To appear in the *International Journal of Remote Sensing*
Vol. 00, No. 00, Month 20XX, 1–30

IJRS Research Paper

A Two-level Markov Random Field for Road Network Extraction and its Application with Optical, SAR and Multitemporal Data

T. Perciano[a*], F. Tupin[b†], R. Hirata Jr.[c‡] and R. M. Cesar Jr.[c§]

[a]Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley-CA, 94720, US

[b]Signal and Image Processing Department of Télécom ParisTech, LTCI, 75634 Paris Cedex, France

[c]Institute of Mathematics and Statistics of University of São Paulo, São Paulo, SP 05508-090, Brazil

(*Received 00 Month 20XX; accepted 00 Month 20XX*)

This paper introduces a method for road network extraction from satellite images. The proposed approach covers a new fusion method (using data from multiple sources) and a new Markov random field (MRF) defined on connected components along with a multilevel application (two levels MRF). Our method allows the detection of roads with different characteristics and decreases by around 30% the size of the used graph model. Results for synthetic aperture radar (SAR) images and optical images obtained using the TerraSAR-X and Quickbird sensors, respectively, are presented demonstrating the improvement brought by the proposed approach. In a second part, an analysis of different types of data fusion combining optical/radar images, radar/radar images and multitemporal SAR (TerraSAR-X and COSMO-SkyMed) images is described. The qualitative and quantitative results show that the fusion approach improves considerably the results of the road network extraction.

Keywords: Road detection; Markov Random Fields (MRF's); SAR data; optical data; multitemporal data; data fusion; connected components

*Corresponding author. T. Perciano was with the Signal and Image Processing Department of Télécom ParisTech, 75634 Paris Cedex, France and also with the Institute of Mathematics and Statistics of the University of São Paulo when this work was developed. Email: tperciano@lbl.gov. Phone: (+1) 5104805060
†Email: florence.tupin@telecom-paristech.fr. Phone: (+33) 145817245
‡Email: hirata@ime.usp.br. Phone: (+55) 1130915376
§Email: cesar@ime.usp.br. Phone: (+55) 1130916135
This work was supported by CAPES PDEE under Grant #0310-10-7; FAPESP under Grant #2011/50761-2; EFIDIR under Grant #ANR-2007-MCDC0-04; and NAP eScience - PRP - USP

1. Introduction

Since the last few years, new optical (Quickbird, Pleiades) and synthetic aperture radar (SAR) (TerraSAR-X, COSMO-SkyMed) sensors that provide high-resolution images have been launched. The arrival of these sensors opens new perspectives for pattern recognition problems and, specifically, for road network extraction. Indeed, new approaches exploring the characteristics of these high-resolution sensors have been proposed in the literature (Poulain, Inglada, and Spigai 2008; Tison et al. 2004; Negri et al. 2006; Hedman et al. 2010; Lacoste, Descombes, and Zerubia 2010; Türetken et al. 2013; Miao et al. 2015a,b). Moreover, the need of multi-sensor image processing methods to tackle problems related to scene interpretation is increasing (Sportouche, Tupin, and Denise 2009; Poulain et al. 2009; Brunner, Lemoine, and Bruzzone 2008; Tupin and Roux 2003; Hinz and Baumgartner 2003).

The problem of road network extraction in remote sensing images has been studied for a long time because of its importance in many applications such as urban planning, map making and updating, traffic management, industrial development, cartography and so on. Besides, the volume of collected imagery is increasing rapidly in recent years and, consequently, manual processes become very time consuming. This problem is considered a very difficult task mainly because the spatial and spectral features of the roads can be very complex and variable. Despite of what may be inferred, using images with better resolution does not make the problem easier. Indeed, the presence of many objects caused by vehicles, traffic signs and buildings, usually mask parts of the roads present in the image.

There are many works in the literature proposing different approaches, either automatic or semi-automatic. An approach that is used in general is a two-step analysis. The first level is dedicated to a feature extraction step and the second level involves some methods for grouping analysis, i.e., a high level step aiming to extract the final road network by using structural/contextual information. Concerning the first step, different methods proposed in the literature use techniques such as edge detection (Zhou, Venkateswar, and Chellappa 1989; Canny 1986; Steger 1998), road mask filters (Gamba, Dell'Acqua, and Lisini 2006), adaptive directional filtering (Dell'Acqua, Gamba, and Lisini 2005), statistical analysis (Skriver et al. 2005), morphological operators (Chanussot, Mauris, and Lambert 1999; Katartzis et al. 2001) and others. In the case of the second step, there are works in the literature using minimisation of a global cost function applying Markov random fields (MRF) (Tupin et al. 1998; Katartzis et al. 2001; Negri et al. 2006; Tupin, Houshmand, and Dactu 2002), tracking methods (Vosselman and Knecht 1995; Zhou, Bischof, and Caelli 2005) and dynamic programming (Barzohar and Cooper 1996). Other works also use genetic algorithms for this step (Jeon, Jang, and Hong 2002).

Concerning data fusion, Hall and Llinas introduced the multi-sensor data fusion (Hall and Llinas 1997), where some applications, process models and identification of applicable techniques were covered. When it comes to image fusion to tackle the problem of road network extraction, some works can be found in the literature. Hedman et al. (2010) and Lisini et al. (2006) used feature-level fusion in order to improve road extraction in high-resolution images. Lisini et al. propose an approach using optical and SAR data for extracting roads with different characteristics and provides the possibility to make the fusion of results coming from different sources. Huang et al. propose a road line extraction method based on a new feature fusion approach, i.e., the complementary information of two kinds of line feature is used. A fusion method was also used by Lisini, Gamba, and Luebeck (2011) where the backscattering characteristic of X-band and P-band are employed

to extract roads from airborne dual-band SAR image. Recently, a new stereoscopic road network extraction framework based on the decision-level fusion of optical and SAR imagery was proposed by He et al. (2013).

Although many achievements have been made on the extraction of road networks from SAR imagery, this kind of data is seriously affected by related noise, which makes difficult to obtain good results using existent methods in the literature (Sun et al. 2014). There are still problems to be tackled due to the high complexity and difficulty of this task. Some very recent efforts in this sense can be found such as the work by Lu et al. (2014), where the authors propose a new method based on the region growing algorithm, which is suitable for different resolution SAR images. In this method, a weighted ratio line detector is used as a first step and the region growing approach is used as the second step to optimise the network. The work by He et al. (2014) uses a multi-scale linear feature detector and beamlet framework. Improvements for the representation of circle roads are also proposed. Even though new methods for road extraction from SAR images have been proposed such as the ones described before, there are not very recent efforts to improve fusion based methods.

Regarding graphical models, the use of graphs have been emerging rapidly as a consolidated representation for image processing and analysis (Lezoray and Grady 2012). Indeed, many different concepts can be defined using graphs and a large number of real-world problems can be modelled using this approach. In particular, probabilistic graphical models have found many developments and applications in different areas of research, due to their appropriateness to represent discrete data by modelling neighbourhood relationships. Besides, these models also provide an interesting way to define theoretical algorithms for the processing of functions related to graphs. Graphical models have become a unified representation for image processing and analysis, mainly because of its discrete nature and mathematical simplicity.

Graphical models, specifically MRFs and related models, have been used to tackle main problems on image processing such as segmentation/classification, image registration, and feature detection. If we consider works related to satellite images, some very recent works can be found in the literature. Salberg and Trier (2012) propose a hidden Markov model for the analysis of multi-source time series or remote sensing images in order to detect changes in the spatial coverage of forests. Moser and Serpico (2012) propose a novel unsupervised change detection technique based on MRFs, line processes, and a dictionary of SAR-specific probability density models. Different techniques for image classification can be found in (Xu et al. 2013; D'Elia et al. 2014; Hoberg et al. 2015). Xu et al. use random fields of latent topics to propose an efficient unsupervised semantic classification method for high-resolution satellite images. D'Elia et al. also explore the problem of classification using features extracted from a segmentation algorithm based on tree structured Markov random fields (TS-MRFs). Finally, Hoberg et al. propose a method for multi-temporal and contextual classification of remote sensing images based on Conditional Random Fields (CRFs). Works using graphical models for feature extraction and registration can be found in (Song, Huang, and Zhang 2014; Papila, Kent, and Kartal 2014; Yang et al. 2014).

In this context, we develop a unified framework that is easy to use and to extend for integrating information of multiple optical and radar satellite sensors (Perciano et al. 2011). The present paper introduces the following original and relevant contributions. The use of connected components to represent road primitives is developed for the first time, simplifying important steps of the detection process and making

the framework more reliable. We highlight that the use of connected components changes completely the modelling behind the framework. Besides, a second high level step is also introduced (two levels MRF). And last but not least, as an applied contribution, data fusion is used in the new method taking advantage of the information provided by different sources (optical, radar and multi-temporal images), improving the road detection results. The aim of the paper is to describe the proposed framework with its different steps and to present experimental results and perspectives/scenarios for its usage. Together, these features improve the results provided by other previous approaches from the literature.

2. Markovian framework for road extraction

2.1. *Introduction*

Figure 1 presents a flowchart that summarises the steps of the proposed approach. We briefly mention all the steps that make up the framework (the details are worked in the next sections) containing two main blocks (dashed lines rectangles):

- The first block is a multi-scale approach including 3 main steps:
  - **Low level step (line detection)**: from the input image (or images from different sources), a process of line detection is carried out in order to find road candidates, which are grouped by connected components (**1**);
  - **Graph construction**: using the connected components found in the previous step, an attributed relational graph is built (**2**);
  - **Graph labelling**: a MRF is defined on the built graph, and a minimisation process is done in order to find the optimal binary labelling of the graph. Here the data coming from different sources is also used (if applied). The labelling is mapped back to the image to find the final road network (**3**).
- The results of each scale are merged together to form the input of the next block:
  - **Road sections level**: the previous merged result is processed to obtain only road sections and crossroads (**4**);
  - **Second high-level step**: a new graph is built and a second minimisation process is carried out which leads to the final road extraction (**5** and **6**).

It can be observed that the proposed fusion approach is done at the low-level step, where the line detection is applied to all the used images. The multiple source data is also used in both high-level steps. In the next sections each step mentioned above is explained in detail.

