A First Step Toward Automatic Interpretation of SAR Images Using Evidential Fusion of Several Structure Detectors

Florence Tupin, Isabelle Bloch, and Henri Maître

***Abstract*— We propose a method aiming to characterize the spatial organization of the main cartographic elements of a synthetic aperture radar (SAR) image and thus giving an almost automatic interpretation of the scene. Our approach is divided into three main steps which build the whole image interpretation gradually. The first step consists of applying low-level detectors taking the speckle statistics into account and extracting some raw information from the scene. The detector responses are then fused in a second step using Dempster–Shafer theory, thus allowing the modeling of the knowledge that we have about operators, including possible ignorance and their limits. A third step gives the final image interpretation using contextual knowledge between the different classes. Results of the whole method applied to different SAR images and to various landscapes are presented.**

***Index Terms*—Dempster–Shafer evidence theory, image interpretation, Markov random fields, SAR images.**

I. Introduction

THE increasing number of synthetic aperture radar (SAR) sensors, and as a consequence, of SAR data, calls for the development of automatic or semiautomatic tools to help the human interpreter. The aim of the method we propose here is to give an elementary but almost automatic interpretation of a SAR image. By interpretation, we mean giving the spatial organization of the main cartographic elements in a SAR image: road and hydrological networks, urban areas, forest or sea areas, relief, etc. This tool must be able to work on many different geographical areas, with different possible soil occupations, and with images originating from different radar sensors. In order to adapt the detection to the signal properties (resolution, noise level, ...), we allow a supervised learning stage. After this stage, we expect the program to work blindly on many different scenes of the same sensor. Such a tool is useful for many applications. It can be used as an initial analysis of the image to select some particular areas of interest (urban areas, for instance). The obtained interpretation can also be a starting point to automatically register data coming from different sources (radar or SPOT images, but also symbolic information like that present in a map).

Many works have been dedicated to the problem of image interpretation in aerial or satellite imaging. Those with the closest objectives to ours are mostly using rule-based systems. SPAM [26] for aerial images of airports, SIGMA [25] for aerial urban images, MESSIE [6] for satellite images, ICARE [9] for multisensor images, are instances of such systems where declarative statements are made, first so as to deduce conclusions from observed features, second to propose strategies to link together successive statements. But rule-based systems may be rather difficult to develop and in order to be efficient, require a large basis of rules that is hardly available in the case of SAR image processing. Indeed, in the case of SAR satellite images with a resolution of 22.5 m like ERS-1 images, only a few cartographic elements can be detected: main hydrological or road networks, industrial areas, and a few kinds of vegetation. This is not exactly the case for polarimetric radars or aerial images for which rule-based systems may be better adapted.

Another family of methods is made of graph-based techniques which aim at labeling the regions of the image (nodes of the graph) and their relationships (arcs). These approaches may be used either in matching the image with a model of the scene when available [13], [12], [20], [5], [14], or, when no model is available, to give a global consistency to the graph interpretation by some kind of relaxation on the graph [1], [23], [22]. In our case, the second track is followed to introduce contextual relationships between the different classes (i.e., between the different geographical structures).

In any interpretation system, pieces of information extracted from different detectors or sources have to be combined to provide the elements of interpretation. This combination may be performed by numerical or symbolic methods, depending on the degree of decision provided by the detector. Radar images are coarse and noisy. Reliable detections are very difficult to obtain, and we expect to benefit from the combination of several detectors to improve the decision. Many theoretical frameworks have been proposed in this case: the oldest and most popular one is the probability theory, but other frameworks can be used like fuzzy set theory [36], [10], possibility theory [37], [11], and evidence theory [29]. In our case, numerical fusion of results obtained from low-level operators designed for specific structure detection is done in the evidence theory framework. This theory is particularly adapted for operators which are imprecise and unable to distinguish all the classes (which is often the case in SAR imagery).

Many interpretation systems exist to deal with optical satellite images, but not in the case of SAR data, for which very few works aim at developing global interpretation methods using all the available information. The rule-based system MESSIE, for instance, combines radar and optical-range images to locate urban areas or man-made objects [6], but according

Manuscript received June 2, 1998; revised November 17, 1998.

The authors are with École Nationale Supérieure des Télécommunications, Department TSI, CNRS URA 820, 75013 Paris, France (e-mail: tupin@ima.enst.fr; bloch@ima.enst.fr; maitre@ima.enst.fr).

Publisher Item Identifier S 0196-2892(99)03449-X.

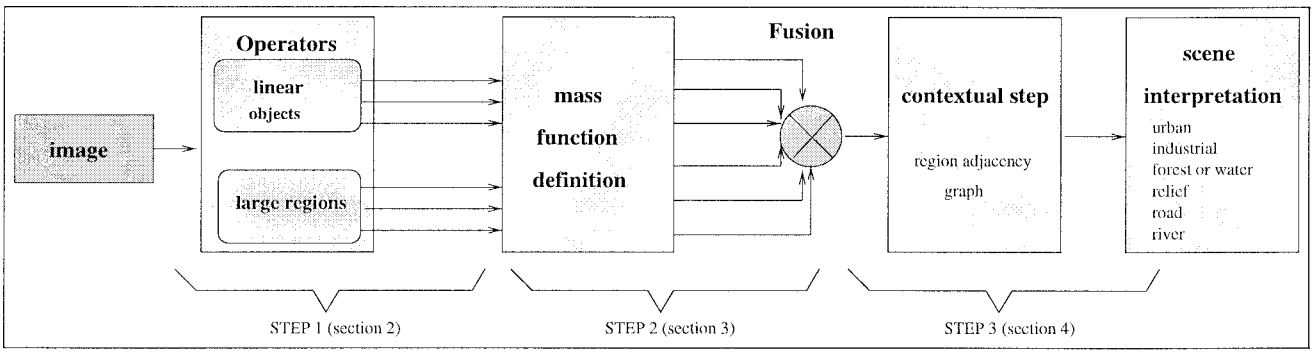


Fig. 1. Diagram of the proposed approach.

to the authors themselves, the radar data is essentially used as complementary data. Recent works on radar images are dedicated to the detection of particular features (road networks [27], [18], urban areas [19]). Other works take place in the classification framework, working at the pixel level, particularly with polarimetric and multifrequency radar images [28], [24], [7]. As opposed to these approaches, the aim of this paper is to present a global scheme dealing simultaneously with all the available information in order to provide a global interpretation of the scene. A complete method is presented from the low-level stage to the final interpretation result, which allows us to provide a new kind of result on SAR images.

The proposed approach is divided into three main steps. In the first step, some low- or intermediary-level tools have been developed. Each of them aims at giving information on a particular scene structure: linear features like roads, rivers, or the bright lines appearing in areas with high variations in elevation, or larger objects like urban areas, forests, or water areas. These operators are presented briefly in Section II. More details can be found in [17] and [32]. Then in the second step, which constitutes the core of this paper, each tool (or operator) is considered as a source of knowledge which gives its confidence on the possible presence of the object it is able to detect, and all these sources are combined in the evidence theory framework (Section III). Instead of working at the pixel level, which would not be justified because of the low operator accuracy, a set of regions is considered, simplifying the problem and speeding up the fusion process. The decision step, which actually classifies each region of the image, is eventually achieved using contextual relationships between the regions in a Markovian framework (Section IV). The classes we consider in this application are the following: *urban*, *industrial*, *homogeneous* $= \{$*forest* or *sea*$\}$, *relief*, *road*, *river* and *bright.field*. As will be seen further, the *bright.field* class is only introduced to avoid false alarms of the urban class. The diagram of the method is presented in Fig. 1. Results on real radar images are then analyzed in Section V.

