

Feature Fusion to Improve Road Network Extraction in High-Resolution SAR Images

Gianni Lisini, *Member, IEEE*, Céline Tison, Florence Tupin, and Paolo Gamba, *Senior Member, IEEE*

Abstract—This letter aims at the extraction of roads and road networks from high-resolution synthetic aperture radar data. Classical methods based on line detection do not use all the information available; indeed, in high-resolution data, roads are large enough to be considered as regions and can be characterized also by their statistics. This property can be used in a classification scheme. Therefore, this letter presents a road extraction method which is based on the fusion of classification (statistical information) and line detection (structural information). This fusion is done at the feature level, which helps to improve both the level of likelihood and the number of the extracted roads. The proposed approach is tested with two classification methods and one line extractor. Results on two different datasets are discussed.

Index Terms—Data fusion, road network extraction, synthetic aperture radar (SAR) image interpretation, urban remote sensing.

I. INTRODUCTION

SATELLITE-BASED high-resolution synthetic aperture radar (SAR) sensors are about to be launched and may provide data at a spatial resolution near 1 m and at reasonable costs, useful for a variety of land and sea applications. This letter is thus aimed at providing a possible solution for one of the most relevant applications of these data: road network detection in dense urban areas. In the past 20 years, many approaches have been developed to deal with road (or, generalizing, linear feature) detection in radar images. Due to the usually coarse resolution of SAR data, most of them exploit a local criterion evaluating the radiometry on some small neighborhood surrounding a target pixel to discriminate lines from background [1]. These segments are eventually connected into a network by introducing some large-scale knowledge about the structures to be detected [2]. The local criterion is related to the need to extract edges between roads and the surrounding environment. Generalizing this idea, Chanussot *et al.* [3] extracted roads by a combination of multiple edge detectors in a fuzzy framework.

Things are different when considering high-resolution SAR data. Contextual relationships between pixels become very important as objects are not restricted to few pixels anymore. Road

detectors change from *line* (edge) to *object* detectors, where the characteristics of road objects may be exploited. For instance, Huber and Lang [4] propose a road extraction operator jointly considering the presence of road lateral edges and the road center continuity. By this way, two different geometrical properties of the road object are simultaneously tested. The road context may take into account also the layover due to buildings, vehicles moving or stopping, bridges, and traffic signs. This allows refining the road model for different environments, from rural to built-up areas, and change adaptively the extraction process [5]. The knowledge of the context may also help in discriminating linear features belonging to other classes of cartographic elements, like power lines and railroads. A knowledge-based system designed to this aim is proposed in [6].

It is interesting to note that these approaches refer to the geometrical/structural context of a road, neglecting or undervaluing its radiometric properties as a region. This point is instead considered in [7] and [8], where clustering of pixels assigned to the “road” class by a classifier is proposed. There the authors try and discriminate the roads by grouping pixels classified as “roads” into linear or curvilinear segments using modified Hough transforms or dynamic programming. The dual approach is proposed in [9], where segmentation is used to discard uniform areas and allow the extraction of edges where statistical homogeneity is lost.

In summary, it is clear that road extraction in high-resolution SAR data by linear or curvilinear line detection is an approach only partially using the available information. Roads may be instead modeled as image segments with a distinct statistical behavior, which can be exploited through classification. Moreover, exploiting their context improves the detection performances. Therefore, in the present work, we aim at using all the information by fusing the output of a line detector and the map obtained by applying a classifier to the same data. The method is therefore based on the joint analysis of the line elements detected by means of SAR image filtering and the classification map obtained by SAR data clustering.

The conceptual workflow of the proposed procedure is described in Fig. 1. The basic elements of the procedure are a classification and a line detection tool in a multiresolution framework to improve the recognition of roads with different widths. Thus the complete road network extraction and reconstruction is made by three steps: first comes the detection of segment candidates using the fusion of a local line detector and a classifier. Then, the reconstruction of the network using Markov random fields (MRFs), with the graph built using the skeleton of the previously selected segments and the likelihood term of the MRF given again by the fusion of a line detector and a classifier. Eventually, the fusion of the road networks reconstructed at different resolutions.

