

Graph Cuts Approach to MRF Based Linear Feature Extraction in Satellite Images

Anesto del-Toro-Almenares^{1,2}, Cosmin Mihai¹,
Iris Vanhamel¹, and Hichem Sahli¹

¹ Vrije Universiteit Brussel, Dept. of Electronics and Informatics,
VUB-ETRO Pleinlaan 2, B-1050, Brussels-Belgium

² Univ. Central de Las Villas, Center for Studies on Electronics and
Information Technologies, Carr. a Camanjuaní Km 5 $\frac{1}{2}$,
CP-54830, Villa Clara - Cuba
{aalmenar, cmihai, iuvanham, hsahli}@etro.vub.ac.be
<http://www.etro.vub.ac.be>

Abstract. This paper investigates the use of graph cuts for the minimization of an energy functional for road detection in satellite images, defined on the Bayesian MRF framework. The road identification process is modeled as a search for the optimal binary labeling of the nodes of a graph, representing a set of detected segments and possible connections among them. The optimal labeling corresponds to the configuration that minimizes an energy functional derived from a MRF probabilistic model, that introduces contextual knowledge about the shape of roads. We formulate an energy function modeling the interactions between road segments, while satisfying the regularity conditions required by the graph cuts based minimization. The obtained results show a noticeable improvement in terms of processing time, while achieving good results.

Keywords: road detection, graph cuts, MRF-MAP labeling.

1 Introduction

Several approaches have been proposed for linear feature extraction, most of them dealing with the problem of road extraction in either synthetic aperture radar (SAR) images or optical images (visible range). These approaches can be classified according to the preset objective, the extraction technique applied, the type of sensor used, etc. [1].

Most of the reported schemes are based on two criteria: a local criterion, involving the use of local operators, and a global criterion, incorporating additional knowledge about the structure of the objects to be detected. The methods based on local criteria evaluate local properties on the image by using either an edge or line detector [2] or morphological operators [3]. The performance of these methods can be greatly increased by using techniques that introduce some global constraints in the image analysis process. These techniques lead to an optimal solution through the minimization of a cost function by using dynamic

programming [4], tracking methods [5] or the Bayesian MRF framework [2], [6], [7], [8].

Previous work at VUB-ETRO [8] proposed a method that combines both local and global criteria for the identification of the medial axis of roads and paths in satellite images. The approach consists of two steps Fig. 1.

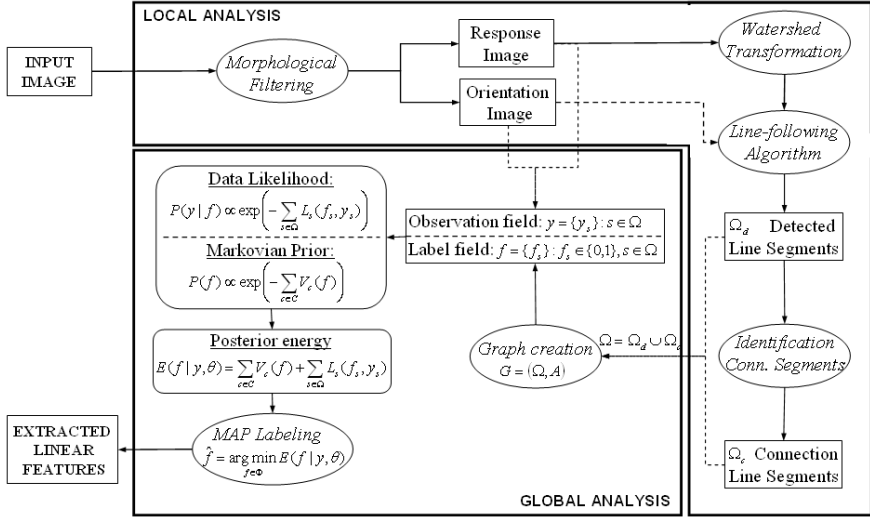


Fig. 1. General algorithm of linear features detection [8]