2.2. *Line detection and low-level fusion*

2.2.1. Line detection

Considering SAR imagery, there are several classical edge detectors in the literature, such as T distribution method (Cook et al. 1994), log-likelihood hypothesis testing method (Caves and Quegan 1994), ratio method (Touzi, Lopes, and Bousquet 1988), ratio line detector and cross-correlation line detector (D1D2) (Tupin et al. 1998), the unbiased detector of curvilinear features (Steger 1998) and so on. Ratio detector (Touzi, Lopes, and Bousquet 1988) provides the relation between the false alarm and the image edge obtained from the statistical properties of the SAR image. This detector is widely used for its statistical based deduction of threshold. The D1D2 detector (Tupin et al. 1998) is based on the ratio detector, as highlighted by R. Touzi. The D1D2 detector proposed by Tupin improves the ratio detector by adding the capability to detect linear objects with a certain width. In this work we use D1D2 detector for the low-level step. This detector is defined as the symmetrical sum (Bloch 1996) of the responses of two detectors, for each pixel of the image:

$$\lambda(r,\rho) = \frac{r\rho}{1-r-\rho+2r\rho}, \tag{1}$$

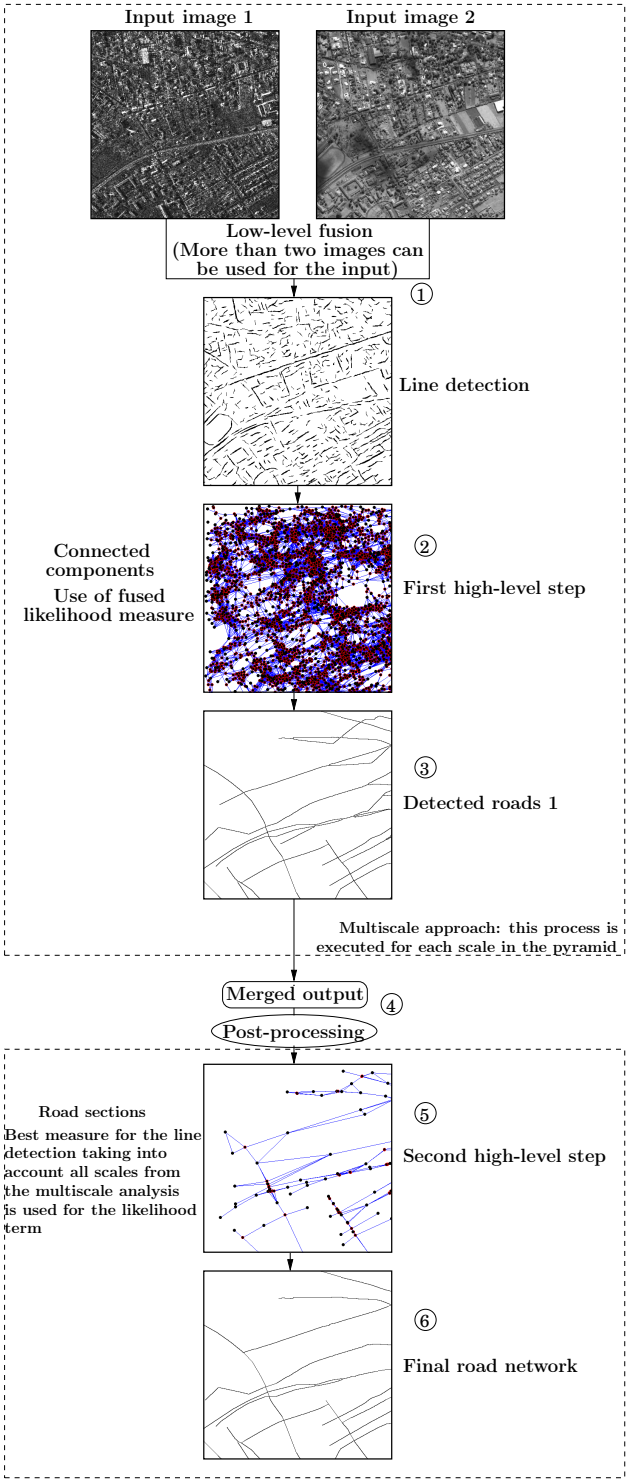


Figure 1.: Flowchart presenting all the steps of the proposed framework. It is composed by two main blocks (dashed lines): a multi-scale approach including a low-level step, a graph construction and a graph labelling; and a second high-level step where only road sections are considered. The framework can be used with images from different sources, from which a low-level fusion approach is carried out and a fusioned likelihood is used for the high-level step.

where r is the D1 detector, a ratio edge detector, which is defined as:

$$r = \min(r_{1,2}, r_{1,3}) \tag{2}$$

with $r_{i,j} = 1 - \min(\mu_i/\mu_j, \mu_j/\mu_i)$, μ_i and μ_j are the empirical mean values of the regions i and j (see Figure 2), respectively, and ρ is the definition of D2 detector, a cross-correlation line detector:

$$\rho = \min(\rho_{1,2}, \rho_{1,3}) \tag{3}$$

with

$$\rho_{i,j} = \sqrt{\frac{n_i n_j (\mu_i - \mu_j)^2}{1 + (n_i + n_j)(n_i \sigma_i^2 + n_j \sigma_j^2)}}, \tag{4}$$

where σ_i^2 denotes the variance of the amplitudes in region i and n_i is the number of pixels in region i. The final response of the detection is a magnitude image which

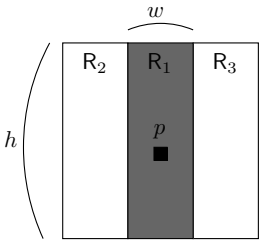


Figure 2.: Line detection mask used for a specific pixel p given a direction (vertical here), a width w and a height h. The ratio and the cross-correlation between region R_1 and each region around it, R_2 and R_3, are calculated and the symmetrical sum in Equation (1) is used to combine the two measures.

is thresholded and thinned, resulting in a binary output.

In this work we also explore optical images for the experiments. In this case, for the low-level detection, the same criterion (D1 and D2 detectors) is used. Although not optimal for optical images, the same detection approach can be applied.

2.2.2. Low level fusion

In the proposed approach we introduce the idea of data fusion using the same procedure for line detection described above. For this purpose, the line detectors are calculated for all images from different sources. Suppose that $\mathrm{I} = \{I_1, I_2, \ldots, I_m\}$ is a set of images of the same region, acquired from different sensors and appropriately registered. After calculating the responses r and ρ for each image in I, a set of responses is obtained, $\Lambda_\mathrm{I} = \{r_{I_1}, \rho_{I_1}, r_{I_2}, \rho_{I_2}, \ldots, r_{I_m}, \rho_{I_m}\}$. The symmetrical sum presented in Equation (1) is used in the following general form to combine all the measures in Λ_I:

$$\Lambda(x_1, \ldots, x_n) = \frac{g(x_1, \ldots, x_n)}{g(x_1, \ldots, x_n) + g(1 - x_1, \ldots, 1 - x_n)}, \tag{5}$$

where $g(x_1, \ldots, x_n) = x_1 x_2 \ldots x_n$. The idea behind this procedure is that the higher the line detector response, the greater the chance that a linear structure is present

in the images. A clear advantage about fusing the detections from different sources is to increase the chance to detect roads that do not appear in all the images. For instance, optical images depend on good weather conditions to clearly capture all the roads, which is not always possible. Adding the information of a radar image of the same area may help solving this problem.

The final result from the detection, after fusing the responses of all images, is a magnitude image with the fused line detection response of each pixel of the image. In order to obtain a binary result, the magnitude image is thresholded and thinned. The thinning process may result in small spur lines, but those are disregarded during the graph modelling process. After this process, all the structures found are mapped to a graph as described in the next section.

2.3. *Graph modelling*

We now describe the approach used to build the proposed graph model. There are different ways to represent an image using a graph: the nodes of the graph can represent pixels, borders, regions or any other structure obtained from an image. In this work, we use a broadly known graphical model, the MRF, and our graph is built using connected components obtained from the low-level step as explained in the following.

2.3.1. Connected components' graph

Differently from the works in the literature, this work proposes to treat the structures detected at the low-level as *connected components*. Starting from the thinned binary image (Figure 3a), all the connected components are detected and (1) for each component all the extremities and crossings are detected and (2) for each crossing, the connected component is subdivided (Figure 3b). This process aims to obtain connected components as curves with only two extremities. Then, all possible connections between the connected components are found (Figure 3c) and a graph is built where each connected component is a node of the graph (including the possible connections) and two nodes of the graph are connected if their corresponding components share an extremity (Figure 3d). Notice that because of the use of the connected components, no step of polygonal approximation, as used in (Tupin, Houshmand, and Dactu 2002) and other works, is needed. Possible connections between the connected components must be found according to some proximity and alignment constraints as explained in details in the next section.