Manuscript received January 14, 2005; revised April 20, 2005. This work is an enlarged version of the paper “Improving Road Network Extraction in High Resolution SAR Images by Data Fusion,” presented at CEOS SAR Workshop 2004. The research was supported in part by the Italian Space Agency (ASI) under Contract I/R/177/02.

G. Lisini and P. Gamba are with the Dipartimento di Elettronica, Università di Pavia, 1-27100 Pavia, Italy (e-mail: gianni.lisini@unipv.it).

C. Tison is with the Centre National d’Etudes Spatiales, 31000 Toulouse, France.

F. Tupin is with Telecom Paris, Department TSI, Centre National de la Recherche Scientifique UMR 5141, 75013 Paris, France.

Digital Object Identifier 10.1109/LGRS.2005.862526

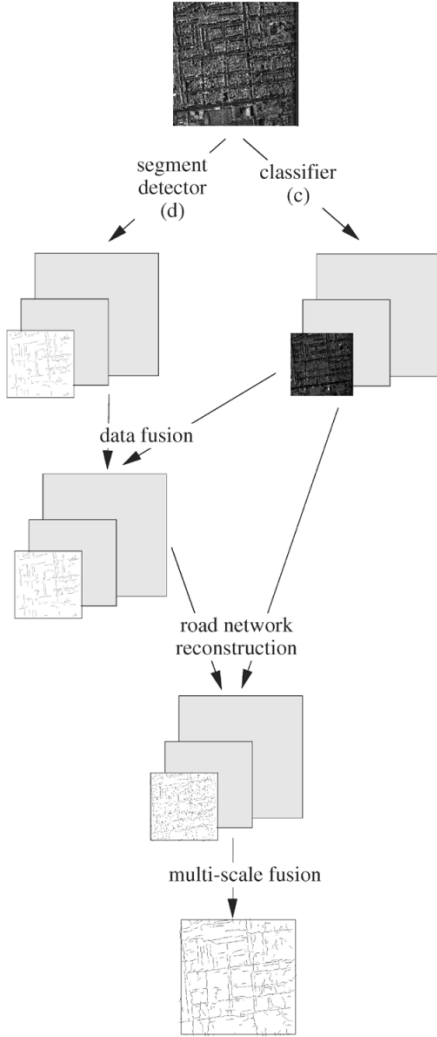


Fig. 1. Conceptual workflow of the proposed procedure.

II. FUSION PROCEDURE

Generally speaking, roads and streets may be extracted from SAR images taking into account two characteristics. First of all, they are laterally bounded by one or more edges, depending on the ground spatial resolution of the data, and therefore they can be individuated by edge extractors. Moreover, roads form a rather homogenous area with a backscatter which is different, in a statistical sense, from the other land cover classes. This is difficult to be accomplished in coarse-resolution SAR images, such as those by satellite sensors, but it is clearly visible in high-resolution SAR data [12].

Finally streets can be individuated by line detection algorithms, but also by suitable classification procedures. In this letter, some possible algorithms for classification and line detection are proposed. They should only be considered as examples and can easily be replaced by other methods, which is why they are just outlined here. In Section III, it appears that results improve without any respect to the classifier used.

1) *Line Detector*: The ratio and correlation detector was introduced in [16], and is based on the statistical properties of Gamma-distributed amplitude image (assumption of fully developed speckle). Indeed, the fully developed speckle assumption is not valid for high-resolution data, either because there are too few scatterers in the resolution cell, or more usually because

some very strong scatterer is present in the resolution cell [10]. However, the presence of these strong scatterers increases the contrast between the “road part” and the surrounding, making easier the detection and giving acceptable results in practice. Anyway, this detector results from the fusion of ratio-based detector $d_{(w,\theta)}^{(1)}(x,y)$ and a correlation-based detector $d_{(w,\theta)}^{(2)}(x,y)$ (see [11] for more details). For each position (x,y) the two detectors are computed as a function of the direction θ and the width w of the edge. For each pixel, (x,y) only the maximum value computed within all possible couple of values (w,θ) is stored. The outputs of the two detectors are merged to improve the reliability of the extraction by using an associative symmetrical sum [13]

$$d_{(w,\theta)}(x,y) = \frac{d_{(w,\theta)}^{(1)}(x,y)d_{(w,\theta)}^{(2)}(x,y)}{1 - d_{(w,\theta)}^{(1)}(x,y) - d_{(w,\theta)}^{(2)}(x,y) + 2d_{(w,\theta)}^{(1)}(x,y)d_{(w,\theta)}^{(2)}(x,y)} \quad (1)$$

2) Classifiers:

1) *Markovian Classifier*. A very precise classification approach recently developed in [10] is based on a Markovian segmentation. The class distributions are modeled by Fisher distributions, and the learning is supervised. This result is improved, when interferometric data are available, by merging it with coherence and interferogram. The classification map gathers features with similar backscattering behavior in statistical sense and with similar architectural meaning.