During a *local analysis*, the detection of elongated structures is performed using a set of soft morphological operators, and a dedicated algorithm is employed for line segment extraction. This results in a set of segments, $\Omega = \Omega_d \cup \Omega_c$, with Ω_d the extracted line segments and Ω_c the set of all possible connections between the segments of Ω_d . The elements of Ω are then organized in a graph $G = (\Omega, A)$, where each node $s \in \Omega$ is a line segment, to which we associate a normalized segment length ($l_s \in [0, 1]$), an observation value y_s (defined below), and a label $f_s = 1$ if s belongs to a road, $f_s = 0$ otherwise. An arc, $A_{st} \in A$, between two nodes s and t , correspond to a shared extremity. To each arc A_{st} is associated the angle θ_{st} between the two segments.

A segment linking process is then performed in the *global analysis*. This is based on the Bayesian framework incorporating an observation measure that reflects the likelihood value of each segment as belonging or not to a road, $L(f_s, y_s)$, and a formulation of the potential functions, $V_c(f)$, which describes the probability distributions of the prior model. The identification of the roads is carried out with an appropriate labeling of the graph $G = (\Omega, A)$, in accordance with the observation process $y = (y_1, \dots, y_m)$; $m = |\Omega|$. A Markov random field (MRF) is defined on the graph and the optimum configuration (labeling) $f = (f_1, \dots, f_m)$, of the segments of Ω given the observation process y , can be estimated based on the Bayes rule and a MAP criterion that maximizes the *posterior probability*

$P(y|f)$. The *conditional probability distribution* $p(y|f)$ depends on the observation measurements, whereas the *prior probability of labelings* $P(f)$ is based on a Markovian model of road-like objects. From the equivalence between MRF and Gibbs fields [11], both of them can be described with a set of potentials that associate an energy function to the different configurations. The minimization of this energy function gives the optimal solution to the problem.

One of the major drawbacks of the approach in [8] is the time required for solving the MAP estimator by means of the simulated annealing algorithm, which requires exponential time in theory and is extremely slow in practice. The computational task of minimizing the energy associated to a particular problem is usually quite difficult, as it generally requires minimizing a non-convex function in a space with thousands of dimensions. However, recently a new approach for energy minimization has been developed based on graph cuts [9]. The basic idea is to construct a specialized graph for the energy function to be minimized such that the minimum cut on the graph also minimizes the energy (either globally or locally). The minimum cut, in turn, can be computed very efficiently by max flow algorithms [10].

This paper investigates the use of graph cuts for the minimization of an MRF based energy functional of [8]. To this end, we reformulate the Bayesian-MRF framework previously described to fit the graph cuts theory. This energy form turns out to be sufficient to model all the interactions between road segments, while satisfying the regularity conditions required by the graph cuts based minimization.

The paper is organized as follows. Section 2 defines the MAP function to be minimized (subsection 2.1), summarizes the graph cuts theory (subsection 2.2), and formulates the functional of subsection 2.1 in terms of graph cuts (subsection 2.3). Section 3 describes the experiments conducted and a preliminary discussion of the obtained results. Finally, Section 4 exposes the final conclusions and addresses future improvements for the proposed approach

2 Materials and Methods

2.1 MRF-MAP Formulation for Linear Feature Extraction

The MAP-MRF estimation belongs to the general family of variational methods, where the main objective is the minimization of an energy functional that conveys the dependencies on observed data and a series of a priori constraints, according to the MAP criterion [11]. The MAP-MRF framework facilitates the formulation of such an energy term by considering certain independence assumptions for the data likelihoods and introducing prior knowledge in the form of Gibbs distributions. The energy functional for linear feature extraction, defined in [8], is stated in Fig. 1, being θ the set of model parameters. From this expression, the MAP estimation is obtained to be:

$$\hat{f}^{MAP} = \arg \min_{f \in \mathbb{F}} \left(\sum_{s \in \Omega} L(f_s, y_s) + \sum_{c \in \mathcal{C}} V_c(f) \right). \quad (1)$$

