponding constraint on parameters.

Note that the different parameters of the model have been calibrated, first by reducing the parameter space by computing the energy associated with local configurations, which provide some inequalities between the different parameters (see table 5), and then by trial and errors on the synthetic image. On table 5, we compare local configurations plotted in the left column with local configurations plotted in the middle column. For all cases except in the last line, the configurations in the middle column is preferred, whatever the data are. For the case described in the last line, the configuration in the middle column is preferred if the LRD does not detect a road under the dashed line. The corresponding inequalities between the parameters are given in the right column. The same values have been used for the real data (both aerial and satellite images), which demonstrates the robustness of these values. These values are the following;  $U_{of} = 0.3$ ,  $U_- = -1.0$ ,  $U_+ = 10$ ,  $E_1 = 2.0$  and  $E_2 = 1.5$ . The average distance between points is taken as  $D = 15$  whereas  $\theta_0 = \arccos(0.3)$ .

On the final results, we can notice that points are added to the configuration even in the parts of the network where data are missing (see roads in the top of the image on Fig. 6) due to tree shadow or in the omitted part of the synthetic image (see Fig. 5). Junction and close parallel roads are also restored. From these results (points and interaction), a post-treatment, consisting in selecting the real road network from the different interactions, remains to be done. Figure 6 shows the result obtained with the Quality Candy model proposed in [7] on the image presented on figure 6. This model, based on segments, presents a slightly better behavior for connecting the network in case of large occlusion, compared to the result we obtained on figure 6. However, junction are not maken

into account and are not detected. Moreover, the computation time using the Quality Candy model is 23 minutes against 4 minutes for the proposed model.



**Fig. 3.** Synthetic image (top) and extracted road network, the links represents interactions between points of the solution bottom).

## 6. CONCLUSIONS

In this paper, we have proposed a new fast algorithm for road network extraction. We considered a point process given by prior information on connectivity, curvature and junctions of a road network. The data term is a result of a new local road detector. This new model performs well for recognition of junctions and parallel roads as well as for the reconstruction of hidden parts of roads. The optimization based on a recently proposed multiple birth and death algorithm allows the results to be computed in a few minutes, that is much faster than other RJMCMC based algorithms, e.g. Quality Candy model, etc. Finally, the parameter calibration has shown some robustness with respect to images of different sensors.

## СПИСОК ЛИТЕРАТУРЫ

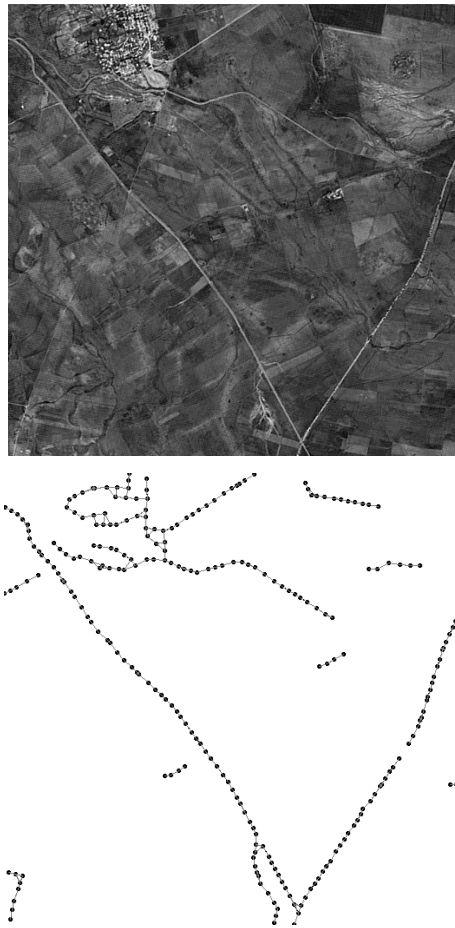
1. Barzohar M., Cooper D.B.: Automatic finding of main roads in aerial images by using geometric-stochastic models and estimation, *IEEE Trans. on Pattern Analysis and Machine Intelligence* **18** (1996) 707–721



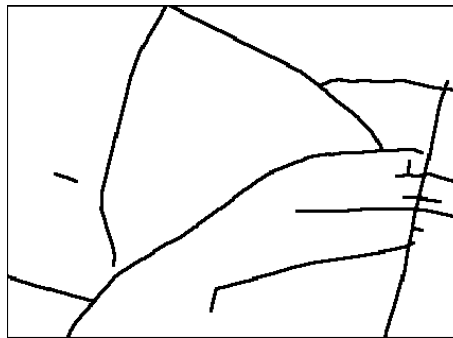


**Fig. 4.** Aerial image © IGN (top) and extracted road network, the links represents interactions between points of the solution (bottom).

2. Descombes X., Minlos R.A., Zhizhina E.: Object extraction using stochastic birth-and-death dynamics in continuum, *Journal of Mathematical Imaging and Vision* **33(3)** (2009) 347–359
3. Fischler M.A., Tenenbaum J.M., Wolf H.C.: Detection of roads and linear structures in low-resolution aerial imagery using a multisource knowledge integration technique, *Computer Graphics and Image Processing* **15** (1981) 201–223
4. Geman D., Jedynak B.: An active testing model for tracking roads in satellite images, *IEEE Trans. on Pattern Analysis and Machine Intelligence* **18** (1996) 1–14
5. Green, P.: Reversible Jump Markov Chain Monte Carlo computation and Bayesian model determination, *Biometrika* **57(1)** (1995) 711–732
6. Grun A., Li, H.: Road extraction from aerial and satellite images by dynamic programming, *ISPRS Journal of Photogrammetry and Remote Sensing* **50(4)** (1995) 11–20
7. Lacoste, C., Descombes, X., Zerubia, J.: Point Processes for Unsupervised Line Network Extraction in Remote Sensing, *IEEE Trans. on Pattern Analysis and Machine Intelligence* **27(10)** (2005) 1568–1579
8. Merlet N., Zerubia J.: New prospects in line detection by dynamic programming, *IEEE Trans. on Pattern Analysis and Machine Intelligence* **18(4)** (1996) 426–431
9. Neuenschwander W.N., Fua P., Iverson L., Szekely G., Kubler O.: Ziplock snakes, *International Journal of Computer Vision* **25(3)** (1997) 191–201
10. Peng, T., Jermyn, I.H., Prinnet, V., Zerubia, J.: Incorporating generic and specific prior knowledge in a multi-scale phase field model for road extraction from VHR images, *IEEE Trans. on Geoscience and Remote Sensing* **1(2)** (2008) 139–146
11. Srivastava, A., Miller, M., Grenander, U.: Multiple target direction of arrival tracking, *IEEE Trans. on Signal Processing* **43(5)** (1995) 282–285



**Fig. 5.** SPOT image © Spotimage/CNES (top) and extracted road network, the links represents interactions between points of the solution (bottom).



**Fig. 6.** Extracted network using the Quality Candy model [7] from the aerial image presented on figure 6.

12. Tupin, F., Maitre, H., Mangin, J.-F., Nicolas J.M., Pechersky, E.: Detection of Linear Features in SAR Images: Application to the Road Network Extraction, *IEEE Trans. on Geoscience and Remote Sensing* **36(2)** (1998) 434-453
13. Vosselman G., de Knecht J.: Road tracing by profile matching and Kalman filtering, in *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Ascona, Switzerland, (1995) 265–274
14. Wand D., He D., Wang L., Morin D.: Road urban network extraction from HRV SPOT images, *International journal of Remote Sensing* **17(4)** (1996) 827-833