2) *Fuzzy ARTMAP Classifier*. In urban areas and especially for multiband data a fuzzy ARTMAP classifier [18] has shown to provide excellent results. This neurofuzzy classifier requires a training step, during which it collects spectral and spatial patterns for the pixels in training areas. During the classification step, it compares via a fuzzy AND logical operator the stored patterns (called *memories* with the input pattern, assigning one of the output class to the corresponding pixel).

Let us assume then that both line detection and classification map can be considered, providing for the generic position (x,y) in the original SAR data matrix two values $d_{(w,\theta)}(x,y)$ and $c(x,y)$. They provide the “likelihood” (however defined) of the pixel being a road pixel because it belongs to a detected edge or to the “road” class, respectively. We may merge the two values using again the associative symmetrical sum $s(d,c)$, where

$$s(d,c) = \frac{dc}{1 - d - c + 2dc} \quad (2)$$

and use this new value to decide if the pixel is actually a road pixel or not. The results will be thresholded by an arbitrary coefficient included between 0 and 1. In our test it was placed to 0.5. The choice of this kind of formula is suggested in [11].

However, this choice looks at each pixel separately. We want to make a further step forward based on the assumptions that line (or edge) detectors are usually more reliable than classifiers for SAR images and that the classification is independent from the linear shape. As a result, it may be more efficient to perform the above mentioned associative symmetrical sum at

the feature (segment) level, instead than at the pixel level. Formula (1) is computed joining likelihood levels d and c referring to a linear element and the road segment associated to it in the classification map. Therefore, for each possible linear element this segment is individuated checking for many different areas around it, with increased widths and different orientations. So, for the i th linear element rs_i , $d_{w,\theta}(rs_i)$ is the mean value of the detector output on its pixels. Instead, $c_{(w,\theta)}(rs_i)$ is computed as the percentage of “road” pixels in a region $2w$ pixels wide ($w = 1, \dots, 5$) around this element, with the possibility to check also for slightly different orientations θ . So, the overall result $s(rs_i)$ is dependent on two parameters, w and θ , and a search for the global optimum is in order, to find the best combination of orientation and width for the given road element. Only the highest likelihood value is eventually retained.

Finally, not all the segments are selected. Only those with high likelihood are considered, using the above-mentioned 0.5 threshold. As a result, an image made by all the selected segments is obtained and a skeletonization and a linearization step are subsequently applied to this image in order to extract the best set of road candidates. The skeletonization procedure extracts for each blob in the thresholded image a center line, just one pixel wide. Then, the linearization algorithms reduces this line to a sequence of joint segments.

To further improve network reconstruction, a methodology to exploit network topology is required. To this aim, a closure method based on a Markovian approach defined on a graph of network elements is performed [11]. This step is essentially a labeling of the network element graph with labels “road” and “not-road,” in order to minimize an energy function. This function is derived from probabilities and from a Markovian hypothesis made on the label field. It takes both original data (likelihood term) and *a priori* knowledge about the road shape—probability of crossings and bending limitations—into account, as detailed in [11]. In the present procedure the prior (regularization) term on the network element clique is not modified compared to previous work. The likelihood term is instead defined to exploit both the line detection and the classification map. Therefore the observation field is defined as $s(d, c)$ along the considered network element, merging the line detection response d and the classification response for road class c . In other words, the joint analysis of the classification and edge extraction results is introduced in this algorithm, too. Eventually, a road network is recovered as the minimization of the labeling process.