II. Operator Description

All the operators used in our interpretation scheme are described in the two following subsections.

A. Operators for Linear Structures

This part is based on our previous work on the road and hydrological network detection [32], [17]. The main steps of the method are summed up here. They explain how confidence measures are defined. The method is divided into two main steps.

- In a local step, a line detector adapted to the speckle statistics of SAR images is applied (thresholding and linking provide segments that are candidates to belong to the network).
- In a global step, a closure method based on a Markovian approach defined on a graph of segments is performed; this step is a labeling of the segment graph with labels "road" and "not-road" (or "river" and "not-river") minimizing an energy function; this function, derived from probabilities and from a Markovian hypothesis made on the label field, takes both original data and *a priori* knowledge about the road shape (probability of crossings and bending limitations) into account.

The reader may refer to [32] to have a detailed description of the method.

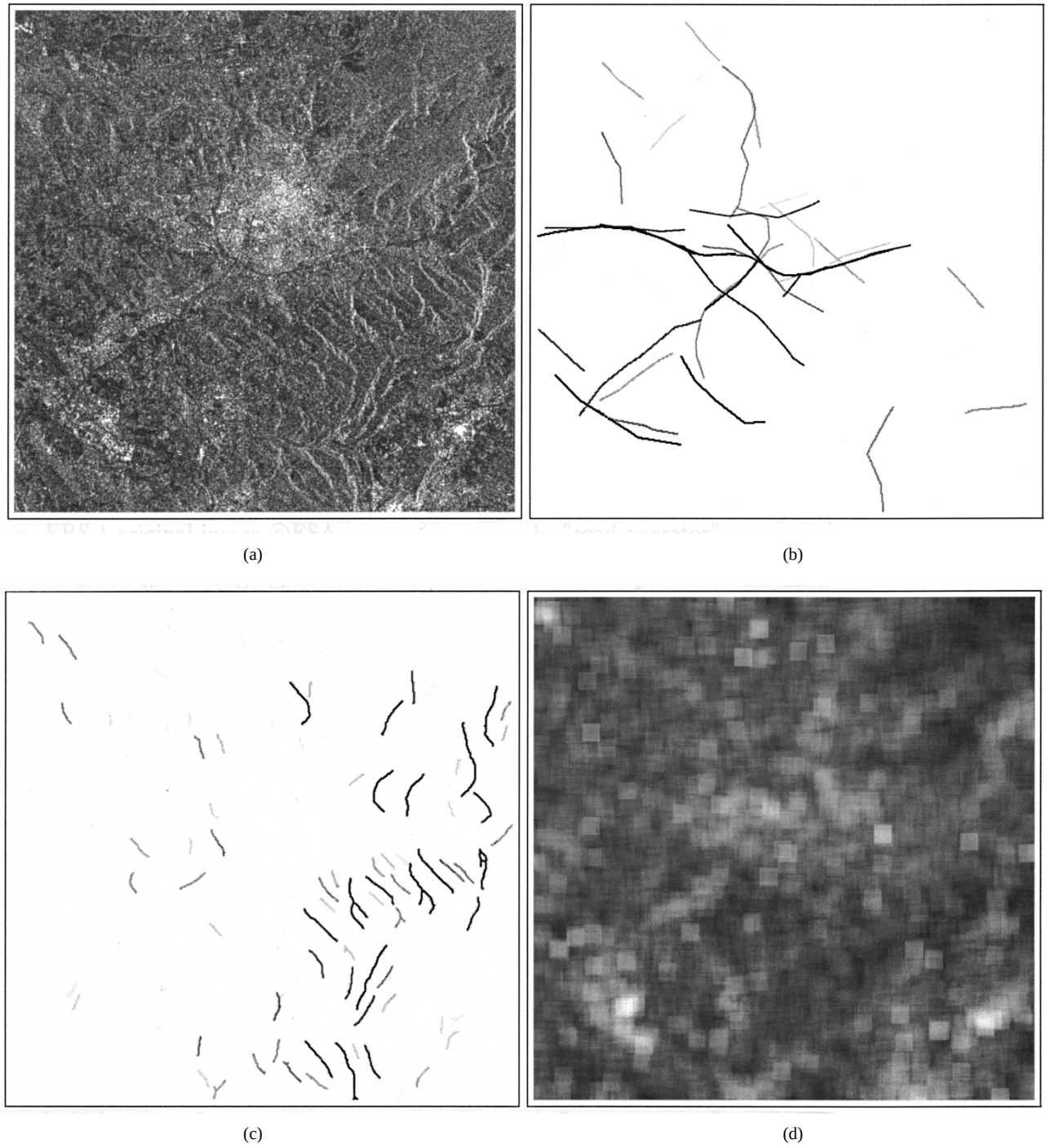
◇ "**road operator**:" the two steps mentioned above are applied to obtain a "map" of the roads which are defined as chains of connected segments. The energy U_{road} of a chain can then be computed using the defined Markovian field as well as the energy $U_{\text{not-road}}$ of the chain with the segments labeled "not-road." The energy variation $\Delta U = U_{\text{not-road}} - U_{\text{road}}$ is positive on a road,[1] and the larger this variation, the stronger the confidence we have in the detected road. So we associate to each detected chain i a confidence measure ΔU_i taking its value in $\mathbf{R}^+ = [0, +\infty[$. The result of this operator applied on an ERS-1 image [Fig. 2(a)] is shown in Fig. 2(b).

◇ "**river operator**:" the method is very similar to the road detection method, but since a river may have different widths along its course, we used the detection algorithm in a multiscale way [33]. The same confidence measure as before is defined, also taking its value in $\mathbf{R}^+$.

◇ "**relief operator**:" bright lines due to signal fold-over are typical of radar images in relief areas. They appear when the slope of the terrain facing the radar exceeds a limit related to the beam incidence angle. In order to detect these lines, only the low-level line detection algorithm is applied with a modification to select only the bright lines of the image (on the contrary, rivers and most of the roads appear as dark).[2] We

[1] Indeed, the final labeling must correspond to the global energy minimum.

[2] Roads can also appear bright on radar images in some very particular cases of orientation.



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

Fig. 2. Operators on a SAR ERS-1 image of Aix-en-Provence (France); (a) ERS-1 original image © ESA, (b) "road operator," (c) "relief operator," (d) "σ_{MAR} operator."

define the confidence measure associated with each detected line by combining the length of the line and the mean of the line detector responses along the line. The combination operator is an associative symmetrical sum [2] reinforcing the response of each parameter. This measure is defined in the interval $[0, 1]$. The result of this operator is shown in Fig. 2(c).

B. Operators for Large Areas

The aim of the following operators is to measure the degree of heterogeneity or homogeneity of the regions appearing in the radar image. Our objective is to distinguish very homogeneous areas, like forests or sea areas with fully developed speckle, and very heterogeneous areas like industrial ones with a high density of bright points due to reflections on buildings. Since we also want to detect urban areas and since very dense urban areas are very bright but rather homogeneous regions (differing on this point from "industrial" ones [17]), we also detect regions with very high radiometry. Therefore, the three following operators are used in our interpretation scheme.

◇ "σ_{MAR} **operator**:" is a textural operator derived from a modeling of SAR images by a log-normal multiplicative

autoregressive model (MAR) proposed by Frankot and Chellappa [15]. From this model, the standard deviation σ_{MAR} of the white Gaussian noise is a good indicator for region homogeneity: heterogeneity increases with this value. Fig. 2(d) presents σ_{MAR} values on a SAR image. The top right corner of the image, corresponding to forest areas, is darker than other regions (low values of σ_{MAR}). This parameter is computed on a 4×4 averaged image and using analysis windows of 9×9 pixels. This value is by itself a confidence value since it increases with heterogeneity and takes its values in $\mathbf{R}^+$.