Formally, let $\mathrm{C}_{\mathrm{detected}}$ be the set of connected components detected at the low-level step, considering that each component is a simple curve with two extremities. In a real detection problem, these components may or may not represent real roads in the image. We assume that the final road network is composed by those components detected at the low-level step and by the connections between them, removing appropriately the false detections. Consequently, we consider also the set of all the possible connections between the connected components, $\mathrm{C}_{\mathrm{connections}}$. Let i and j be any two connected components in $\mathrm{C}_{\mathrm{detected}}$, let E_i^k with $k \in \{1, 2\}$ be the extremities of i ($i = \widetilde{E_i^1 E_i^2}$), and let E_j^l with $l \in \{1, 2\}$ be the extremities of j ($j = \widetilde{E_j^1 E_j^2}$) where $\widetilde{AB}$ denotes the connected component joining A and B. Denoting by $i\mathrm{R}j$ the relation of a possible connection (which will be defined in Section 2.3.2), we have:

$$\mathrm{C}_{\mathrm{connections}} = \{\widetilde{E_i^k E_j^l}, i \in \mathrm{C}_{\mathrm{detected}}, j \in \mathrm{C}_{\mathrm{detected}} \text{ and } i\mathrm{R}j\}. \quad (6)$$

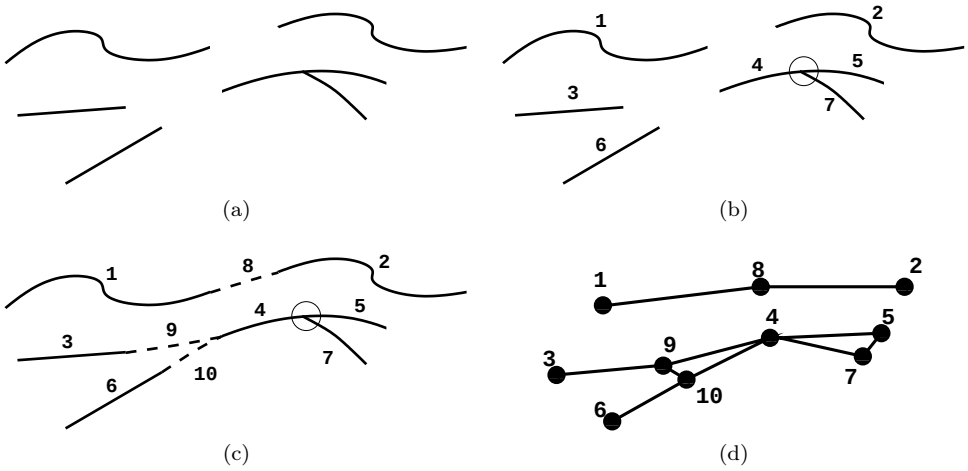


Figure 3.: Proposed graph modelling where connected components are used instead of line segments (numbers represent identified connected components). (a) Example of line detection response after threshold. (b) Detection of connected components and crossings. (c) Identification of the possible connections between connected components. (d) Graph generated from the connected components.

The computation of $\widetilde{E_i^k E_j^l}$ in $\mathrm{C}_{\mathrm{connections}}$ will be described in the next section. The complete set of connected components is given by $\mathrm{C} = \mathrm{C}_{\mathrm{detected}} \cup \mathrm{C}_{\mathrm{connections}}$. This set C is endowed by a graph structure where each node of the graph G represents a connected component and there is an edge between two nodes i and j if their corresponding components share an extremity as presented in Figure 3d. According to the previous construction, the following neighbourhood can be defined for G:

$$\partial_i = \{j \in \mathrm{C} \mid \exists (l,k) \in \{1,2\}, E_j^l = E_i^k, j \neq i\}. \tag{7}$$

Notice that this neighbourhood covers two connection cases: a possible connection between components or a connection by crossroads, as can be observed in Figure 3b. According to this neighbourhood, we define an irregular grid, differently from regular grids defined by 4-connected or 8-connected neighbourhoods for instance. G is an attributed relational graph: each node i of G stores the attribute $\mathcal{L}_i = \min(1, (\mathrm{Length}_i / \mathcal{D}_{\max})) \in [0,1]$, where Length_i is the length of the connected component and $\mathcal{D}_{\max}$ is a normalisation factor (chosen empirically). We introduce also for each node i an attribute of local homogeneity $\mathcal{H}_i = \sigma_i / \mu_i$, where σ_i and μ_i are the standard deviation and the mean of the intensities of the pixels in the region where the connected component is passing through (area around the connected component). Each edge of G connecting two nodes i and j stores the attribute related to the angle between their corresponding connected components, which is denoted by $\alpha = \measuredangle ij$ ($\alpha \in [0, \pi[$). This angle is calculated considering the direction vector at the extremities of each component (see Figure 4). Despite of the disconnected appearance of the components i and j in Figure 4, due to illustration purpose, these components are connected in practice. The cliques of the graph, $\mathrm{C_G}$, are the complete subgraphs of G that correspond to all subsets of connected components sharing an extremity.

The aim of using connected components instead of segments is to maintain, in general, complete structures detected at the low-level without subdividing them into many segments. As a consequence, without those subdivisions, a lower number

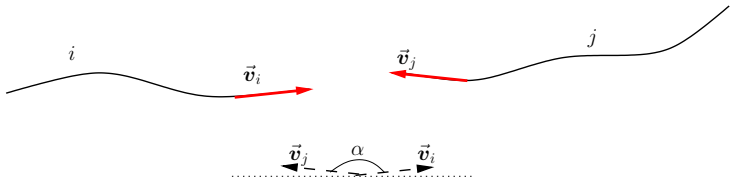


Figure 4.: Example on how to calculate the angle between two connected components i and j ($\measuredangle ij$) considering the direction vectors $\vec{\boldsymbol{v}}_i$ and $\vec{\boldsymbol{v}}_j$ of the extremities of each component.

of primitive structures has to be processed during the next steps of the processing chain and the number of possible false paths is decreased. Besides, the built graph is simplified as the number of nodes decreases considerably. Finally, the use of connected components makes the framework more robust in the sense that it can be applied to the detection of different and more complex structures beside roads, as it is more flexible with the curvature of the structures.

2.3.2. Possible connections - best path approach

As this work considers connected components instead of line segments, the intermediate connections $\widetilde{E_i^k E_j^l}$ (Equation (6)) between connected components are also represented as components.

Firstly, there is a possible connection $i\mathrm{R}j$ between two connected components i and j if the Gestalt principles of proximity and alignment apply, i.e., the two extremities being connected are close enough and if the alignment between the components is acceptable. These two principles aim to obtain connections resulting in ϵ-meaningful events (Desolneux, Moisan, and Morel 2000), i.e., the connection is "visually acceptable" in terms of visual perception.

Secondly, instead of tracing a straight line, the possible connection between two connected components corresponds to the best path between the two closest extremities. The best path is calculated using a dynamic programming algorithm taking into account the intensities of the pixels of the input image. For this purpose, a geodesic distance is used as the work by Perciano, Hirata, and de Castro Jorge (2010). The distance between the two pixels (extremities) p_i and p_j is the minimum path between these two pixels according to the image topology. Let us consider a path $\mathrm{P} = \{p_1, p_2, \ldots, p_n\}$, where p_i and p_{i+1} are connected neighbours and $i \in \{1, 2, \ldots, n-1\}$. The length of P is

$$l(\mathrm{P}) = \sum_{i=1}^{n-1} d_\Omega(p_i, p_{i+1}). \tag{8}$$

The neighbourhood (Ω) used in this work is 8-connected and $d_\Omega(p_i, p_{i+1}) = |\mathcal{I}(p_i) - \mathcal{I}(p_{i+1})|$, where $\mathcal{I}(\cdot)$ refers to the intensity value of a pixel. Even though the minimum path is mainly calculated using the intensity values of the pixels, we also add the information about the angle between the start point and the end point. This means that the algorithm will prefer a path that deviates less the direction towards the end point, unless the intensity variation is too high.

Using this approach, a more accurate connection can be ensured in complex cases where a straight line does not match the searched structure as in the example presented in Figure 5. The connection is improved with the use of connected com-

ponents instead of segments. In Figure 5a, a circular intersection is present in the selected region of interest of a radar image. The line detection result for this region is shown in Figure 5b. The original method used by Tupin et al. traces straight lines for all possible connections between the road primitives as shown in Figure 5c. In doing so, the details of the circular intersection are lost. On the other hand, using the proposed method, all the best paths between the connected components are calculated, thus retaining the shape of the circular intersection, as presented in Figure 5d.

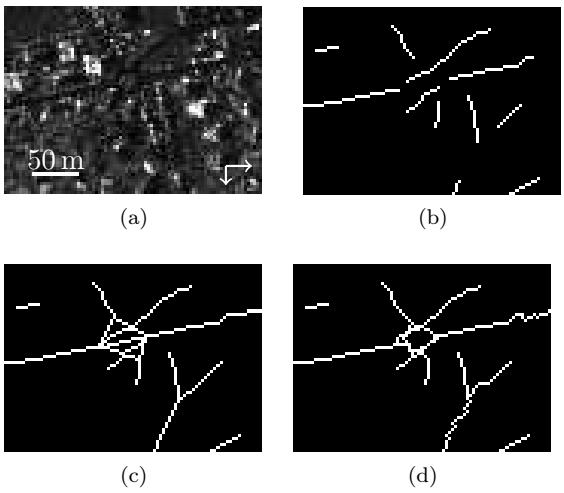


Figure 5.: Example of the application of best path method to obtain the possible connections between the road primitives. (a) Region of interest from an original radar image. (b) Line detection result. (c) Connections obtained using the method by Tupin et al.. (d) Connections obtained using the proposed method.

The definition of the geodesic distance as shown before is flexible in terms of the type of image. In order to apply the same method to multispectral images, for example, the intensity values of the pixels from all (or the most representative) spectral bands can be combined in the calculation of the geodesic distance. The way the values are combined is of great impact, so different methods should be considered depending on the type of image. The most simple approach, for example, would be to consider the mean value of the intensities of all bands and using that value as $\mathcal{I}(p_i)$.

The network reconstruction process is held as a labelling process of the graph G, i.e., the identification of the nodes belonging to a road or not. This process is explained in the following section.

2.4. *Network reconstruction*

2.4.1. Markovian model

Let $\mathrm{L} = \{l_1, l_2 \dots l_N\}$ be the set of all the labels, $l_i (i = 1 \dots N)$, of the nodes (connected components) i of the graph G. L is a binary labelling where $l_i = 1$ if the node i belongs to a road and $l_i = 0$ otherwise. The road network corresponds to the optimal labelling that is obtained by the minimisation of an energy function

derived from a probabilistic model (Tupin et al. 1998). The energy function can be written as:

$$U(\mathrm{L}) = U_{\mathrm{likelihood}}(\mathrm{L}, \mathrm{D}) + U_{\mathrm{prior}}(\mathrm{L}), \tag{9}$$

where D is the set of measures calculated from the data.