Until now, the proposed procedure may be considered as a generalization of a road network extraction algorithm based on segment extraction routines only. This is already an improvement over existing methodologies, but a further processing step is added. As a matter of fact, the previous procedure recovers the road network at a single resolution, the original fine one. Different road widths are considered but, in order to maintain reasonable the number of options to be considered for the (w, θ) pair, it is useful to limit w to the range from one to five pixels. If high-resolution data are considered, down to 1 m or less, this may be insufficient. So, to widen the search range for road width, it may be useful to analyze the data at different resolutions [14]. This is easily done by a very simple decimating approach. It is a box filter applied to the detected data in a 2×2 window. Each window is replaced by one pixel whose value is the square root of the averaged intensity. This way was origi-

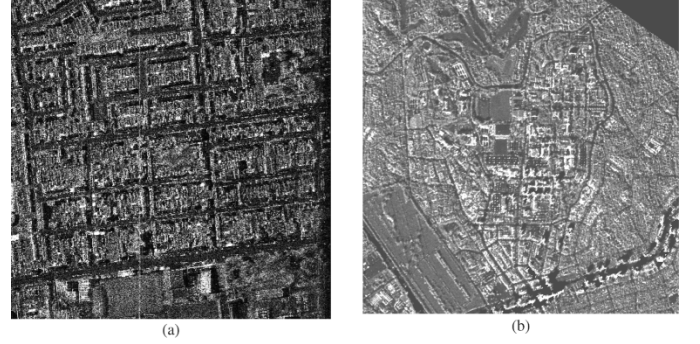


Fig. 2. High-resolution SAR images processed in this work. (a) RAMSES data over Dunquerque. (b) Star-3i data over Los Angeles.

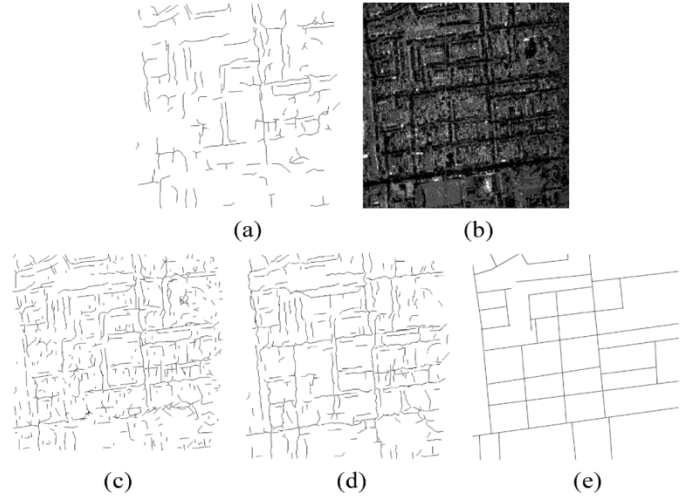


Fig. 3. Extraction results for the SAR image in Fig. 2(a). (a) Line detection. (b) Classification map. (c) Fusing the line detection and the classification map at the original fine resolution. (d) After the proposed multiresolution road network fusion. (e) Road network ground truth.

nally chosen for coarser data because it preserves the Gamma model for data distribution, and only the number of looks is modified). By this way, instead of detecting all the segment candidates and building a large graph for the connection step (and thus mixing all the networks), we prefer extracting the road network at different resolutions and then merging them. This method has the advantage of preserving the coherence of each network and produces less noisy results.

More precisely, the fusion step is made first by considering all the extracted segments at multiple resolutions. Then, in order to delete as much as possible the redundant segments in the post fusion network, a pruning procedure is applied [15]. The algorithm discards network elements extracted at different resolutions but that correspond to the same (part of a) road and preserves the longest one. This is done by comparing their directions, starting and end points. Finally, in the last step each network element is superimposed to the original full-scale image trying to move it over a little range from its original location to get the best position that covers the maximum number of dark pixels. In this way it is possible to correct small positioning errors due to the multiresolution fusion step.