◇ "**bright reflector operator**:" this operator aims at detecting regions with a high density of bright points. First of all, a Markovian segmentation method is applied on the radar image [21] to provide radiometrically homogeneous regions. The brightest areas are then selected if their size is less than a fixed value (ten pixels); these areas are considered in the following as strong reflectors (for instance, a specular reflection due to a building). In order to select the areas where these points are very dense, the Voronoi diagram of the set of points is computed [Fig. 3(a)]. Selecting small Voronoi regions corresponding to areas with many bright reflectors (with a size less than 500 pixels) and grouping them with a morphological closing gives us industrial zone candidates. At last, the large enough candidates are selected (size of more than 1000 pixels). The confidence measure of a candidate is proportional to the number of bright points in a region and takes its values in $[0, 1]$.

◇ "**ffmax operator**:" as said before, in really dense urban areas, the previous detector will fail since these areas are rather homogeneous. Therefore, we make use of a radiometric detector due to Gouinaud [17]. It analyzes the histogram queues and detects very bright areas in an image. This method is based on a local histogram splitting defined in such a way that the population of the upper part equals the mode of the histogram. The output of the "ffmax operator" is this value where the histogram is split. It is expressed as a multiple of the signal standard deviation and may be quantized on ten levels.

The three operators which are defined above do not precisely locate the detected areas since they use large windows to compute local statistics, or Voronoi cells. Therefore, our interpretation scheme does not work at the pixel level, but on a set of regions. These regions are defined using an over-segmentation based on the Markovian segmentation we mentioned above [21] [see Fig. 3(c)]. The Markovian model used for the label field is a Potts model, thus linear features (roads, rivers) are not very well detected. Therefore, the linear objects, detected using the linear operators of Section II-A are added to the Markovian segmentation and superimposed on the set of regions to define the primitives of our interpretation scheme [Fig. 3(d)].

The image structure is now ready for information fusion. The image is represented as a graph of regions with an adjacency relationship between regions. For each and every region, each and every operator provides a numerical value which reflects the confidence the operator has in the region to belong to the object class that the operator is qualified to detect. The value for a region is taken as the average of the detector responses on the pixels belonging to the region. From now on this value is called the confidence value of the operator about the region.

III. Numerical Fusion Step in the Evidence Theory Framework

This section presents the combination of the different detector responses by mean of Dempster–Shafer theory to interpret the SAR images. First, we briefly recall the evidence theory principles. Then we describe how we define the mass functions associated with the operators which are presented in Section II, starting from the operator confidence measure. The unnormalized Dempster rule of combination is then applied, the empty set representing a rejection class. A justification of the choice of the evidence theory is eventually given with an illustration in a simplified case.

A. Evidence Theory Principles

One of the main advantages of the evidence theory as proposed in [8] and [29] is its capability of taking both imprecision and uncertainty into account. Let Θ be the set of discernment (for us it is the set of all the classes mentioned above), and 2^Θ be the power set of Θ, which contains all the possible unions of classes. The evidence theory is based on three functions defined from 2^Θ onto $[0, 1]$: mass function m, belief function *Bel* and plausibility function *Pls* [29]. The quantity $m(A)$ corresponds to the measure of belief that is exactly committed to A, *Bel*(A) the measure of the total belief committed to A, and *Pls*(A) the extent to which A is plausible. In the case of many sources of information defined on a same frame of discernment by their mass functions, the resulting mass function of the information fusion is obtained by the orthogonal rule of Dempster [29]. As we will see in the following section, in the case of a nonexhaustive frame of discernment, the orthogonal rule can be used without normalization [30], thus allowing a natural definition of a reject class.

The orthogonal rule is commutative and associative which permits combining information sources in any order without changing the result (it will be a useful property for the application we deal with).

The last step of the fusion process is to take a decision. Since for each hypothesis, both belief and plausibility can be used, the choice is closely related to the application at hand. The maximum of belief or plausibility on the simple hypotheses is often chosen. Since we will use a Markovian framework within the probability theory, we propose here to use the "pignistic probability" *BetP* defined by Smets [31]:

$$BetP(\{h_i\}) = \sum_{A/\{h_i\} \subset A} \frac{m(A)}{card(A)} \tag{1}$$

where $\{h_i\}$ is a simple hypothesis, and *card* stands for cardinality. This transformation shares out the mass of a compound hypothesis on its singletons.