The first term of the energy function, $U_{\mathrm{likelihood}}(\mathrm{L}, \mathrm{D})$, is the data attachment term and it is calculated using the line detectors from the low level step. In our method the measure is calculated along the connected component taking into account the changes of directions (see Figure 6). As exemplified in Figure 6a, according to Tupin et al., the measure is calculated for each line segment separately. On the other hand, in the present method the measure is calculated once for each connected component considering the variation of directions along the structure. So, the observation d_i

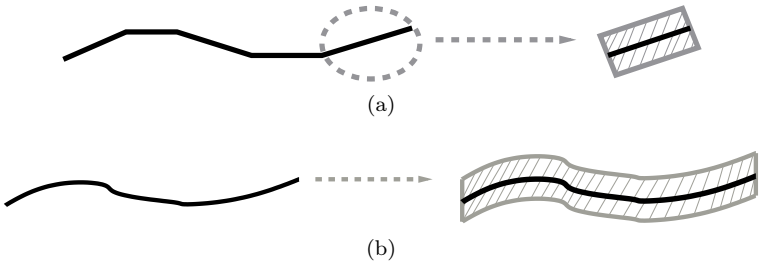


Figure 6.: Comparing the calculation of the measure for the likelihood term. (a) Using line segments as proposed by Tupin et al., the measure must be calculated for each line segment (b) Using connected components as proposed, the measure is calculated along the connected component considering the variation of directions along the structure.

related to each connected component is calculated as the mean of the line detector responses along it.

Higher responses mean that the connected component has more chance to be a road. In this work, the potentials $V(d_i|l_i)$ related to the observation d_i and label l_i used by Tupin, Houshmand, and Dactu are maintained, which were experimentally obtained after using manual segmentation of roads:

$$V(d_i, l_i = 0) = 0, \quad \text{if } d_i < t_1 \tag{10}$$

$$V(d_i, l_i = 0) = \frac{d_i - t_1}{t_2 - t_1}, \text{ if } t_1 < d_i < t_2 \tag{11}$$

$$V(d_i, l_i = 0) = 1, \quad \text{if } d_i > t_2 \tag{12}$$

$$V(d_i, l_i = 1) = 0, \quad \forall d_i, \tag{13}$$

where $t_1 \leq t_2 \in [0, 1]$. Consequently, if there is a good association between d_i and l_i, the value of the potential is low. In order to respect the normalisation constraint (Tupin et al. 1998), the constant $\ln Z$ is added to the potentials $V(d_i, l_i = 0)$, with $Z = t_1 + (1 - t_2)(1/\mathrm{e}) - (t_2 - t_1)((1/\mathrm{e}) - 1)$. Since $Z < 1$, we have $\ln Z < 0$. Besides, in order to take into account the length of the connected components, the potentials are multiplied by $\mathcal{L}_i$. Using the potentials above, the energy term for the

data attachment is defined as

$$U_{\text{likelihood}}(\mathrm{L},\mathrm{D}) = \sum_{i=1}^{N} V(d_i, l_i). \tag{14}$$

The second term, $U_{\text{prior}}(\mathrm{L})$, represents the structural information about roads. The local configurations, taking into account the cliques of the graph, should express the contextual knowledge (*a priori* information). In the context of the proposed framework, the user has the option to choose between the prior knowledge about roads considered in (Tupin et al. 1998) and (Tupin, Houshmand, and Dactu 2002), i.e., the user may choose between a configuration for non-urban or urban areas. This term can be written as the sum of the local clique potentials, $V_{\mathrm{C}}(l_i)$:

$$U_{\text{prior}}(\mathrm{L}) = \sum_{\mathrm{C}\in\mathrm{C_G}} V_{\mathrm{C}}(l_i, i \in \mathrm{C}) = \sum_{\mathrm{C}\in\mathrm{C_G}} V_{\mathrm{C}}(\mathrm{L}). \tag{15}$$

There are four parameters used in the definitions of the clique potentials: K_{e} penalises extremities, i.e., small and isolated roads; K_{l} ponders the length of the connected components; K_{c} considers the curvature between components and K_{i} penalises cliques with too many intersections.

Choosing $K_{\mathrm{e}} > 0$ and $K_{\mathrm{l}} > 0$ favours long structures. If $K_{\mathrm{c}} > 0$, structures with high curvature are penalised and if $K_{\mathrm{i}} > 0$, isolated structures are penalised. All potentials $V_{\mathrm{C}}(\mathrm{L})$ are null except for the cliques of higher order which correspond to the set of connected components that have a common extremity. For a clique C of this kind, the potentials are defined as in (Tupin et al. 1998; Tupin, Houshmand, and Dactu 2002):

$$\begin{aligned}
\forall i \in \mathrm{C} \mid l_i = 0 &\rightarrow V_{\mathrm{C}}(\mathrm{L}) = 0 \\
\exists! i \in \mathrm{C} \mid l_i = 1 &\rightarrow V_{\mathrm{C}}(\mathrm{L}) = K_{\mathrm{e}} - K_{\mathrm{l}}\mathcal{L}_i \\
\exists (i,j) \in \mathrm{C}^2 \mid l_i = l_j = 1, &\rightarrow V_{\mathrm{C}}(\mathrm{L}) = -K_{\mathrm{l}}(\mathcal{L}_i + \mathcal{L}_j) \\
\measuredangle ij > \frac{\pi}{2} \quad &\quad +K_{\mathrm{c}} \sin \measuredangle ij \\
\exists! (i,j,k) \in \mathrm{C}^3 \mid l_i = l_j = &\rightarrow V_{\mathrm{C}}(\mathrm{L}) = -K_{\mathrm{l}}(\mathcal{L}_i + \mathcal{L}_j + \\
= l_k = 1, i\mathrm{R}j, i\mathrm{P}k, j\mathrm{P}k \quad &\quad \mathcal{L}_k) + K_{\mathrm{c}}\,(\sin \measuredangle ij + \\
&\quad + {}^1\!/_2\,(\cos \mathcal{R}_{i,k} + \cos \mathcal{R}_{j,k})) \\
\exists! (i,j,k,l) \in \mathrm{C}^4 \mid l_i = \cdots = l_l = &\rightarrow V_{\mathrm{C}}(\mathrm{L}) = -K_{\mathrm{l}}(\mathcal{L}_i + \mathcal{L}_j + \\
= 1, i\mathrm{R}j, kRl, i\mathrm{P}k, j\mathrm{P}l \quad &\quad +\mathcal{L}_k + \mathcal{L}_l) + K_{\mathrm{c}}(\sin \measuredangle ij \\
&\quad + \sin \mathcal{R}_{k,l}) \\
\text{otherwise} &\rightarrow V_{\mathrm{C}}(\mathrm{L}) = K_{\mathrm{i}} \sum_{i \mid i \in \mathrm{C}} l_i
\end{aligned} \tag{16}$$

where $i\mathrm{R}j$ is the connection between two components, $i\mathrm{P}k$ denotes the parallelism between two components and C^m denotes an m-clique or a clique of size m. The four and five rules are optional, depending on the kind of structure present on the image. These rules are appropriate in the case of roads in urban areas for instance.

As said before, in order to find the road network, a minimisation process is carried out using Eq. (9). Simulated Annealing (Kirkpatrick, Gelatt, and Vecchi 1983) is

used along with the Gibbs sampler to execute this process.

2.4.2. Multi-scale approach

After the previous step, a road network is obtained. However, as the line detector used in the low-level only takes into account structures with width less or equal than 5 pixels, some roads may not be detected in this first step. In order to prevent that and to detect larger roads, a multi-scale approach is used. An image pyramid is created degrading the resolution by averaging the intensities of $n \times n$ pixel blocks; in this work $n = 2$ and $n = 4$ are used. The extraction process is carried out for each scale, and a single result is obtained making the union of the results of each scale.

A cleaning step is applied to the merged result in order to remove some overloaded data due to the detection of some roads in more than one scale.

2.4.3. Additional MRF step based on road sections

The final road network extracted after the execution of the previous steps has a more structured characteristic, because of the contextual information included in the MRF. However, this result still carries a reasonable amount of false detections and some parts of the roads are not detected. As an attempt to improve this result, an additional step is proposed in this work. The idea is to take the output of the previous step and pass it through a new high-level optimisation step as explained below (see Figure 7):

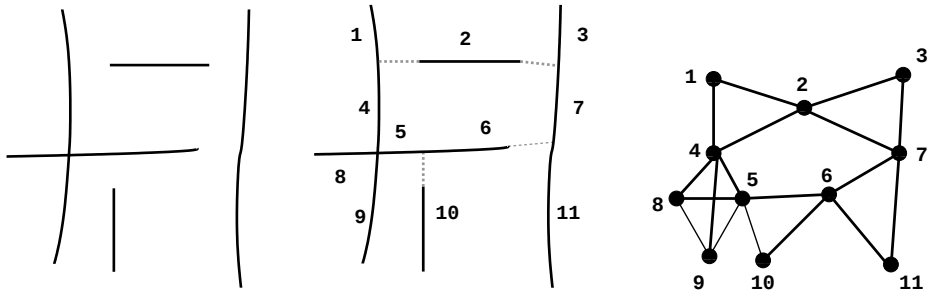
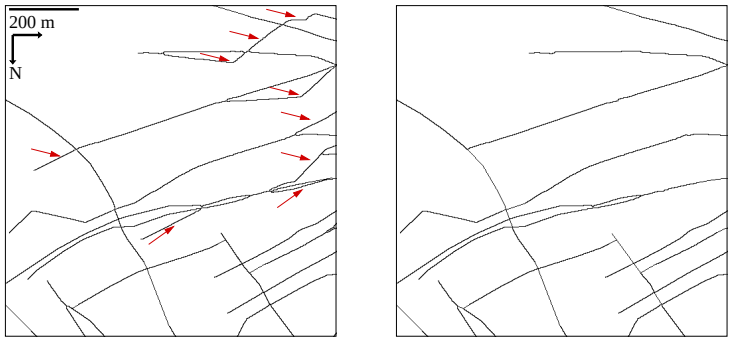


Figure 7.: Graph modelling for the additional high-level step (numbers represent identified connected components). (a) Example of output of the first high-level step. (b) Extension of the connected components. (c) Final graph representation.

- Process the image in order to obtain only road sections and crossroads: for each extremity of a connected component, extend the component to the closest connected component or to the limits of the image;
- Build a graph similarly as explained in Section 2.3 where each road section is a node of the graph and two nodes are connected if their corresponding sections share a crossroad;
- Proceed with the network reconstruction step as explained in Section 2.4. For the calculation of the likelihood term in Eq. (9), analysis is retained.

Figure 8 presents an example of zoomed results before and after the second high-level step where false alarms remaining from the previous step are removed, as indicated by the red arrows.