where \mathbb{F} is the space of all possible configurations $f = \{f_s\} : s \in \Omega, f_s \in \{0, 1\}$. \mathcal{C} is the set of all the cliques or clique space of the model. For each detected segment, two cliques are added to \mathcal{C} ; those conformed by the connection segments sharing one extremity with the detected segment. For more details refer to [8]. The observation, y_s , of a segment s is a function of a saliency measure r_s defined as:

$$r_s = \frac{\overline{R}_s}{|\theta_s - \overline{\alpha}_s| + 1}, \tag{2}$$

where θ_s is the line segment orientation, \overline{R}_s and $\overline{\alpha}_s$ are the mean values, along the line segment, of the response and orientation images, respectively, obtained using soft morphological operators as described in [8] (Fig. 1). The observation values y_s are defined as:

$$y_s = \max_{t \in \mathcal{N}_s} \left\{ \frac{(r_s + r_t)}{2} \right\}, \tag{3}$$

where \mathcal{N}_s is the neighborhood of segment s . After computed, the y_s are linearly normalized to fit the $[0, 1]$ interval. Analogously:

$$L(f_s, y_s) = (1 - y_s) \left(\min\left(\frac{y_s}{\tau}, 1\right) + \log Z_0 \right). \tag{4}$$

where τ is a model parameter to threshold observation values, and Z_0 is a normalization factor found to be equal to $Z_0 = (1 - \tau)\left(\frac{1}{e}\right) - \tau\left(\frac{1}{e} - 1\right)$, with $e = \exp(1)$. The potential functions $V_c(f)$ are defined for each clique according to its size and composition (number and types of segments that conform the clique) as follows:

$$V_c(f_s) = \begin{cases} 0 & \text{if } \forall s \in c, f_s = 0; \\ K_1 + 1 - l_s + \log Z_0 & \text{if } \exists! s \in c \wedge s \in \Omega_d : f_s = 1; \\ \sin(\theta_{st}) + 1 - l_s + l_t + 2 \log Z_0 & \text{if } c1; \\ K_2 \sum_{s: s \in c} f_s & \text{otherwise.} \end{cases} \tag{5}$$

where $c1 \equiv \exists!(s, t) \in c \times c \wedge s \in \Omega_d, t \in \Omega_c : f_s = f_t = 1$. K_1 and K_2 are parameters of the labeling model, defined by the prior road assumptions, and l_s denotes the normalized length of a segment s .

2.2 Energy Minimization Using Graph Cuts

In [9], a systematic and general formulation for energy minimization using graph cuts was presented. The energy form is mainly restricted to functions of binary variables, although it is easily applicable to problems that involve large numbers of labels. Next, we describe the type of energy we are interested to minimize using graph cuts, according to the method described in [9]. Let $x = (x_1, \dots, x_m) : x_s = \{0, 1\}$ be a set of binary-value variables. The type of energy considered has the form:

$$E(x) = \sum_s E_s(x_s) + \sum_{s < t} E_{s,t}(x_s, x_t). \quad (6)$$

That is, the sum of functions involving up to two binary values at a time. A more general energy form, including functions involving up to three binary values, is described in [9].

In order to minimize $E(x)$ (Eq. 6) using graph cuts, a specialized energy graph is created such that the minimum cut on the graph also minimizes $E(x)$ (either globally or locally). The form of the graph depends on the exact form of $E(x)$ and on the number of labels. Let $G = (V, \mathcal{A})$ be a directed graph with non negative edge weights that have two special vertexes (terminals), namely, the source p , and the sink q . A p - q -cut referred simply as cut, $R = \{P, Q\}$ is a partition of the vertexes of V into two disjoint sets P and Q , such that $p \in P$ and $q \in Q$. The total cost of the cut is the sum of the cost of all edges that go from P to Q . The minimum p - q -cut problem is to find a cut R with the smallest cost. The solution to this problem is equivalent to computing the maximum flow from the source to the sink [9]. There are several algorithms that solve this problem in polynomial time with small constants [10]. A cut R is conveniently associated with a labeling f , mapping from the set of vertexes $V - \{p, q\}$ to $\{0, 1\}$, where $f_s = 0$ means that $s \in P$ and $f_s = 1$ means that $s \in Q$. Note that a cut is a binary partition of a graph viewed as a labeling; it is a binary-valued labeling.