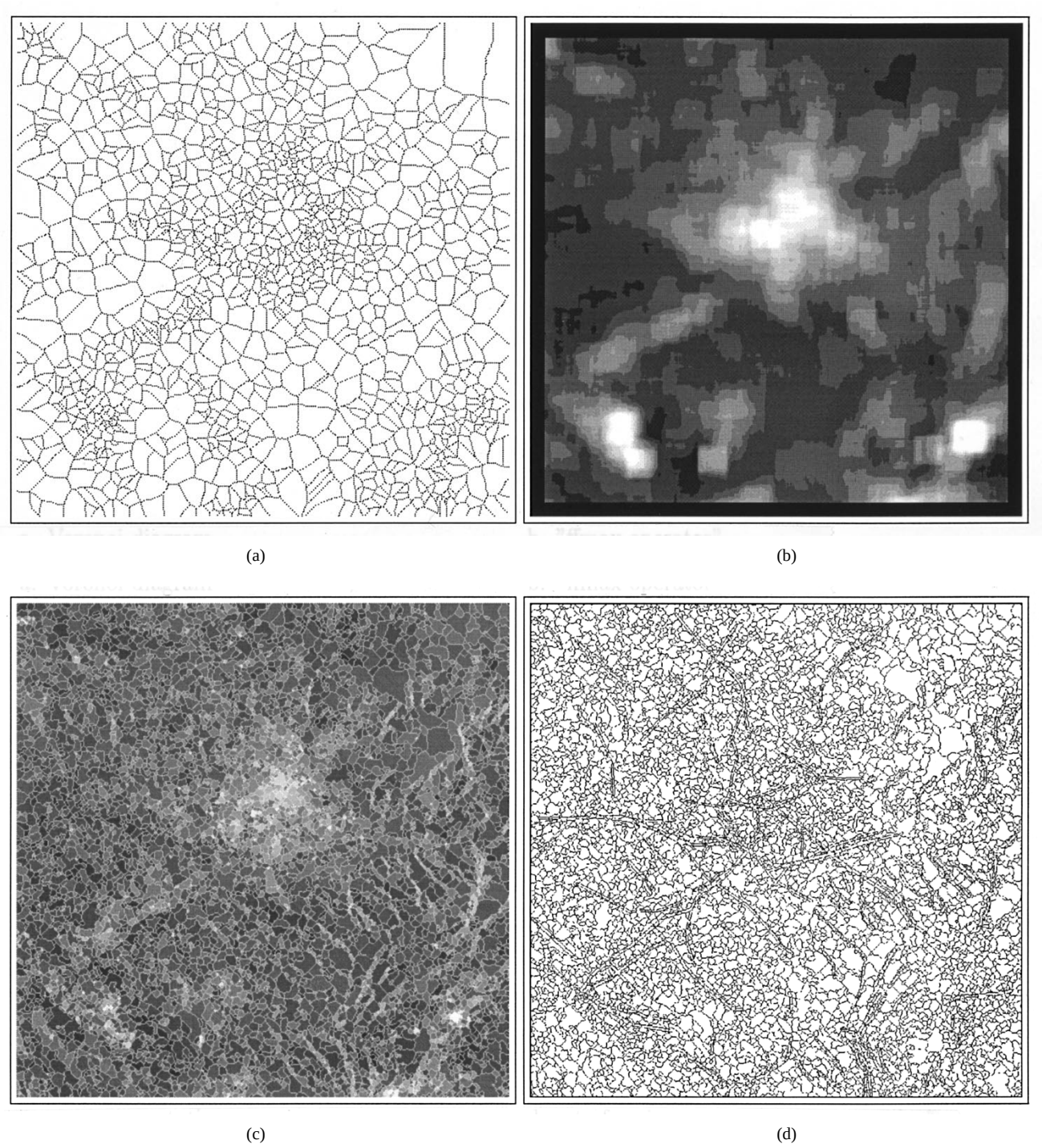


Fig. 3. Operators and region definition on a SAR ERS-1 image of Aix-en-Provence (France); (a) Voronoi diagram, (b) "ffmax operator," (c) over-segmentation, (d) set of regions.

B. Mass Function Definition

The operators or sources of knowledge we presented previously (Section II) give information on linear structures ("road operator," "river operator," "relief operator") and on large areas ("ffmax operator," "bright reflector operator," "σ_{MAR} operator"). Each of them gives a measure of confidence in the detection of the objects it is dedicated to. Therefore, we now have to define the mass functions starting from the operator confidence measures.

The choice of the mass functions is the crucial step of our fusion scheme. In fact, all the imprecision of the data must be introduced in this step since the fusion operator (Dempster rule) is always conjunctive. It is a general feature of Dempster–Shafer fusion that most of the flexibility lies at the modeling level [3].

We have chosen to define the mass functions in a supervised way using our knowledge of the operator behaviors. This modeling step has to be done only once and the same mass functions have proved to be well adapted to many radar sensors

(if the same operators are used), as will be seen with the results on real radar images (Section V).

Let us note that unlike most of the image applications of Dempster–Shafer, the focal elements and mass functions are not defined by the data properties [22], [24], but by the properties of the low-level operators we defined in the first stage.

1) Focal Element Definition: First of all, we have to define the focal elements for each operator. For an operator dedicated to the detection of the class C_i, the simplest idea is to take as focal elements C_i and its complementary $\bar{C}_i$. In practice, unfortunately, the operators are often not precise enough to use only these two focal elements, and other classes have to be taken into account. Many situations occur: for instance, the operator is defined in such a way that many classes are confused and indistinguishable; besides, many operators are not able to precisely localize the classes and can also detect the surrounding classes. Let us justify for each previous operator the choice of the focal elements.

- "Road operator," "relief operator," and "river operator" are quite precise for the detection of the structure they are dedicated to; besides, all that can be said about regions with low responses is that they belong to "not road" (or "not relief," "not river") without distinction among the classes composing "not road" ("not relief," "not river," respectively); thus, the focal elements are, respectively, *road* and $\overline{road}$, *relief* and $\overline{relief}$, and *river* and $\overline{river}$ (where $\bar{A}$ denotes the complement of A).
- "Bright reflector operator" is not accurate to localize the classes since it groups large regions around specular reflectors to characterize industrial areas. Therefore, this operator has high responses not only in *industrial* regions, but also in *urban* areas and on the roads which go through these regions. Thus, it is not able to distinguish the classes *industrial*, *urban*, and *road*. On the other hand, the regions with weak responses surely do not belong to *industrial* and are in the $\overline{industrial}$ class. This operator gives information on (*industrial* $\cup$ *urban* $\cup$ *road*) and $\overline{industrial}$ which are thus the focal elements.
- "ffmax operator" has imprecisely localized responses since the histograms are computed on windows of 40 $\times$ 40 pixel size. With the same argument as before, it is unable to distinguish *urban*, *industrial* and *road* classes. Besides, by definition, this operator detects very bright regions, which means urban areas but also very bright fields. Therefore the focal elements are (*industrial* $\cup$ *urban* $\cup$ *road* $\cup$ *bright.field*) and ($\overline{urban \cup bright.field}$).
- "σ_{MAR} operator" is not very accurate since its responses are computed on a 9 $\times$ 9 pixel window, and it is applied on a radar image of reduced size (obtained by block averaging). Thus, fine roads going through forests or homogeneous areas are not detected. On the other hand, regions of high response for this operator are surely inhomogeneous. Focal elements are thus: (*road* $\cup$ *homogeneous*) and $\overline{homogeneous}$.

2) Mass Function Definition: The choice of the focal elements is the first part of the mass function definition. We now have to associate to each operator response and to each focal element the mass function value. To do so, a learning step on selected areas is used.

Since our operators are especially dedicated to the detection of a class of interest, the task is essentially reduced to the definition of the good thresholds between "weak" and "high" responses and the transition area we allow between these two extreme values. We select regions of interest to perform the learning step and deduce these thresholds. Mass functions of trapezoidal shape are used and give good results. Since no difficulties are related to this point, we do not describe it in detail, but only relate the process for the example of the "ffmax operator."

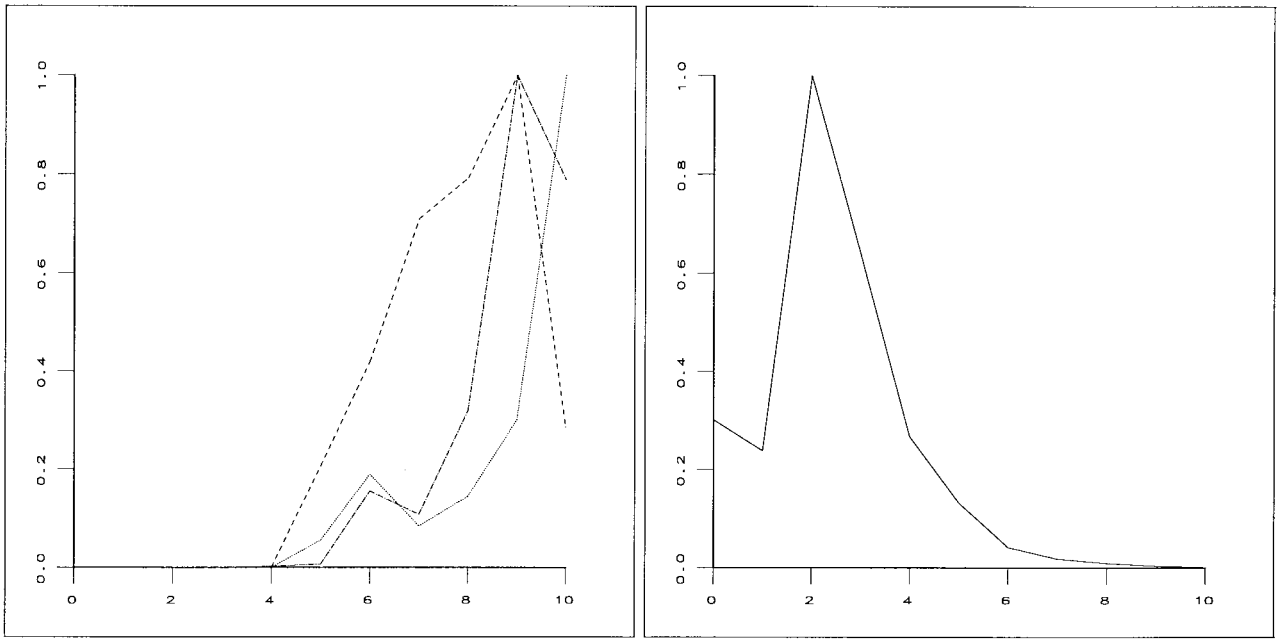
For this operator, samples of bright and homogeneous areas have been selected by the user and the normalized responses on these samples and on the rest of the image are shown in Fig. 4(a) and (b), respectively. The corresponding mass functions are simply and experimentally deduced from the histograms by fitting trapezoidal functions.