(a) (b)

Figure 8.: Comparison of results of a zoomed region of Toulouse area, centre at (43.5985º N, 1.3880º E), before (a) and after (b) the proposed second high-level step. The red arrows indicate the false alarms removed.

A consideration should be made about the first step explained above: one connected component is extended only if the distance of the extension is lower than a certain value and the likelihood measure of the extension is above a threshold (both values are parameters chosen empirically). The important point is that no connected component is isolated from the main road network. This additional step is much simpler and faster than the previous one, as the graph is much smaller. So, the extra computational time is not considerably increased and the improvements encountered by the results justify its use.

3. Results and discussion

This section presents the results obtained using the proposed framework. First, the results using optical and radar images separately are presented in Section 3.1. Section 3.2 presents the results using the proposed low-level fusion approach, explained in Section 2.2.2, considering optical and radar data, and also multitemporal data. The reference data used for all the experiments were determined manually.

3.1. *Images from a single source*

3.1.1. Optical images

For the experiments using optical data, panchromatic images with 0.7 m resolution obtained from the Quickbird sensor are used. The two regions used, corresponding to an area from Toulouse, France, are shown in Figures 9a(i) and 9b(i) along with their reference data in Figures 9a(ii) and 9b(ii). The reference data were obtained manually using the road atlas of each region as reference. The results of applying the proposed framework for road detection without fusion for each region are presented in Figures 9a(iii) and 9b(iii). Roads detected correctly are marked in green, while those detected incorrectly are marked in red and the missing roads are marked in black.

3.1.2. Radar images

For the experiments using radar data, images with 1 m resolution obtained from the TerraSAR-X sensor are used. The same regions chosen for the optical data are used here. They are shown in Figures 10a(i) and 10b(i) along with their reference data in Figures 10a(ii) and 10b(ii). The results of applying the proposed framework for road detection without fusion for each region are presented in Figures 10a(iii) and 10b(iii). As before, roads detected correctly are marked in green, those detected incorrectly are marked in red and the missing roads are marked in black.

We can notice the different aspects of the responses in each case relative to the characteristics of each image, optical or radar. The extraction process using the radar image seems to be efficient for detecting the main roads of the regions. However, some details (small or thin roads) are lost during the process. On the other hand, the extraction result using the optical image contains more false detections, although it is also capable to detect the main roads. This happens because we deal with a high-resolution optical image and the low-level detection is very sensitive to the details present in the image. Thus, more road candidates are detected during the low-level step. Besides that, one can notice that some other structures are mistaken with roads, such as rivers. Another interesting aspect is that depending on the region, it can be difficult to detect some roads in the optical image, because sometimes the contrast of the road region and its surrounding is too low. For instance, in the result presented for the first region, we have an entire road in the middle of the region that was not detected using the optical image. This can happen also on the other way around, where we have not detected roads in the radar image that can be detected in the optical one depending on the local contrasts.

In Table 1, some comparative results between the proposed method and the one in (Tupin et al. 1998) are presented considering the same regions analysed before. The values for the number of vertices using line segments (Tupin et al. 1998) and connected components (proposed) are compared. The decrease is about 30%. Besides, the Matthews Correlation Coefficient (Matthews 1975) value, given

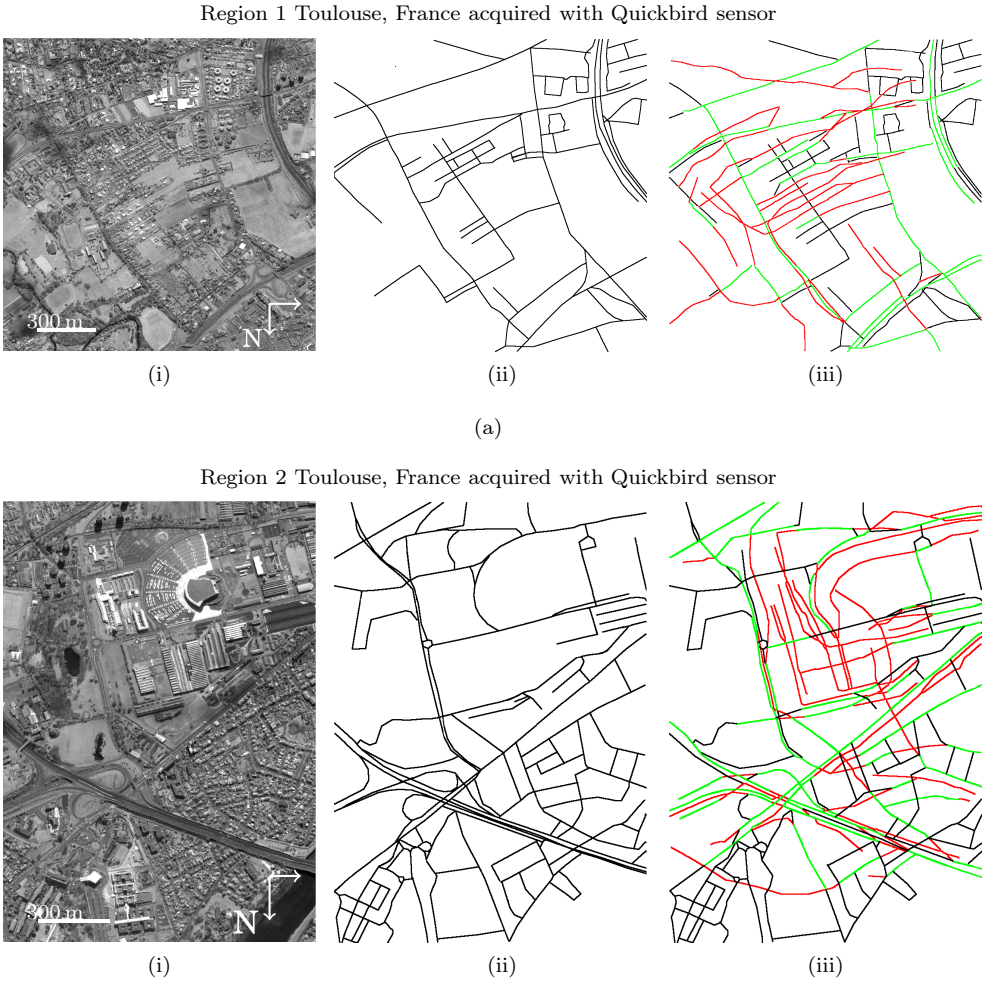


Figure 9.: Results for optical images: (i) Original images (ii) Reference data (iii) Detection results where the roads detected correctly are marked in green, those detected incorrectly are marked in red and the missing roads are marked in black. The center of the regions are at (43.6003° N, 1.3889° E) and (43.6025° N, 1.4059° E) respectively.

by

$$\mathrm{MCC} = \frac{(\mathrm{TP})(\mathrm{TN}) - (\mathrm{FP})(\mathrm{FN})}{\sqrt{[(\mathrm{TP}) + (\mathrm{FP})][(\mathrm{TP}) + (\mathrm{FN})][(\mathrm{TN}) + (\mathrm{FP})][(\mathrm{TN}) + (\mathrm{FN})]}}, \qquad (17)$$

is also compared for each region. TP, TN, FP and FN refer to the number of true positives, true negatives, false positives and false negatives, respectively. The MCC may be obtained directly from the confusion matrix of a binary classification, resulting in a value between -1 and 1, where 1 indicates perfect classification and -1 indicates total discordance between the obtained classification and the reference data. The proposed method outperforms the other one for all cases.

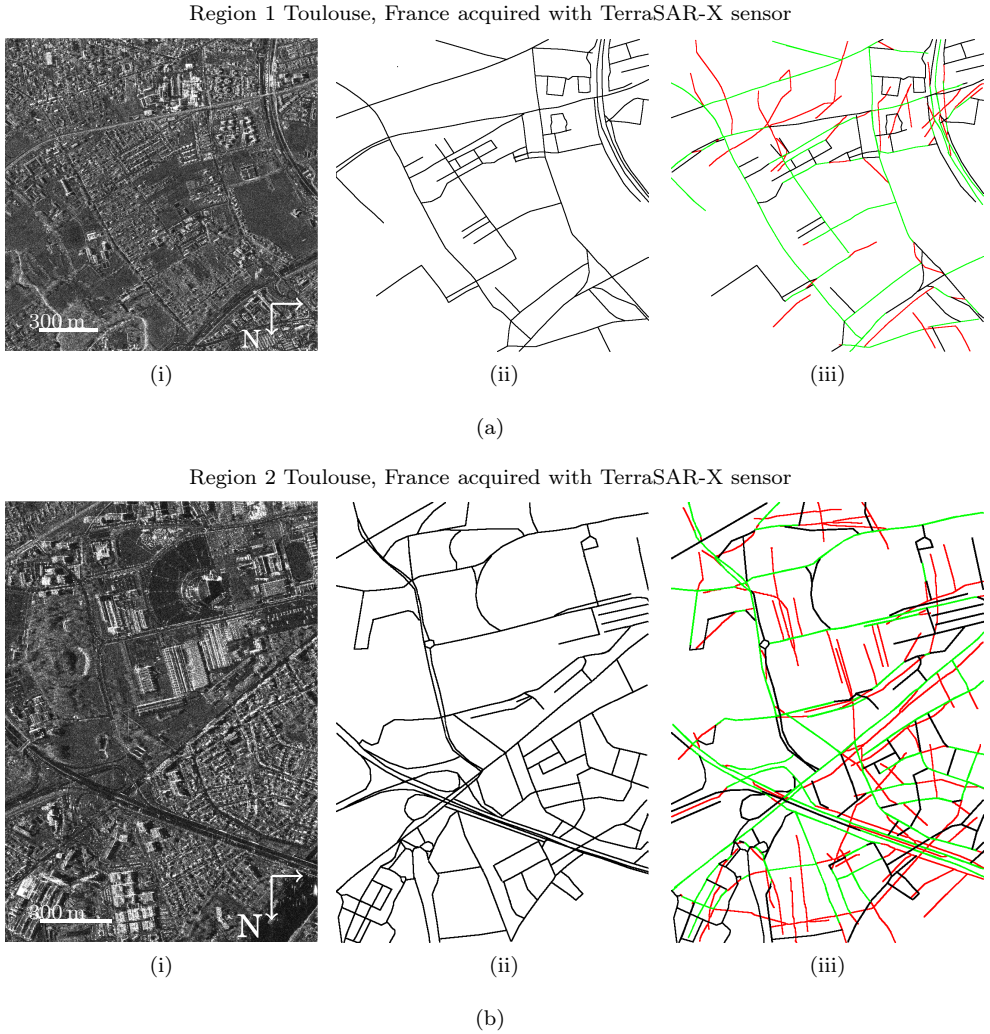


Figure 10.: Results for radar images: (i) Original images (ii) Reference data (iii) Detection results where the roads detected correctly are marked in green, those detected incorrectly are marked in red and the missing roads are marked in black. The center of the regions are at (43.6003° N, 1.3889° E) and (43.6025° N, 1.4059° E) respectively.