III. RESULTS

The proposed method is illustrated by means of actual radar images. The first one [Fig. 2(a)] was acquired by the RAMSES

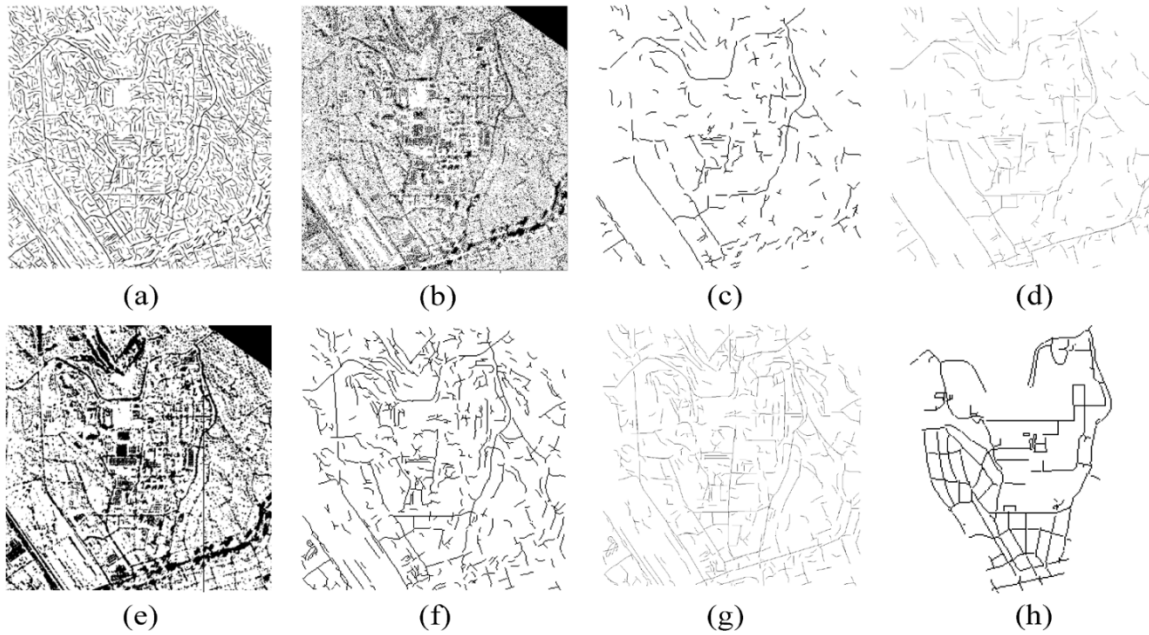


Fig. 4. Extraction results for the SAR image in Fig. 2(b). (a) Line detection. (b) Classification map by fuzzy ARTMAP. (c) Fusing the line detection and the fuzzy ARTMAP map at the original fine resolutions. (d) The corresponding multiresolution network fusion results. (e) Classification map by the Markovian classifier. (f) Fusing the line detection and the Markovian map at the original fine resolution. (g) The corresponding multiresolution road network fusion. (h) Road network ground truth.

sensor over Dunkerque (north of France) and represents a dense urban environment with mostly straight roads. The SAR data are complex, single-look slant range digital numbers with high spatial resolution (<1 m) and the area covers a nearly $2000 \text{ m} \times 2000 \text{ m}$ scene. The second image [Fig. 2(b)] was recorded by the Star-3i system, operated by Intermap Technologies Inc. It depicts an area around the University of California, Los Angeles (UCLA) campus. These data are multilook and in ground range and cover a $1250 \text{ m} \times 1250 \text{ m}$ area, with 1.25-m posting.

The proposed procedure was applied to the Dunkerque data using the line detector introduced in Section II, and the Markovian classifier. The line detection results are proposed in Fig. 3(a), while the map obtained in output to this classifier for the area of interest is shown in Fig. 3(b). As discussed above, by combining the line detectors and the classification map the most reliable segments are extracted, using the joint likelihood sum values. The road network reconstruction routine is then applied, and multiple resolutions are considered by downsampling the image by a factor of 2, 4, and 8 in each direction. Networks are finally merged and the final repositioning step is considered. The final result of this procedure is shown in Fig. 3(d), while an intermediate step is provided in Fig. 3(c), where the proposed methodology is applied without the multiresolution analysis, using only the data at the original, finest spatial resolution.

As a first comment, we may visually observe that the number of actual roads in the images increases from left to right. This strengthens our assumption that the proposed procedure improves the road network reconstruction. Furthermore, Fig. 3(d) does not present as many small segments as Fig. 3(c). The roads appear “cleaner” and more continuous, and a lot of little spurious segments has been deleted.