The parameter values we used for the other mass functions in the case of ERS-1 images are given in Table II using the notations of Table I with the definitions of g_1: $g_1(x) = 0$ for $x < a$, $g_1(x) = 1$ for $x > b$, $g_1(x) = \frac{x-a}{b-a}$ for $x \in [a, b]$ and g_2, $g_2(x) = 1 - g_1(x)$. Since there are only two focal elements for each operator, the type of the mass function and its parameters a and b are given only for one focal element; the complementary function is chosen for the other focal element with the same parameters (for instance if g_1 is used for the focal element *Re* of the "relief operator," g_2 is the mass function of $\overline{Re}$). The result shown in Section V is obtained using these values.

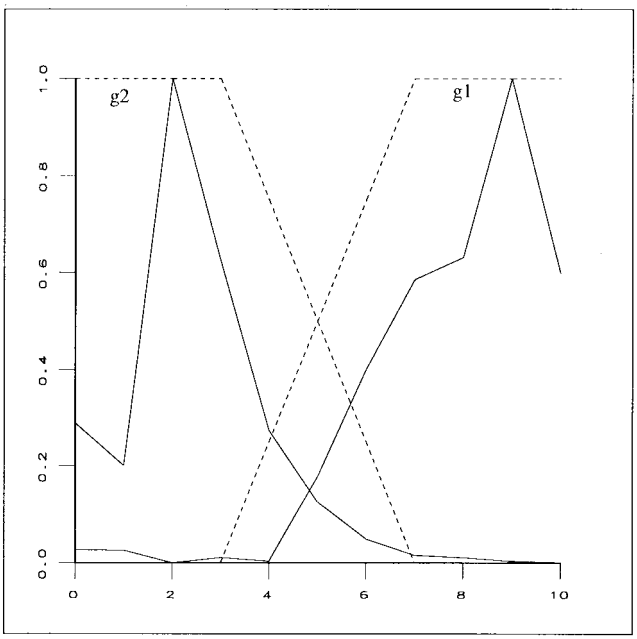
The parameters we used are rather robust. Indeed, they have been learned on a reduced set of samples provided by only two ERS-1 images (on the whole four forest samples, eight industrial samples, 11 dense urban samples, and the "complementary" samples provided by the rest of the image for each focal element). They have then been tested on all our ERS-1 data base and have given satisfying results. As for the other sensors, the set of learned parameters has been modified using our *a priori* knowledge about the considered sensor (particularly its resolution). The main modification concerns the "σ_{MAR} operator" for forest detection. The robustness of the parameters is due both to the operator behavior (each operator detecting few classes) and to the transition areas of the trapezoidal functions.

3) Discussion on the Mass Function Choice: The choice of the focal elements and the mass functions has been done in a supervised way using our knowledge about the information provided by each operator. We argue that for the application at hand, i.e., the automatic interpretation of SAR images, these modeling and learning steps are justified. It is done only once and then any radar image can be treated. Of course, this reasoning approach is only possible because we are interested in few classes, and we deal with relatively few operators.

Automatic methods have been proposed for the choice of the focal elements, or the mass function definition. For instance, a method based on the comparison of two classification results is used in [24] to automatically define the focal elements. But



(a) (b)



(c)

Fig. 4. Learning step for the "ffmax operator." Histogram of the normalized responses of "ffmax operator" on samples corresponding to its focal elements (a) ($industrial \cup urban \cup road \cup bright.field$) (three samples) and (b) ($\overline{urban \cup bright.field}$) (one sample). (c) Histograms of the normalized responses for ($industrial \cup urban \cup road \cup bright.field$) and ($\overline{urban \cup bright.field}$) samples (the three samples for the learning of the first focal element of (a) are merged). The mass functions of trapezoidal shape are derived from the histograms $m_{\mathrm{ffmax}}(industrial \cup urban \cup road \cup bright.field)$ and $m_{\mathrm{ffmax}}(\overline{urban \cup bright.field})$.

such a reasoning is not adapted to our problem since too many classes would be introduced (for instance, the class "*road in urban area*," "*road in forest*," etc.) and many of them would not have a cartographical meaning. As for the choice of the mass functions, this coarse model with functions of trapezoidal shape was found to be sufficient and is robust since it allows some variations for the parameters of the functions. This step could be done in an automatic way [4] and some preliminary tests have shown good results.

C. The Use of the Orthogonal Rule for Combination

The fusion of the operator responses is done using the Dempster rule of combination. Let m_j be the mass function of

TABLE I
CLASS AND OPERATOR NOTATIONS

Re	*relief* class
Ro	*road* class
Ri	*river* class
H	*homogeneous* = *forest* ∪ *sea* class
U	*urban* class
I	*industrial* class
BF	*bright.field* class
O_1	"relief operator"
O_2	"road operator"
O_3	"bright reflector operator"
O_4	"river operator"
O_5	"ffmax operator"
O_6	"σ_{MAR} operator"
$O_{1\ldots k}$	operator resulting of the combination of k operators

TABLE II
OPERATORS AND MASS FUNCTION PARAMETER (g_1: $g_1(x) = 0$ FOR $x < a$, $g_1(x) = 1$ FOR $x > b$, $g_1(x) = \frac{x-a}{b-a}$ FOR $x \in [a, b]$, AND g_2, $g_2(x) = 1 - g_1(x)$

Operator	Focal element	mass function	a	b
O_1	Re	g_1	0	0.2
O_2	Ro	g_1	0	0.3
O_3	$Ro \cup V \cup I$	g_1	0	1
O_4	Fl	g_1	0	0.4
O_5	$Ro \cup V \cup I \cup Cl$	g_1	3	7
O_6	$Ro \cup H$	g_2	50	100

source j and A an element of 2^Θ, the resulting mass function $m_{1\ldots n}$ for n distinct sources is then

$$m_{12\ldots n}(A) = \sum_{B_1, B_2 \ldots, B_n / B_1 \cap B_2 \cap \cdots \cap B_n = A} m_1(B_1) m_2(B_2) \ldots m_n(B_n)$$

Since our hypotheses do not entirely describe a radar image (many pixels do not belong to the considered classes), normalization is not adapted to our case and could lead to incoherent situations as described by many authors [34], [35], [11], [30]. Besides, we can define a rejection class in a very natural way, using the conflict value as the mass of the empty set.

Table III shows the focal elements resulting from the combination of the operators. Since the Dempster rule is commutative and associative, the resulting elements do not depend on the order of the combination. We performed a combination in five steps following the given indications. We use the notations indicated in Table I.

Thus, many classes appear during the combination process, although they are never detected directly by the operators. Besides, the operators allow us to distinguish between all the classes we consider (except between *urban* and *bright.field*), since these classes are focal elements at the end of the combination.

The final decision is made using contextual information. Since this is done in a Markovian framework (see Section IV), we need to reduce the final mass functions to a probability function. This is performed using the pignistic transformation using the Smets formula given in Section III-A [(1)].