The necessary conditions for an energy function to be minimized using graph cuts are described in terms of graph representability, which is conditioned by the regularity of the terms conforming the energy $E(x)$. The regularity condition that must be satisfied is stated as follows:

$$E_{s,t}(0, 0) + E_{s,t}(1, 1) \leq E_{s,t}(0, 1) + E_{s,t}(1, 0). \quad (7)$$

For a given regular energy function $E(x)$, of the form in Eq. (6), the construction of the energy graph is done for each term separately and then all the sub-graphs merged together. The graph will contain $m + 2$ vertexes: $V = \{p, q, v_1, \dots, v_m\}$; p and q are terminal vertexes, and the non-terminal vertexes v_s will encode the binary variable x_s . For a detailed description of the graph construction process refer to [9].

2.3 Graph Cuts for Linear Feature Extraction

The goal of the graph cuts approach for linear feature extraction is to express the energy of the *global analysis* step, in the form of Eq. (6), while conveying for the assumptions of the prior model. According to the method described in [8], the clique space can contain cliques of size greater than two, since it is possible to have more than two connections at one extremity of a detected segment. This makes infeasible the direct use of graph cuts for the energy form defined by the potential functions described in Eq. (5).

In order to make use of the graph cuts theory, we propose the use of a clique space composed by cliques of size one and two, $\mathcal{C} = \mathcal{C}_1 \cup \mathcal{C}_2$, as follows. \mathcal{C}_1 contains

cliques of size one, defined only by a detected segment and \mathcal{C}_2 contains cliques of size two, conformed by a detected and a connection segment, sharing one extremity. Thus, the prior energy term in Eq. (1) can exhibit the following form:

$$E_{prior}(f) = \sum_{s \in \Omega_d} V_1(f_s) + \sum_{(s,t) \in \Omega_d \times \Omega_c : t \in \mathcal{N}_s} V_2(f_s, f_t). \quad (8)$$

The key facts used to model the prior knowledge rely on the following assumptions [8]: 1) roads are long structures, 2) roads have low curvature, and 3) multiple connections are rare. From the previous assumptions three constraints should be imposed to the prior model: continuity, co-linearity and penalization of multiple connections. $V_1(f_s)$ is used to account for the extremity penalization in the case of isolated segments, and detected segments with a free extremity situated far from the border of the image. Free extremities at the border of the image are though to belong to a long road not captured by the image, thus no extremity penalization is included.

Let $Ex(s) : \Omega \rightarrow \{0, 1, 2\}$ be a function that return the number of free extremities of a segment s , excluding the cases when the extremity is at the border of the image, then:

$$V_1(f_s) = f_s \cdot K_{ex} \cdot Ex(s), \quad (9)$$

where K_{ex} is the extremity penalization model parameter, that accounts for assumption 1). Analogously, the use of size two cliques is intended to penalize non co-linear segments while favoring the co-linear ones, and to include a penalization term for unconnected extremities. The definition of $V_2(f_s, f_t)$, for size two cliques, is presented in Table 1.