3.2. *Images from different sources*

3.2.1. SAR and optical fusion

In order to show the results of the proposed fusion approach, we use here the same regions shown before. These regions are previously co-registered using geometric and acquisition information. Figure 11 presents the results for the first and the second regions by applying the proposed fusion approach.

As expected, the result using the fusion of the two images exploits their complementary information, thus leading to an improved result.

Table 2 presents a quantitative analysis of the results of the two regions, named R1 and R2 in the first column, using the *correctness*, *completeness* and MCC indexes. The *correctness* measures the detection accuracy relative to the reference data. The *completeness* measures may be interpreted as the inverse of the failure error. As the actual road network was manually obtained, the values in Table 2 have more relative than absolute meaning. The most important thing is to notice the improvements

	Region 1			
	Optical		Radar	
	Our method	Tupin et al.	Our method	Tupin et al.
No. of vertices	**14754**	20205	**18265**	23991
MCC	**0.2872**	0.256	**0.378**	0.3274
	Region 2			
	Optical		Radar	
	Our method	Tupin et al.	Our method	Tupin et al.
No. of vertices	**9242**	14442	**14384**	20071
MCC	**0.2368**	0.2098	**0.2779**	0.2249

Table 1.: Comparison between the results of the proposed method and the method in (Tupin et al. 1998). Results of our method are emphasised in bold text.

shown by the results using the fusion.

Data		Completeness (%)	Correctness (%)	MCC
R1	Optical image	36.6	33.7	0.2872
	Radar image	48.6	39.4	0.3780
	Fused	**65.6**	**55.1**	**0.5601**
R2	Optical image	34.8	33.7	0.2368
	Radar image	41.5	35.6	0.2779
	Fused	**47.0**	**45.3**	**0.3744**

Table 2.: Quantitative evaluation of the results comparing the single source approach and the data fusion for regions 1 and 2 (R1 and R2). Best results are emphasised in bold text.

3.2.2. Multitemporal fusion

The last set of experiments is concerned with the fusion of multitemporal images. We use two multitemporal stacks of images from a region of Saint Gervais – France, and one stack of images from Chamonix – France, obtained in different days:

- 13 images obtained using the TerraSAR-X sensor of 1 m resolution
- 4 images obtained using the COSMO-SkyMed sensor of 3 m resolution
- 13 images obtained using the TerraSAR-X sensor of 1 m resolution, Chamonix region

The idea here is to compare the result of applying the low-level fusion approach using all the images of the stacks to the result applying the framework only to the image of the quadratic mean of all the images. The image of the mean is considered an ideal result, as it has a high quality, similar to a filtered one, because the roads are stable features and the incidence angles were similar.

Figure 12a presents one of the images of the TerraSAR-X stack, along with the quadratic mean image and its reference data, respectively. Figure 12b presents one of the images of the COSMO-SkyMed stack, and also the corresponding quadratic

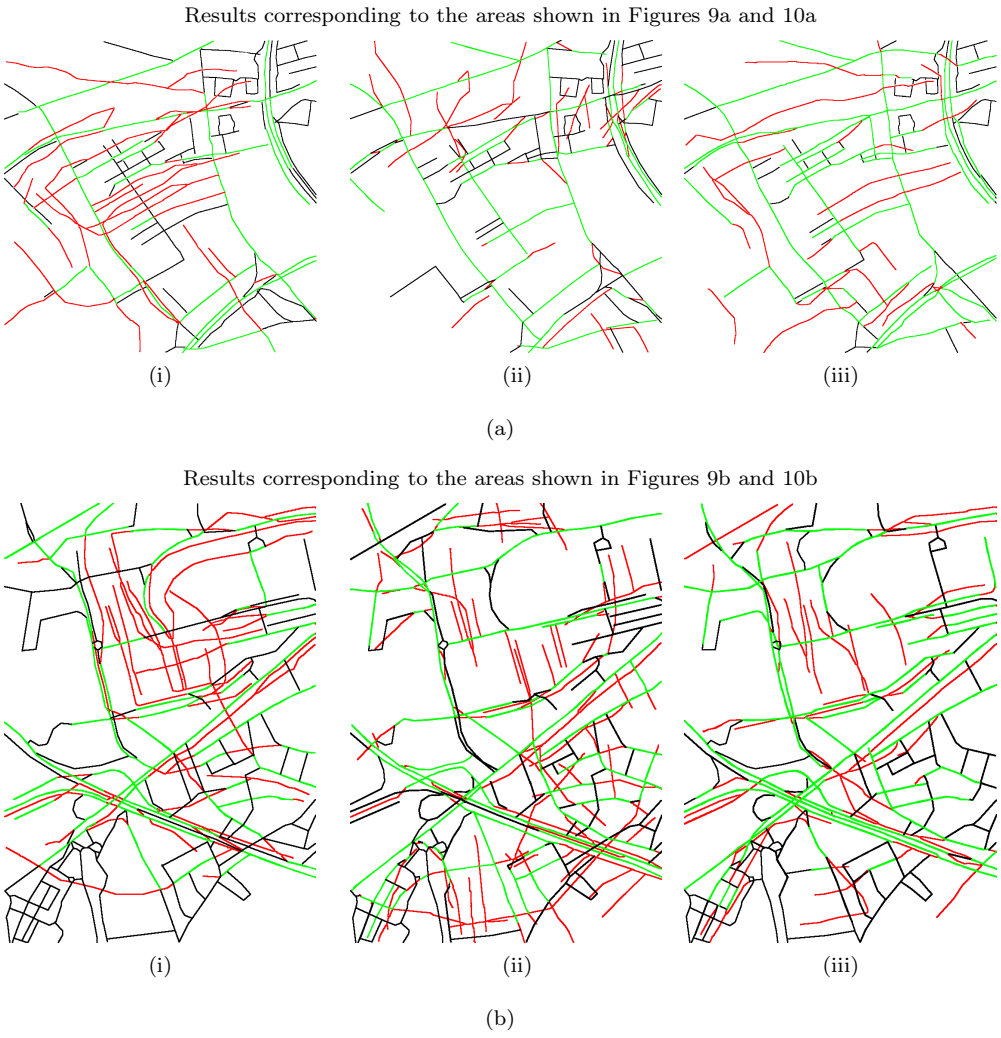


Figure 11.: Results fusing optical and radar data: (i) Results using only the optical images (ii) Results using only the radar images (iii) Results using the proposed fusion approach. Green: correct detection. Red: incorrect detection. Black: missing roads.

mean image and the reference data, respectively. Similarly, figure 12c presents one of the images of the TerraSAR-X stack (Chamonix region), and also the corresponding quadratic mean image and the reference data, respectively. The results obtained using those images are shown in Figure 13. For the three stacks, the road networks obtained using the quadratic mean image (Figures 13a(i), 13b(i) and 13c(i)) and the low-level fusion approach (Figures 13a(ii), 13b(ii) and 13c(ii)) lead to similar MCC indexes, as observed in Table 3. On the other hand, also for the three stacks, the best MCC indexes are achieved by the road networks obtained after merging the results from the quadratic mean and the low-level fusion (Figures 13a(iii), 13b(iii) and 13c(iii)). The merging is done by a simple OR combination. Even though more false detections are observed in those cases, the number of true detections is considerably increased.

From the quantitative results observed in Table 3, it is possible to note that the results for both stacks are very similar despite of the higher number of TerraSAR-X images and the better resolution of this stack. Similarly to other problems explored in remote sensing, images with better resolution do not always ensure a better

Images from TerraSAR-X multitemporal stack of a region in Saint Gervais – FR

(i) (ii) (iii)

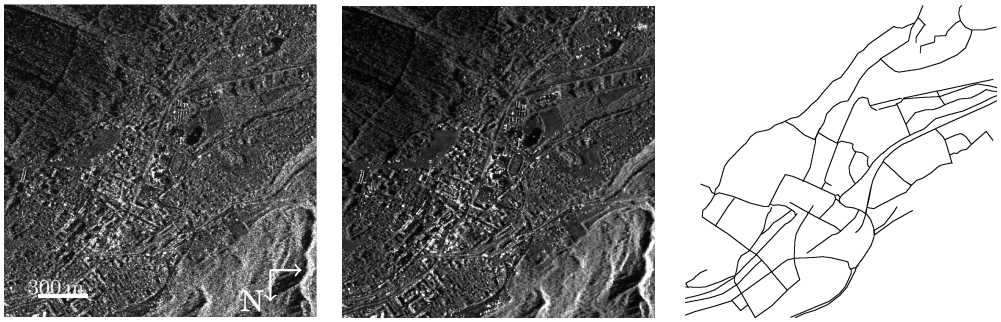
(a)

Images from COSMO-SkyMed multitemporal stack of a region in Saint Gervais – FR

(i) (ii) (iii)

(b)

Images from TerraSAR-X multitemporal stack of a region in Chamonix – FR



(i) (ii) (iii)

(c)

Figure 12.: Multitemporal image data. (i) One of the original images of each multitemporal stack (ii) Quadratic mean of all the original images from each multitemporal stack (iii) Reference data of the corresponding regions. The center of the regions are at (45.9130° N, 6.7068° E), (45.9137° N, 6.7067° E) and (45.9272° N, 6.8727° E) respectively.

result. In order to understand better these results, an analysis of the fusion of the 13 TerraSAR-X images was carried out. The low-level fusion was accomplished for a varying number of images, from 1 to 13, in chronological order of acquisition. For each experiment the MCC was calculated, as presented in Figure 14. It is noticed from the analysis that the result using 6 images gives a quality measure very close to