TABLE III
OPERATOR COMBINATION

Step 1.relief⊕road

$O_1 \setminus O_2$	Re	$\overline{Re}$
$\overline{Ro}$	$\emptyset$	Ro
$\overline{Ro}$	Re	$\overline{Ro} \cap \overline{Re}$

Step 2. (relief⊕road)⊖river

$O_3 \setminus O_{12}$	Ro	Re	$\overline{Ro} \cap \overline{Re}$
Ri	$\emptyset$	$\emptyset$	Ri
$\overline{Ri}$	Ro	Re	$\overline{Ro} \cap \overline{Re} \cap \overline{Ri}$

Step 3. (relief⊕road⊕river)⊕bright.reflector

$O_4 \setminus O_{123}$	Ro	Re	Ri	$\overline{Ro} \cap \overline{Re} \cap \overline{Ri}$
$Ro \cup U \cup I$	Ro	$\emptyset$	$\emptyset$	$U \cup I$
$\overline{I}$	Ro	Re	Ri	$H \cup U \cup BF$

Step 4. (relief⊕road⊖river⊕bright.reflector)⊖ffmax

$O_5 \setminus O_{1234}$	Ro	Re	Ri	$U \cup I$	$H \cup U \cup BF$
$Ro \cup U \cup I \cup BF$	Ro	$\emptyset$	$\emptyset$	$U \cup I$	$U \cup BF$
$\overline{U} \cap \overline{BF}$	Ro	Re	Ri	I	H

Step 5.(relief⊖road⊕river⊕bright.reflector⊕ffmax)⊕σ_{MAR}

$O_6 \setminus O_{12345}$	Ro	Re	Ri	I	H	$U \cup I$	$U \cup Cl$
$Ro \cup H$	Ro	$\emptyset$	$\emptyset$	$\emptyset$	H	$\emptyset$	$\emptyset$
$\overline{H}$	Ro	Re	Ri	I	$\emptyset$	$U \cup I$	$U \cup BF$

Resulting focal elements	Ro	Re	Ri	I	H	$U \cup I$	$U \cup BF$

D. Advantages of Dempster–Shafer Theory

This section aims at justifying the choice of Dempster–Shafer theory to fuse the operators in the case of SAR image interpretation. One of the main advantages of the evidence theory is its capability to take into account compound hypotheses, i.e., union of classes here [3]. In the Bayesian framework, the degree of belief we have on a union of classes (without being able to discriminate between them) should be shared by all the simple hypotheses, thus penalizing the good one. This point is particularly important in image interpretation, since the operators are usually unable to distinguish all the different classes in a precise way. Let us present an example to illustrate the different behaviors of the fusion step in these two frameworks.

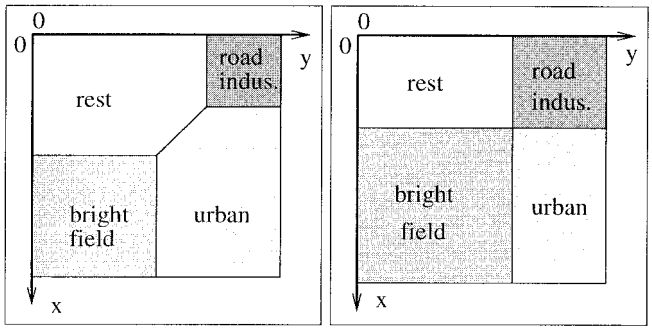
Let us consider the operators O_1 and O_2 and the classes U, I, *BF*, Ro, and R (R stands for the rest, i.e., all the other classes). Let us suppose that O_1 is able to discriminate the union $(U \cup BF)$ against the other classes, and O_2, $(Ro \cup I \cup U)$ against $(BF \cup R)$.

We can express these properties in the evidence framework by

$$\begin{aligned} m_1(U \cup BF) = x, \quad m_1(I \cup R \cup Ro) = 1 - x, \quad \text{and} \\ m_1(A) = 0 \quad \text{if } A \notin \{U \cup BF, I \cup R \cup Ro\} \\ m_2(Ro \cup I \cup U) = y, \quad m_2(BF \cup R) = 1 - y, \quad \text{and} \\ m_2(A) = 0 \quad \text{if } A \notin \{Ro \cup I \cup U, BF \cup R\} \end{aligned}$$

TABLE IV
PIGNISTIC AND BAYESIAN PROBABILITIES

	U	BF	I	Ro	R
$BetP$	xy	$x(1-y)$	$\frac{1}{2}(1-x)y$	$\frac{1}{2}(1-x)y$	$(1-x)(1-y)$
$P_{bay} \propto$	$\frac{1}{6}xy$	$\frac{1}{4}x(1-y)$	$\frac{1}{9}(1-x)y$	$\frac{1}{9}(1-x)y$	$\frac{1}{6}(1-x)(1-y)$



(a) (b)

Fig. 5. Classes maximizing the probability using the pignistic probability after (a) fusion in the evidence theory framework and (b) using the Bayesian probability.

On the other hand, in the Bayesian framework, and supposing the same probability for the simple hypotheses that an operator is unable to distinguish :

$$p_1(U) = p_1(BF) = \frac{x}{2} \quad \text{and}$$
$$p_1(I) = p_1(R) = p_1(Ro) = \frac{1-x}{3}$$
$$p_2(U) = p_2(I) = p_2(Ro) = \frac{y}{3} \quad \text{and}$$
$$p_2(Cl) = p_2(R) = \frac{1-y}{2}.$$

The following Bayesian and pignistic probabilities are then obtained after the fusion step (considering that the sources of knowledge are independent) and given in Table IV.

The classes having the maximum probability in the two frameworks are shown on Fig. 5 depending of x and y values. Results are different and as claimed before, the evidence theory favors the classes belonging to compound hypotheses (*urban* and *rest*) on which the degree of belief can be displaced. This property is not straightforward obtained in the standard Bayesian framework.

IV. CONTEXTUAL STEP IN A MARKOVIAN FRAMEWORK

The fusion of the operators in the evidence theory provides a resulting mass function which has been converted into a probability function by the pignistic transformation. Many schemes can then be adopted to take the final decision. In our case of radar image interpretation, we propose the introduction of some contextual information (which has not been used before in our fusion step) to attribute a class to each primitive. The Markovian framework through the choice of the *a priori* probability is well adapted for this purpose [16].

This step is based on a graph representation of the scene. A node is associated to each region, and an edge between two nodes corresponds to the adjacency of the two corresponding regions. The interpretation step can be considered as the labeling of the graph. We associate a random variable L_i to each node i (i.e., each region of the image) which takes its values in $\mathcal{C}l$ set of all the previously defined classes (with the reject class). Knowing some information d on each region (the values of the operators), we look for the best configuration l of the label field L in the sense of the maximum *a posteriori*, i.e., which maximizes the conditional probability of L given d, $p(L = l \mid D = d)$. Under some independence assumptions and supposing L is a Markovian field, it can be shown [16] that $p(L \mid D)$ is a Gibbs distribution of energy $U(l) = \sum_{c \in \mathcal{C}li} V_c(l)$. $\mathcal{C}li$ is the set of cliques, and the clique system is defined from the graph neighborhood structure. The energy is divided in two terms.

- The first one is the data driven term and is derived from the previous fusion step and the pignistic probability using $V_{\{i\}}(C_i) = -\ln(BetP(C_i))$.
- The second term is the contextual term and is defined using some *a priori* information on the label field. In our case, a simple Potts-like model [16] is defined to favor adjacency between "compatible" regions and penalize adjacency between "incompatible" regions as described below and using *a priori* knowledge about the landscape organization. We only take into account two site-cliques with nonzero potentials defined by

$$V_{c=\{i,j\}}(C_i, C_j) = K_e$$

where e indexes the relationship between the two classes. We distinguish three situations.