For the UCLA dataset, first the fuzzy ARTMAP classification was considered. In the classification map only very basic land cover classes were searched: vegetation, buildings, and roads.

The required, small training set was built using available optical images of the area but is possible to do it even using only the SAR image. The results for the initial detectors are shown in Fig. 4(a) and (b), respectively. In particular, the classification map shown in Fig. 4(b) is clearly affected by “salt-and-pepper” classification noise, due to the single-band input. Applying the proposed procedure, these two outputs were combined for improving both the candidate road selection and the MRF analysis of the road network. Moreover, the multiresolution approach was considered, and the final result is provided in Fig. 4(d), while the single resolution results are presented in Fig. 4(c). For comparison, to this dataset the Markovian classifier was also applied. The initial classification map is shown in Fig. 4(e), while single resolution and multiresolution results after the proposed procedure are depicted in Fig. 4(f) and (g). The final road network looks more precise than the original road network in both cases, but some details are better recognized using the multiresolution approach. See for instance Wilshire Boulevard, which is the large road on bottom right of the image. It is extracted very well in Fig. 4(g), while it is almost invisible in Fig. 4(f).

We also note that, despite the low classification accuracy of both classifiers with respect to the road class (24.3% for Fuzzy ARTMAP and 44.3% for the Markov classifier), the final road network in Fig. 4(d) and (g) shows almost the same amount of correct network elements. Fig. 4(g) looks however more “complete” than Fig. 4(b).

To quantify the analysis of the previous section, two quantitative indexes are shown in Table I. The correctness and completeness indexes [17] are common indexes to validate classification results. Both of them require the knowledge of the true network and provide a means to understand to what extent the extracted network is similar to the reference one. In particular, *completeness* (L_r/L_{gt}) represents the fraction of ground truth length (L_{gt}) extracted (L_r is the extracted road length), while *correct-*

TABLE I
QUANTITATIVE EVALUATION AND COMPARISON OF THE RESULTS

	fig. 3(a)	fig. 3(c)	fig. 3(d)
correctness	51.4%	75.9%	78.4%
completeness	47.6%	40.3%	46.1%
	fig. 4(a)	fig. 4(c)	fig. 4(d)
correctness	34.8%	56.5%	57.1%
completeness	76.6%	41.6%	48.7%
		fig. 4(f)	fig. 4(g)
correctness		43.6%	47.8%
completeness		62.9%	71.9%

ness (L_r/L_N) is the fraction of the total road length (L_N) belonging to actual roads.

Here the actual road network was manually extracted and is shown in Fig. 3(e) and 4(g). For this reason the values in Table I should be considered as having a relative more than an absolute meaning. In Table I, trends are more important and significant than the numbers. It is interesting to observe that correctness increases from left to right, as we could expect. Completeness instead is bigger for the networks extracted with the original algorithm. However, this is mainly due to the fact that less small, spurious roads are present in such images due to a larger number of false detections. Difference in completeness and correctness absolute values in the two examples are due to the different street networks in the area. A look at the UCLA maps shows that we have far less roads in the final stage of our approach, and the remaining ones delineate the structure of the network much clearly than using the original data alone. However, for this site the classification maps are obtained from multilook ground range data, which suffers from the lack of the full information carried by the complex SAR signal. Thus, the correctness results are not as good as in the Dunkerque example. In other words, the differences between the two situations are due both to the data and to the different behaviors of the classification routines. In the second situation the classification does not add much information to the original set of segments, resulting only in a partial advantage. So, the better results obtained using the classification is mostly due to the reduction of false positives, i.e., edges that are classified as streets if no map is available.

A final note is deserved to the use of multiple resolutions. More resolutions means more computations, and therefore more CPU time. However, their exploitation always improves the results with respect to the single-resolution analysis. These examples run in 5 min for a personal computer with a Pentium IV (single resolution) and in 10 min (multiresolution). Generally speaking, we may say that this step is worth the effort.

IV. CONCLUSION

This letter presents a method for road network detection from high-resolution SAR data that includes a data fusion procedure in a multiresolution framework. It takes into account the information available by both a line detector and a classification algo-

rithm to improve the road segment selection and the road network reconstruction. To this aim, in the Markovian approach used for solving the global network optimization problem (with constraints), the clique potentials were modified to exploit all the available knowledge. Moreover, a multiresolution fusion approach was exploited, and we showed that it is able to increase the percentage of actual roads while reducing missing ones. Another advantage is the merging of detections of the same road at different resolutions; gaps due to missed parts are also reduced. This is clear for instance in Fig. 3(d) where the number of these gaps is clearly lower than the Fig. 3(c).