Table 1. Formulation of the potential function for size two cliques

(f_s, f_t)	$V_2(f_s, f_t)$
0 0	0
0 1	$+\infty$
1 0	$K_{ex} \cdot w_{t s}^N$
1 1	$K_{co} \cdot \sin \theta_{st}$

Here, K_{co} is a weight to account for assumption 2), by providing a penalty weight for non co-linear segments. It also serves to penalize multiple connections, assumption 3). $w_{t|s}^N$ is a normalized arc measure, defined as the strength of the path determined by the connection segment t in the direction to which it is connected to the detected segment s . This arc measure is computed from the quantities $w_{t|s}$, given by:

$$w_{t|s} = K_{co} \sin \theta_{st} + \begin{cases} (l_s \cdot r_s + l_t \cdot r_t) \cdot (l_s + l_t)^{-1} & \text{if } l_s > l_t; \\ 0.5 \cdot (r_s + r_t) & \text{otherwise.} \end{cases} \quad (10)$$

The quantities $w_{t|s}$, after computed, are normalized locally to fit in the $[0, 1]$ interval to obtain $w_{t|s}^N$. This normalization is accomplished locally among the connection segments sharing the same extremity of a detected segment. $V_2(f_s, f_t)$ is defined based on the belief that every time a connection segment is labeled as zero then the likelihood that the corresponding detected segment is isolated is increased. Note that in the case of a configuration labeling of a connection segment without its corresponding detected segment, a high penalization is included. This corresponds to the analysis that connection segments will not be included without the evidence of a detected segment in one extremity.

From equations (4), (8), (9) and Table 1, we can express the proposed graph cut energy in the form of Eq. (6) as follows:

$$E(f) = \sum_{s \in \Omega} [L(f_s, y_s) + V_1(f_s)] + \sum_{(s,t) \in \Omega_d \times \Omega_c: t \in \mathcal{N}_s} V_2(f_s, f_t). \quad (11)$$

The proof of the required regularity condition, Eq. (7), for the current definition of the clique potentials is straightforward as shown in Eq. (12):

$$\begin{aligned} V_2(0, 0) + V_2(1, 1) &\leq V_2(1, 0) + V_2(0, 1) \\ 0 + K_{co} \cdot \sin \theta_{st} &\leq +\infty + K_{ex} \cdot w_{t|s}^N \end{aligned} \quad (12)$$

3 Results and Discussion

In order to validate the proposed approach, some experiments were conducted using images corresponding to the blue channel of two scenes from IKONOS imagery, given in 2.a and Fig. 3.a. Images are available at (<http://www.bauv.unibw-muenchen.de/institute/inst10/euroedr/>).

The experiments were carried out as follows: for each image, a graph containing the potential road network was obtained, using the *local analysis* step described in [8] (incises b). Based on this graph, the *global analysis* step (graph labeling) was accomplished using the proposed approach for different sets of parameters settings (incises c and d). Fig. 2.b illustrates the detected candidate road segments of Fig. 2.a, composed of 630 and 441 detected and connection segments, respectively; Fig. 2.c and Fig. 2.d show the obtained road labeling, with parameters $\tau = 0.165$, $K_{co} = 0.01$, $K_{ex} = 0.13$, and $\tau = 0.15$, $K_{co} = 0.008$, $K_{ex} = 0.135$, respectively.

Fig. 3.b illustrates the detected candidate road segments of Fig. 3.a, composed of 870 and 704 detected and connection segments, respectively; Fig. 3.c and Fig. 3.d show the obtained road labeling, with parameters $\tau = 0.15$, $K_{co} = 0.09$, $K_{ex} = 0.12$, and $\tau = 0.155$, $K_{co} = 0.09$, $K_{ex} = 0.135$, respectively.

In both cases, the proposed approach using graph cuts takes less than 1 second for the optimization while the time required by the method reported in [8] takes up to 10 minutes. As can be visually noticed, the obtained results are