1) Adjacency is favored between two classes; this relationship is denoted by $\mathcal{A}$, and corresponds to a negative parameter $K_{\mathcal{A}}$.
2) This adjacency is neutral (denoted by $\mathcal{N}$), accepted but not favored, with the parameter $K_{\mathcal{N}} = 0$.
3) The adjacency is disfavored, denoted by $\mathcal{I}$; $K_{\mathcal{I}}$ is positive.

Let us detail the cases corresponding to these three different situations. Since the neutral adjacency is the most common case between two classes, only situations with either favored or disfavored adjacencies are reported.

- *Favored adjacency*: First of all, this is the case for any class with itself, since the different objects of the scene are either compact or thin, long structures. Adjacency between *urban* and *industrial* classes is also concerned, since these areas are often neighbors in the landscape structure.[3] Since roads or rivers often go through *urban* or *industrial* or even *bright.field* regions, these class adjacencies are favored. Besides, since roads and rivers are very thin, these classes are necessarily disfavored with a Potts-like model. To avoid their suppression, a favored adjacency has also been introduced with the rejection class.

[3] As said before, the distinction between these two classes relies mostly on their appearance in the SAR image (high density of bright points for *industrial* areas, and very high radiometry and possibly dense areas for *urban* class).

TABLE V
ADJACENCY RELATIONSHIPS BETWEEN CLASSES

Classes	Rc	Ro	Ri	H	U	I	BF	$\emptyset$
Rc	$\mathcal{A}$	$\mathcal{N}$	$\mathcal{N}$	$\mathcal{N}$	$\mathcal{I}$	$\mathcal{I}$	$\mathcal{I}$	$\mathcal{N}$
Ro	$\mathcal{N}$	$\mathcal{A}$	$\mathcal{N}$	$\mathcal{N}$	$\mathcal{A}$	$\mathcal{A}$	$\mathcal{A}$	$\mathcal{A}$
Ri	$\mathcal{N}$	$\mathcal{N}$	$\mathcal{A}$	$\mathcal{N}$	$\mathcal{A}$	$\mathcal{A}$	$\mathcal{A}$	$\mathcal{A}$
H	$\mathcal{N}$	$\mathcal{N}$	$\mathcal{N}$	$\mathcal{A}$	$\mathcal{N}$	$\mathcal{N}$	$\mathcal{N}$	$\mathcal{N}$
U	$\mathcal{I}$	$\mathcal{A}$	$\mathcal{A}$	$\mathcal{N}$	$\mathcal{A}$	$\mathcal{A}$	$\mathcal{N}$	$\mathcal{N}$
I	$\mathcal{I}$	$\mathcal{A}$	$\mathcal{A}$	$\mathcal{N}$	$\mathcal{A}$	$\mathcal{A}$	$\mathcal{I}$	$\mathcal{N}$
BF	$\mathcal{I}$	$\mathcal{A}$	$\mathcal{A}$	$\mathcal{N}$	$\mathcal{I}$	$\mathcal{A}$	$\mathcal{A}$	$\mathcal{N}$
$\emptyset$	$\mathcal{N}$	$\mathcal{A}$	$\mathcal{A}$	$\mathcal{N}$	$\mathcal{N}$	$\mathcal{N}$	$\mathcal{N}$	$\mathcal{A}$

- *Disfavored adjacency*: This parameter tries to solve the case of a bad discrimination between some classes using contextual knowledge. Therefore, it is used between *bright.field* and *urban* or *industrial* classes; *relief* and *industrial* or *urban* classes; and *bright.field* and *relief* classes. All of these situations are unusual in a natural scene, and they will mostly occur because of misclassification of these classes in the SAR images.

Table V summarizes the different relationships we used between all the classes.

The energy minimization is done by simulated annealing, but does not take too much time (7 min) since the set of regions is quite small (about 10 000 regions on a 1024×1024 image).

The parameters of the Markovian step are the following: $K_{\mathcal{A}} = -1$, $K_{\mathcal{I}} = 2$, $K_{\mathcal{N}} = 0$, the initial temperature is 5.0, and we used geometric decreasing with a multiplicative coefficient of 0.95. The parameters K_e have been empirically chosen and are fixed once and for all to process any SAR image.

The *bright.field* class, which has only been used to discriminate urban areas, has been suppressed and reclassified in the reject class on the final result (this class has indeed very little interest on its own). The *homogeneous* class is given with its associated probability for each region. In fact, only reliable regions correspond to forest or sea areas with the parameters we set, but we did not want to put too many regions in the rejection class and we preferred to present the results with this degree of confidence. Let us note that the *homogeneous* class does not distinguish between *forest* or *sea* areas. A particular class for urban class with weak probability has been added and called *suburb*.

V. RESULTS ON REAL SAR IMAGES

The general scheme we have previously described has been applied on about 20 radar images taken with different sensors (ERS-1, SIR-C/X-SAR, RADARSAT) in a fully automatic way. The set of parameters has been fixed once and for all for each sensor. Actually, the parameters are not very different from one to another, but the different interpixel spacing influences the statistical properties of the SAR images.

We present and analyze here two interpretation results on real ERS-1 images and then make a synthesis on all the tests we made to evaluate the performances and limits of the proposed scheme. The evaluation of the interpretation results is essentially made by comparing them visually with the maps corresponding to the analyzed SAR areas.

A. Analysis of Two ERS-1 Images

1) ERS-1 Image of Mantes-La-Jolie: This is an image of a flat area in the suburb of Paris with the small town of Mantes-La-Jolie, and the Seine river (Fig. 6). There are many forests and small urban areas as can be seen on the map (Fig. 7).

The interpretation result is shown in Fig. 8. The names of the detected towns and forest areas have been manually added to help the reader to compare the result with the corresponding map of Fig. 7. The Fig. 9 shows the boundaries of the detected features (except forest areas) superimposed on the SAR image.

The results are globally satisfying.

- **Urban areas**: most of the urban areas are well detected (14 of 17 existing in the map: Mantes-La-Jolie, Rosny sur S., Guernes, Limay, Mantes-La-Ville, Dennemont, Follainville, La Roche-Guyon, Vetheuil, Fresneuse, Gasny, Moisson, Sandrancourt, Claudry), although some of them (three towns: Mousseaux sur S., St-Martin-La-Garenne, Méricourt) were missed (for instance the town of Mousseaux-sur-Seine in the middle of the river curve as seen on the map in Fig. 7).
- **Road and hydrological networks**: the river Seine has been well detected, but the big highway on the bottom of the scene has also been classified as *river* instead of *road*; roads are rather incomplete, calling for an improvement of the road operator [32].
- **Forests**: they are well discriminated from the background if we take into account the confidence values: there are four areas with high confidence values which are indicated as forest areas on the map (Forêt de Moisson, Bois du Chenay, the area under the Sandrancourt town, and Forêt de Rosny at the bottom left of the map).
- **Relief**: this class is of course not important on this flat area, but detected lines correspond to difference of height near the river, with some isolated false alarms.

2) ERS-1 Image of Aix-en-Provence: This is a more difficult image since we are near a high hill with important relief [Fig. 2(a)]. The scene is centered on Aix-en-Provence, which is a large town in the South of France, with forest in the bottom right, and many industrial areas and an important road network (Fig. 7). Results are shown in Fig. 8.

The results obtained on this image are the following.