Results corresponding to the stack shown in Figure 12a

(i) (ii) (iii)

(a)

Results corresponding to the stack shown in Figure 12b

(i) (ii) (iii)

(b)

Results corresponding to the stack shown in Figure 12c

(i) (ii) (iii)

(c)

Figure 13.: Results for multitemporal image data. (i) Obtained results using only the quadratic mean image of each stack (ii) Obtained results using the proposed low-level fusion for each stack (iii) Results obtained merging the quadratic mean and the low-level fusion for each stack. Green: correct detection. Red: incorrect detection. Black: missing roads.

the best result. This can be a clue that depending on the combination of the images chosen from the stack, a good result can be obtained with a decreased number of images. Besides, the CSK and TerraSAR-X images, though very similar, are not exactly the same. Additional analysis of multitemporal stacks should be carried out

Data		Completeness (%)	Correctness (%)	MCC
St-Gervais (TSX)	Mean	42.0	57.6	0.4437
	Fusion	41.0	59.4	0.4473
	Fusion+Mean	**50.8**	**51.7**	**0.4607**
St-Gervais (CSK)	Mean	40.6	61.5	0.462
	Fusion	44.0	58.7	0.4684
	Fusion+Mean	**53.1**	**52.1**	**0.4795**
Chamonix TSX)	Mean	23.0	42.8	0.2735
	Fusion	20.6	45.8	0.2700
	Fusion+Mean	**28.1**	**40.1**	**0.2911**

Table 3.: Quantitative evaluation of the results applying the proposed framework to the three different multitemporal stacks. Values are presented for three different fusion approaches: (Mean) quadratic mean of all the images, (Fusion) low-level fusion, and (Fusion+Mean) merging mean and low-level fusion. Best results are emphasised in bold text.

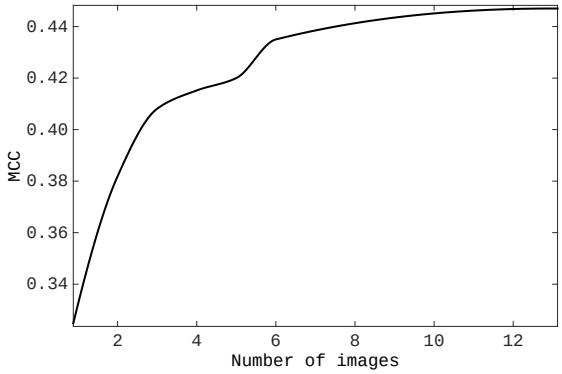


Figure 14.: Analysis of the influence of the number of images used for the fusion of the TerraSAR-X images. It can be observed that fusing the first 6 images of the stack a measure close to the best is obtained.

in future works, and dealing with the fusion of other satellite sensors.

Tables 4 (optical, radar and optical/radar) and 5 (multitemporal fusion) present the values of the parameters used in all experiments shown before. There are few variations for the parameters in the case of optical/radar fusion, mainly for the high-level step, which have more variations in the case of multitemporal fusion. This fact emphasises the stability and flexibility of the proposed method, which turns to be not too sensible to the parameters setting. Indeed, the parameters are set empirically, but the same set can be used for the same kind of landscape, giving similar detection performance. In the case of the low-level step when fusing images, the values of ρ and r are usually higher in order to control the final result of the low-level fusion.

Initial results of the framework described here applied to the fusion of TerraSAR-X and Cosmo-SkyMed high-resolution images for road extraction may be found in (Sportouche et al. 2012).

Scale	Low level			Graph		High level			
	ρ	r	Dimensions (height,width) (pixels)	$\mathcal{L}_{\max}$	$\mathcal{D}_{\max}$	K_{e}	K_{l}	K_{c}	K_{i}
Region 1									
3	$[0.35, 0.5]$	$[0.35, 0.5]$	$(3, 3)$	$[30, 40]$	$[30, 40]$	$[0.3, 0.4]$	-0.12	0.3	0.3
2	$[0.3, 0.45]$	$[0, 3, 0.5]$	$(5, 4)$	$[60, 70]$	$[60, 70]$	$[0.3, 0.4]$	-0.12	0.3	0.3
1	0.3	$[0.3, 0.5]$	$(5, 5)$	$[100, 110]$	$[100, 110]$	$[0.3, 0.4]$	-0.12	0.3	0.3
Region 2									
3	$[0.38, 0.45]$	$[0.35, 0.5]$	$(4, 3)$	$[30, 40]$	$[30, 40]$	$[0.21, 0.3]$	$-0, 12$	0.3	0.3
2	0.35	$[0.35, 0.5]$	$(5, 4)$	70	70	0.3	-0.12	0.3	0.3
1	$[0.3, 0.35]$	$[0.3, 0.5]$	$(5, 5)$	100	100	0.3	-0.12	0.3	0.3

Table 4.: Intervals of parameters' values used for the experiments with optical, radar and optical/radar images. Parameters values are set empirically.

Image	Low level			Graph		High level			
	ρ	r	Dimensions (height, width) (pixels)	$\mathcal{L}_{\max}$	$\mathcal{D}_{\max}$	K_{e}	K_{l}	K_{c}	K_{i}
St-Gervais (TSX)									
Mean	0.45	0.45	$(4, 4)$	120	30	0.2	-0.22	0.35	0.08
Fusion	0.44	0.44	$(5, 4)$	120	30	0.3	-0.4	0.35	0.1
St-Gervais (CSK)									
Mean	0.35	0.35	$(3, 3)$	120	30	0.21	-0.2	0.3	0.1
Fusion	0.45	0.45	$(3, 3)$	120	30	0.25	-0.3	0.3	0.15
Chamonix (TSX)									
Mean	0.3	0.35	$(3, 3)$	110	30	0.21	-0.21	0.3	0.1
Fusion	0.45	0.5	$(3, 3)$	110	30	0.3	-0.35	0.3	0.1

Table 5.: Values of the parameters used for the multitemporal fusion. Parameters values are set empirically.

4. Conclusions

The idea of fusing information for object detection, image classification and, consequently, for SAR and optical image interpretation has been increasingly studied nowadays. The framework proposed here is flexible, and each module can be eventually changed depending on the application. Besides, different feature detectors can be used for the low-level step and the energies of the MRFs can be also improved or changed. In fact, we intend to test other feature detectors for the optical image in order to improve the final result after fusion as, for instance, edge and visual alignments detectors.

Further studies will be carried on multi-sensor and multi-model fusion using this framework, such as using SAR images taken from different angles and from different sensors (as the example of fusing TerraSAR-X and Cosmo-SkyMed images). Another interesting future path is to extend this framework for polarimetric SAR images, for instance relying on recent low level detection methods for this kind of images (Zhou et al. 2011).

Acknowledgement

The authors would like to thank FAPESP grant #2011/50761-2, CNPq, CAPES PDEE grant #0310-10-7, NAP eScience - PRP - USP (Brazilian agencies) for funding. We thank also the project EFIDIR (ANR-2007-MCDC0-04, `http://www.efidir.fr`) to provide the images used. Original images, results and ground-truths are available at `http://perso.telecom-paristech.fr/~tupin/radarteam/DemoPerciano/`.

References

Barzohar, Meir, and David B. Cooper. 1996. "Automatic Finding of Main Roads in Aerial Images by Using Geometric-Stochastic Models and Estimation." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18: 707–721.

Bloch, I. 1996. "Information combination operators for data fusion: a comparative review with classification." *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans* 26 (1): 52–67.

Brunner, D., G. Lemoine, and L. Bruzzone. 2008. "Extraction of Building Heights from VHR SAR Imagery using an Iterative Simulation and Match Procedure." In *Proc. IEEE International Geoscience and Remote Sensing Symposium, 2008.*, Vol. 4IV – 141–IV – 144.

Canny, J. 1986. "A Computational Approach to Edge Detection." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8: 679–698.

Caves, Ronald G., and Shaun Quegan. 1994. "Matching segmentation algorithms to ERS-1 SAR applications." In *Proc. of SPIE,* Vol. 2316148–158.

Chanussot, J., G. Mauris, and P. Lambert. 1999. "Fuzzy Fusion Techniques for Linear Features Detection in Multitemporal SAR Images." *IEEE Transactions on Geoscience and Remote Sensing* 37 (3): 1292–1305.

Cook, Rod, Ian McConnell, Christopher J. Oliver, and Edward Welbourne. 1994. "MUM (Merge Using Moments) segmentation for SAR images." In *Proc. of SPIE,* 92–103.

D'Elia, C., S. Ruscino, M. Abbate, B. Aiazzi, S. Baronti, and L. Alparone. 2014. "SAR Image Classification Through Information-Theoretic Textural Features, MRF Segmentation, and Object-Oriented Learning Vector Quantization." *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of* 7 (4): 1116–1126.

Dell'Acqua, F., P. Gamba, and G. Lisini. 2005. "Road extraction aided by adaptive directional filtering and template matching." In *Proc. of URBAN,* Vol. 34(on CD–ROM). Tempe, AZ. Mar..

Desolneux, Agnès, Lionel Moisan, and Jean-Michel Morel. 2000. "Meaningful Alignments." *International Journal of Computer Vision* 40: 7–23. 10.1023/A:1026593302236.

Gamba, P., F. Dell'Acqua, and G. Lisini. 2006. "Improving urban road extraction in high-resolution images exploiting directional filtering, perceptual grouping, and simple topological concepts." *IEEE Geoscience and Remote Sensing Letters* 3 (3): 387 – 391.

Hall, D. L., and J. Llinas. 1997. "An introduction to multisensor data fusion." In *Proc. of the IEEE,* Vol. 856–23. Jan..

He, Chu, Bo Shi, Yu Zhang, Xin Xu, and Mingsheng Liao. 2014. "Road extraction for SAR imagery based on the combination of beamlet and a selected kernel." In *Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International,* 2257–2260. July.

He, Chu, Fang Yang, Sha Yin, Xinping Deng, and MingSheng Liao. 2013. "Stereoscopic Road Network Extraction by Decision-Level Fusion of Optical and SAR Imagery." *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of* 6 (5): 2221–2228.