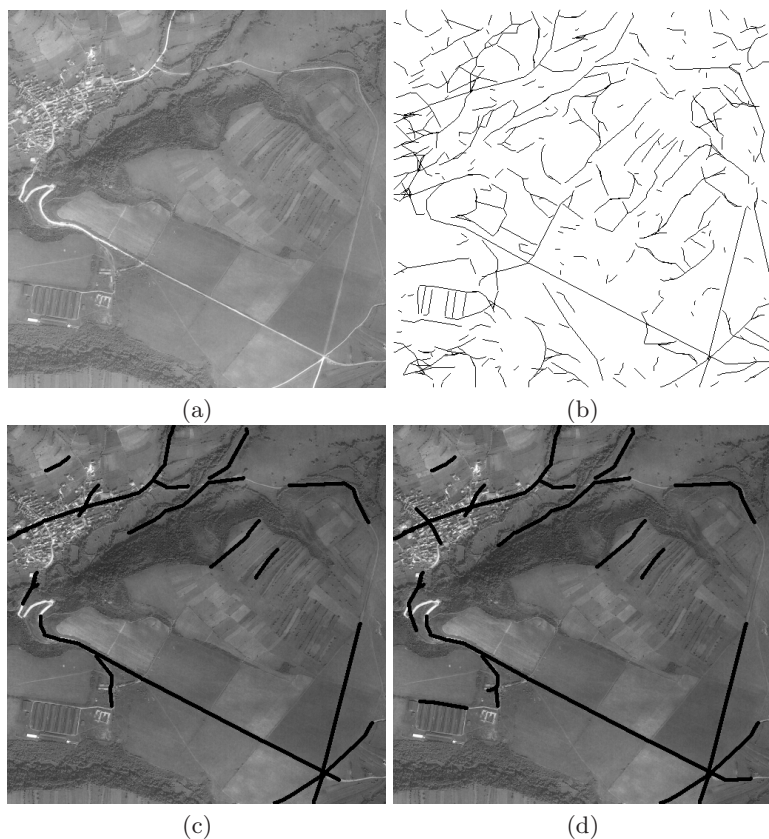


Fig. 2. Image Ikonos1: a) original image; b) candidate road segments; c) road labeling composed by 63 and 51 detected and connection segments; d) road labeling composed by 55 and 45 detected and connection segments

acceptable in terms of false alarms and correct detection of roads. Some false alarms corresponds to the detection of elongated linear structures other than roads, *e.g.* ridges and bright terrain patches. These erroneous detections are thought to be favored by the employed observation model [8], and its ability to effectively describe road and non road linear structures (likelihood value).

Note that the resulting road labeling, is always limited by the quality and completeness of the potential road network, obtained in the *local analysis* step. The setting of the parameters has been made empirically, using the fact that the parameter τ depends on the intensity of the roads. The setting of K_{co} , as expected, was found to be related with the presence of curvilinear roads. Lower values of K_{co} favors the labeling of segments corresponding to curved paths, since its reduces the penalization for non co-linear segments.