REFERENCES

- [1] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 11, pp. 679–698, Nov. 1986.
- [2] M. A. Fischler, J. M. Tenenbaum, and H. C. Wolf, "Detection of roads and linear structures in low resolution aerial imagery using a multisource knowledge integration technique," *Comput. Graph. Image Process.*, vol. 15, no. 3, pp. 201–223, 1981.
- [3] J. Chanussot, G. Mauris, and P. Lambert, "Fuzzy fusion techniques for linear features detection in multitemporal SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 2287–2297, May 1999.
- [4] R. Huber and K. Lang, "Road extraction from high resolution airborne SAR using operator fusion," in *Proc. IGARSS*, Sidney, Australia, 2001, pp. 2813–2815.
- [5] B. Wessel, "Context-supported road extraction from SAR imagery: transition from rural to built-up areas," in *Proc. EUSAR 2004*, Ulm, Germany, May 2004, pp. 399–402.
- [6] L. Pigeon, B. Solaiman, K. P. B. Thomson, B. Moulin, and T. Toutin, "Human-experts rules modeling for linear planimetric features extraction in a remotely sensed images data fusion context," in *Proc. IGARSS*, Hamburg, Germany, Jun. 28–Jul. 2 1999, vol. 5, pp. 2369–2371.
- [7] F. Dell'Acqua and P. Gamba, "Detection of urban structures in SAR images by robust fuzzy clustering algorithms: The example of street tracking," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 10, pp. 2287–2297, Oct. 2001.
- [8] F. Dell'Acqua, P. Gamba, and G. Lisini, "Extraction and fusion of street network from fine resolution SAR data," in *Proc. IGARSS*, Toronto, ON, Canada, Jun. 2002, vol. 1, pp. 89–91.
- [9] D. Borghys, C. Perneel, and M. Achery, "A multivariate contour detector for high-resolution polarimetric SAR images," in *Proc. 15th Int. Conf. Pattern Recognition*, Sep. 3–7, 2000, vol. 3, pp. 646–651.
- [10] C. Tison, J. M. Nicolas, F. Tupin, and H. Maître, "A new statistical model of urban areas in high resolution SAR images for Markovian segmentation," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 10, pp. 2046–2057, Oct. 2004.
- [11] F. Tupin, H. Maître, J.-F. Mangin, J.-M. Nicolas, and E. Pechersky, "Detection of linear features in SAR images: Application to road network extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 36, no. 2, pp. 434–453, Mar. 1998.
- [12] R. Touzi, A. Lopes, and P. Bousquet, "A statistical and geometrical edge detector for SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 26, no. 6, pp. 764–773, Nov. 1988.
- [13] I. Bloch, "Information combination operators for data fusion: A comparative review with classification," *IEEE Trans. Syst., Man, Cybern.*, vol. 26, no. 1, pp. 52–67, Jan. 1996.
- [14] F. Tupin, B. Houshmand, and M. Datcu, "Road detection in dense urban areas using SAR imagery and the usefulness of multiple views," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 11, pp. 2405–2414, Nov. 2002.
- [15] F. Dell'Acqua, P. Gamba, and G. Lisini, "Road map extraction by multiple detectors in fine spatial resolution SAR data," *Can. J. Remote Sens.*, vol. 29, no. 4, pp. 481–490, Aug. 2003.
- [16] J. W. Goodman, "Some fundamental properties of speckle," *J. Opt. Soc. Amer.*, vol. 66, no. 11, pp. 1145–1150, 1976.
- [17] C. Wiedemann and H. Ebner, "Automatic completion and evaluation of road networks," *Int. Arch. Photogramm. Remote Sens.*, vol. 33, pp. 976–986, 2000.
- [18] P. Gamba and F. Dell'Acqua, "Improved multiband urban classification using a neuro-fuzzy classifier," *Int. J. Remote Sens.*, vol. 24, no. 4, pp. 827–834, Feb. 2003.