Hedman, K., U. Stilla, G. Lisini, and P. Gamba. 2010. "Road Network Extraction in VHR SAR Images of Urban and Suburban Areas by Means of Class-Aided Feature-Level Fusion." *IEEE Geoscience and Remote Sensing Letters* 48 (3): 1294 – 1296.

Hinz, Stefan, and Albert Baumgartner. 2003. "Automatic extraction of urban road networks from multi-view aerial imagery." *ISPRS Journal of Photogrammetry and Remote Sensing* 58 (1–2): 83 – 98. Algorithms and Techniques for Multi-Source Data Fusion in Urban Areas.

Hoberg, T., F. Rottensteiner, R. Queiroz Feitosa, and C. Heipke. 2015. "Conditional Random Fields for Multitemporal and Multiscale Classification of Optical Satellite Imagery." *Geoscience and Remote Sensing, IEEE Transactions on* 53 (2): 659–673.

Huang, Zhijian, Jinfang Zhang, Luxiao Wang, and Fanjiang Xu. 2012. "A feature fusion method for road line extraction from remote sensing image." In *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International,* 52–55. July.

Jeon, Byoung-Ki, Jeong-Hun Jang, and Ki-Sang Hong. 2002. "Road detection in spaceborne SAR images using a genetic algorithm." *IEEE Transactions on Geoscience and Remote Sensing* 40 (1): 22–29.

Katartzis, A., H. Sahli, V. Pizurica, and J. Cornelis. 2001. "A model-based approach to the automatic extraction of linear features from airborne images." *IEEE Transactions on Geoscience and Remote Sensing* 39 (9): 2073–2079.

Kirkpatrick, S., C. D. Gelatt, and M. P. Vecchi. 1983. "Optimization by simulated annealing." *SCIENCE* 220 (4598): 671–680.

Lacoste, Caroline, Xavier Descombes, and Josiane Zerubia. 2010. "Unsupervised line network extraction in remote sensing using a polyline process." *Pattern Recognition* 43 (4): 1631 – 1641.

Lezoray, Olivier, and Leo Grady. 2012. *Image Processing and Analysis with Graphs: Theory and Practice*. 1st ed. CRC Press.

Lisini, Gianni, Paolo Gamba, Fabio Dell'Acqua, and Francesco Holecz. 2011. "First results on road network extraction and fusion on optical and SAR images using a multi-scale adaptive approach." *International Journal of Image and Data Fusion* 2 (4): 363–375. `http://www.tandfonline.com/doi/abs/10.1080/19479832.2011.613412`.

Lisini, G., P. Gamba, and D. Luebeck. 2011. "Road extraction in urban and rural environments exploiting a dual-band SAR system." In *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International,* 3610–3613. July.

Lisini, G., C. Tison, F. Tupin, and P. Gamba. 2006. "Feature fusion to improve road network extraction in high-resolution SAR images." *IEEE Geoscience and Remote Sensing Letters* 3 (2): 217 – 221.

Lu, P., K. Du, W. Yu, R. Wang, Y. DENG, and T. Balz. 2014. "A New Region Growing-Based Method for Road Network Extraction and Its Application on Different Resolution SAR Images." *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of* PP (99): 1–12.

Matthews, B. W. 1975. "Comparison of the predicted and observed secondary structure of T4 phage lysozyme." *Biochim. Biophys. Acta* 405: 442–451.

Miao, Z., W. Shi, P. Gamba, and Z. Li. 2015a. "An Object-Based Method for Road Network Extraction in VHR Satellite Images." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 8 (10): 4853–4862.

Miao, Z., W. Shi, A. Samat, G. Lisini, and P. Gamba. 2015b. "Information Fusion for Urban Road Extraction From VHR Optical Satellite Images." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* PP (99): 1–14.

Moser, G., and S.B. Serpico. 2012. "Unsupervised change detection with high-resolution SAR images by edge-preserving Markov random fields and graph-cuts." In *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International,* 1984–1987. July.

Negri, M., P. Gamba, G. Lisini, and F. Tupin. 2006. "Junction-aware extraction and regularization of urban road networks in high-resolution SAR images." *IEEE Geoscience and Remote Sensing Letters* 44 (10): 2962 – 2971.

Papila, Ibrahim, Sedef Kent, and Mesut Kartal. 2014. "An Automated SAR Image Registration Approach Using Hidden Markov Scale Invariant Feature Transform Algorithm." In *EUSAR 2014; 10th European Conference on Synthetic Aperture Radar; Proceedings of,* 1–4. June.

Perciano, Talita, Roberto Hirata, and Lúcio de Castro Jorge. 2010. "Ridge Linking Using an Adaptive Oriented Mask Applied to Plant Root Images with Thin Structures." In *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications,* Vol. 6419 of *Lecture Notes in Computer Science* edited by Isabelle Bloch and Roberto Cesar. 378–385. Springer Berlin / Heidelberg.

Perciano, Talita, Florence Tupin, Roberto Hirata Jr., and Roberto M. Cesar. 2011. "A hierarchical Markov random field for road network extraction and its application with optical and SAR data." In *Proc. IEEE International Geoscience and Remote Sensing Symposium, 2011,* 1159–1162.

Poulain, V., J. Inglada, and M. Spigai. 2008. "High Resolution Remote Sensing Image Analysis with Exogenous Data: A Generic Framework." In *Proc. IEEE International*

Geoscience and Remote Sensing Symposium, 2008, Vol. 2II–1025 –II–1028. july.

Poulain, V., J. Inglada, M. Spigai, J.-Y. Tourneret, and P. Marthon. 2009. "Fusion of high resolution optical and SAR images with vector data bases for change detection." In *Proc. IEEE International Geoscience and Remote Sensing Symposium, 2009,* Vol. 4IV–956 – IV–959. july.

Salberg, A., and O.D. Trier. 2012. "Temporal analysis of multisensor data for forest change detection using hidden Markov models." In *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International,* 6749–6752. July.

Skriver, H., J. Schou, A. A. Nielsen, and K. Conradsen. 2005. "Polarimetric edge detector based on the complex Wishart distribution." In *Proc. IEEE International Geoscience and Remote Sensing Symposium, 2005,* Vol. 73149–3151. Sydney, Australia. Jul..

Song, Huihui, Bo Huang, and Kaihua Zhang. 2014. "Shadow Detection and Reconstruction in High-Resolution Satellite Images via Morphological Filtering and Example-Based Learning." *Geoscience and Remote Sensing, IEEE Transactions on* 52 (5): 2545–2554.

Sportouche, H., F. Tupin, and L. Denise. 2009. "Building detection by fusion of optical and SAR features in metric resolution data." In *Proc. IEEE International Geoscience and Remote Sensing Symposium, 2009,* Vol. 4IV–769 –IV–772. july.

Sportouche, H., F. Tupin, J.-M. Nicolas, T. Perciano, and C.-A. Deledalle. 2012. "How to combine TerraSAR-X and Cosmo-SkyMed high-resolution images for a better scene understanding?." In *Proc. IEEE International Geoscience and Remote Sensing Symposium, 2012,* 178 –181. july.

Steger, Carsten. 1998. "An Unbiased Detector of Curvilinear Structures." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20 (2): 113–125.

Sun, N, J X Zhang, G M Huang, Z Zhao, and L J Lu. 2014. "Review of Road Extraction Methods from SAR Image." In *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International,* 1–5.

Tison, C., J.-M. Nicolas, F. Tupin, and H. Maitre. 2004. "A new statistical model for Markovian classification of urban areas in high-resolution SAR images." *IEEE Geoscience and Remote Sensing Letters* 42 (10): 2046 – 2057.

Touzi, R., A. Lopes, and P. Bousquet. 1988. "A statistical and geometrical edge detector for SAR images." *IEEE Transactions on Geoscience and Remote Sensing* 26 (6): 764 –773.

Tupin, F., Bijan Houshmand, and Mihai Dactu. 2002. "Road Detection in Dense Urban Areas Using SAR Imagery anf the Usefulness of Multiple Views." *IEEE Transactions on Geoscience and Remote Sensing* 40 (11): 2405–2414.

Tupin, F., H. Maître, J.-M. Nicolas, and E. Pechersky. 1998. "Detection of linear features in SAR images: Application to road network extraction." *IEEE Transactions on Geoscience and Remote Sensing* 36: 434–453.

Tupin, F, and M Roux. 2003. "Detection of building outlines based on the fusion of SAR and optical features." *ISPRS Journal of Photogrammetry and Remote Sensing* 58 (1-2): 71 – 82. Algorithms and Techniques for Multi-Source Data Fusion in Urban Areas.

Türetken, E., F. Benmansour, B. Andres, H. Pfister, and P. Fua. 2013. "Reconstructing Loopy Curvilinear Structures Using Integer Programming." In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on,* 1822–1829. June.

Vosselman, George, and Jurrien De Knecht. 1995. "Road tracing by profile matching and Kalman filtering." In *Proc. Workshop on Automatic Extraction of ManMade Objects from Aerial and Space Images,* Vol. 2265–275. Birkhuser Verlag.

Xu, K., W. Yang, G. Liu, and H. Sun. 2013. "Unsupervised Satellite Image Classification Using Markov Field Topic Model." *Geoscience and Remote Sensing Letters, IEEE* 10 (1): 130–134.

Yang, Wen, Hui Song, Xiaojing Huang, Xin Xu, and Mingsheng Liao. 2014. "Change Detection in High-Resolution SAR Images Based on Jensen-Shannon Divergence and Hierarchical Markov Model." *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of* 7 (8): 3318–3327.

Zhou, Guangyi, Yi Cui, Yilun Chen, Jian Yang, H. Rashvand, and Y. Yamaguchi. 2011. "Linear Feature Detection in Polarimetric SAR Images." *Geoscience and Remote Sens-*

ing, IEEE Transactions on 49 (4): 1453–1463.

Zhou, J., W. F. Bischof, and T. Caelli. 2005. "Robust and efficient road tracking in aerial images." In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences (CMRT05),* Vol. XXXVI35–40.

Zhou, Y.T., V. Venkateswar, and R. Chellappa. 1989. "Edge Detection and Linear Feature Extraction Using a 2-D Random Field Model." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11: 84–95.