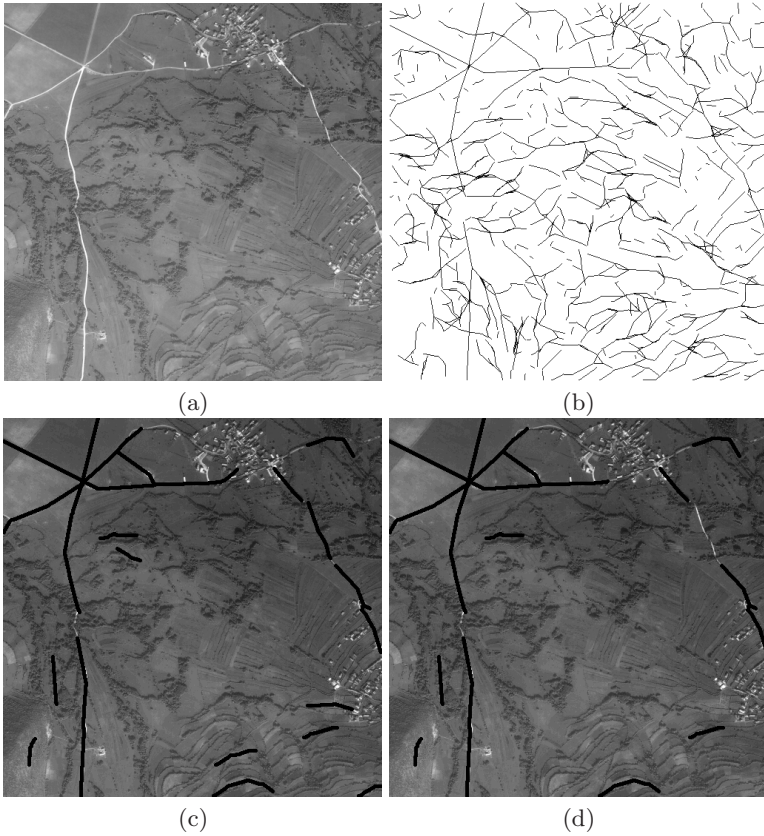


Fig. 3. Image Ikonos3: a) original image; b) candidate road segments; c) road labeling composed by 65 and 49 detected and connection segments; d) road labeling composed by 50 and 39 detected and connection segments

4 Conclusions and Future Work

Despite of its simplicity, the proposed approach shows a noticeable improvement in terms of processing time, while achieving good results. Although the obtained results were satisfactory, these are only preliminary. Improvements should address the observation model, and take into account curvilinear roads as well as the automatic estimation of the involved parameters, by deriving relationships among them. In the future, we aim to tackle these issues and to conduct several experiments to make a comparative analysis with other methods and different optimizers in terms of computational burden and performance.

Acknowledgment

This work has been partially funded by the EU-IST project STREAM Technology to Support Sustainable Humanitarian Crisis Management (contract EU-IST-FP6-2-511 705), and the IncGEO VERA project.

References

1. Mena, J.B.: State of the art on automatic road extraction for GIS update: a novel classification. *Pattern Recognition Letters*. 24(16), 3037–3058 (2003)
2. Tupin, F., Maitre, H., Mangin, J.F., Nicolas, J.M., Pechersky, E.: Detection of Linear Features in SAR Images: Application to road network extraction. *IEEE Trans. on Geoscience and Remote Sensing*. 36(2), 434–453 (1998)
3. Chanussot, J., Lambert, P.: An application of mathematical morphology to road network extraction on SAR images. In: *Proc. 4th Int. Sym. On Mathematical Morphology and its App. to Image and Signal Processing*, Amsterdam, pp. 399–406 (1998)
4. Merlet, N., Zerubia, J.: New prospects in line detection by dynamic programming. *IEEE Trans. Pattern Anal. Mach. Intell.* 18(4), 426–431 (1996)
5. Geman, D., Jedynek, B.: An active testing model for tracking roads in satellite images. *IEEE Trans. Pattern Anal. Mach. Intell.* 18(1), 1–14 (1996)
6. Tupin, F., Houshmand, B., Datcu, M.: Road detection in dense urban areas using SAR imagery and the usefulness of multiple views. *IEEE Trans. on Geoscience and Remote Sensing*. 40(11), 2405–2414 (2002)
7. Katartzis, A., Pizurica, V., Sahli, H.: Application of mathematical morphology and Markov random field theory to the automatic extraction of linear features in airborne images. In: *Proc. 5th Int. Sym. On Mathematical Morphology and its Applications to Image and Signal Processing*. California, pp. 405–414 (2000)
8. Katartzis, A., Sahli, H., Pizurica, V., Cornelis, J.: A model-based approach to the automatic extraction of linear features from airborne images. *IEEE Trans. on Geoscience and Remote Sensing* 39(9), 2073–2079 (2001)
9. Kolmogorov, V., Zabih, R.: What energy functions can be minimized via graph cuts? *IEEE Trans. Pattern Anal. Mach. Intell.* 26(2), 147–159 (2004)
10. Boykov, Y., Kolmogorov, V.: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Trans. Pattern Anal. Mach. Intell.* 26(9) (2004)
11. Li, S.Z.: *Markov random field modeling in computer vision*, 1st edn. Springer-Verlag, New York Berlin Heidelberg Tokyo (1995)