- **Urban areas**: the town and industrial areas are well detected with no false alarms in the relief areas; there are ten towns or industrial areas well detected (of a total amount of 14), two false alarms which correspond to bright reflectors but are not indicated on the map, as well as two omissions.
- **Road network**: it is rather incomplete, although the main axes are detected (highways A-8 and A-51 are detected, as well as some of the national roads but many of them were missed; let us note that they often are hardly visible on the radar image, though the results can certainly be improved).
- **Forests**: the forest Bois de Ligoures situated at the top right of the radar image is also well detected, as well as the forest areas on the Mount Montaiguet slope.

Fig. 6. Original ERS-1 image of Mantes-La-Jolie © ESA.

- **Relief**: the detected crests in the right part of the image indicate the presence of the Mount Montaiguet.

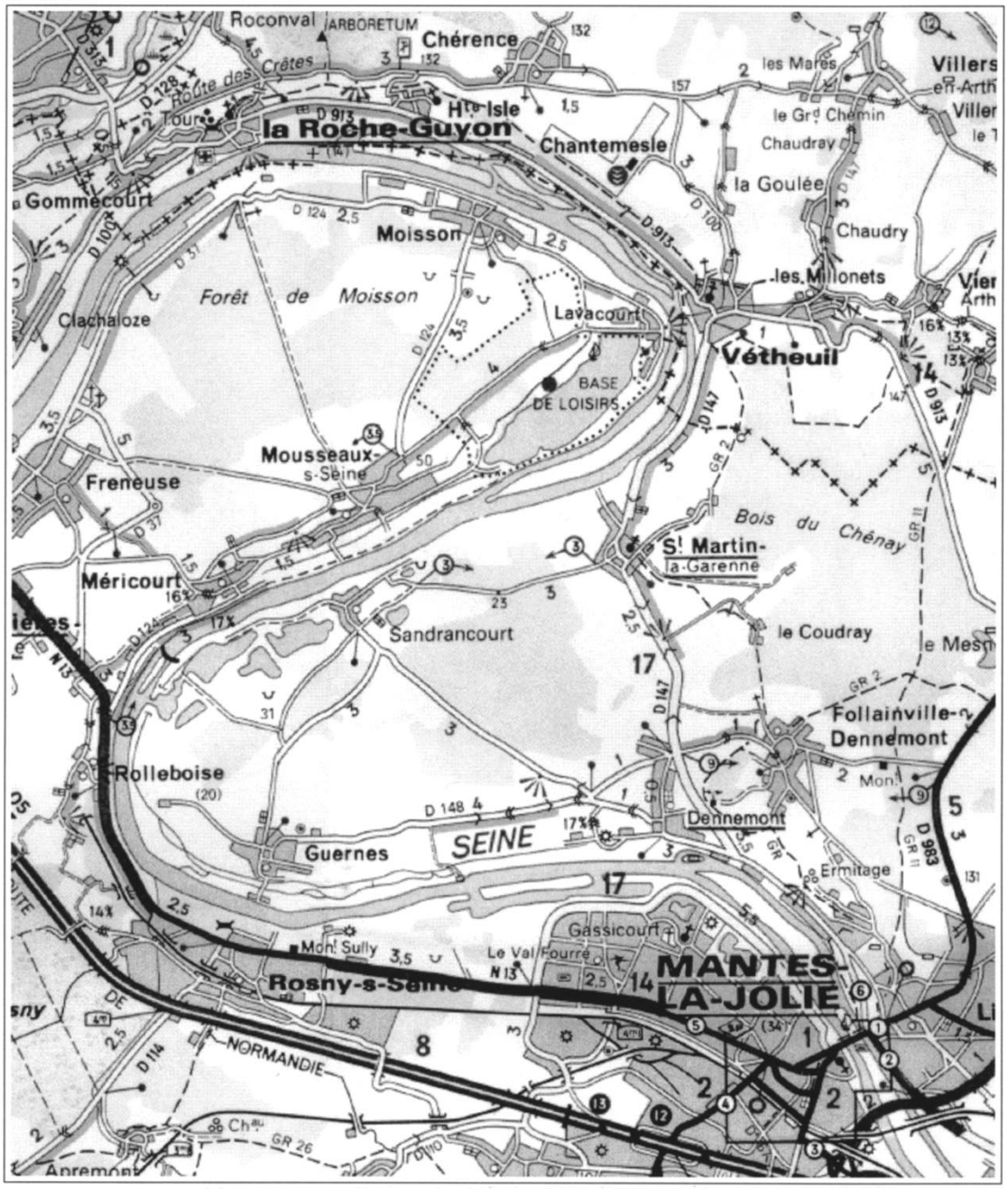
B. Global Analysis of the Results

This section describes the general results we obtained with the proposed method on a large set (about 20) of radar images and comparing them with the maps or with the radar image when no maps were available.

Several SAR images of size 1024×1024 to 2048×2048 pixels have been tested, from ERS-1, RADARSAT, and SIR-C/X-SAR radars. Four ERS-1 scenes were used in which we extracted about 13 images:

- scene of Paris which corresponds to a very large conurbation in a rather flat landscape; many urban areas (Versailles, Mantes-La-Jolie, Roissy, etc.) have been extracted and tested;
- scene of Provence (South of France) with the very large conurbation of Marseille and the town of Aix-en-Provence in a rural and mountainous landscape;
- scene of Brittany in France (Saint-Brieuc, Lorient) with small urban areas in a hilly and agricultural landscape with many forests and fields;
- scene of The Netherlands, with a typical landscape of polder (Wieringermeer, Lelystad) with small towns, seasides, many geometrical fields, and a complex network of channels;
- scene of Kourou in French Guyane, which is a tropical landscape on the seashore with many kinds of vegetation and forests.

As for RADARSAT images, we had at our disposal the scene of The Netherlands with polder landscape. Concerning SIR-



(a)

Fig. 7. Maps of the SAR images. The areas corresponding to forest areas are darker than the background but without black boundaries. (a) Map corresponding to the Mantes-La-Jolie image © Michelin.

C/X-SAR, we only had a scene near Strasbourg corresponding to small urban areas in a rather flat landscape.

All the interpretation results have been analyzed by a visual comparison with the SAR image and with a map of the corresponding area when available. For a smaller set (six ERS-1 images of different landscapes), a registration step has been performed and an exhaustive comparison with the map and the SAR image has been made. All the results given below have been deduced from this analysis.

- **Urban areas:** Urban and industrial areas are generally well detected with very few false alarms due to relief areas or bright fields thanks to the general scheme of interpretation. The introduction of the *bright.field* class and contextual information is a powerful way to avoid detecting *urban* class on very bright fields in most of the case (although it was not sufficient for the Lelystad image analysis). A better discrimination of *industrial*, *urban*, and very dense urban areas should be possible with simple radiometric criteria starting with this coarse detection. There are still some omissions and, depending on the sensor calibration, detection is more or less difficult. Table VI gives some quantitative results on a sample of ERS-1 images for which the maps were available and a precise comparison with registration has been made.

Since an oversegmentation is used to define the regions of the interpretation scheme, boundaries of the urban areas are not precisely located.

(b)

Fig. 7. (*Continued*.) Maps of the SAR images. The areas corresponding to forest areas are darker than the background but without black boundaries. (b) Map corresponding to the Aix-en-Provence image © Michelin.

• **Road and hydrological networks:** Road network detection is still disappointing because the network is often incomplete, although much computation time is dedicated to its extraction. Besides, it is difficult to give quantitative results on the road network detection. Indeed, unlike the industrial areas which are always visible, the roads are sometimes not visible at all due to particular orientation and surrounding. Therefore, the results should be compared to the detection of a human expert without the help of a map since our interpretation scheme did not use it.

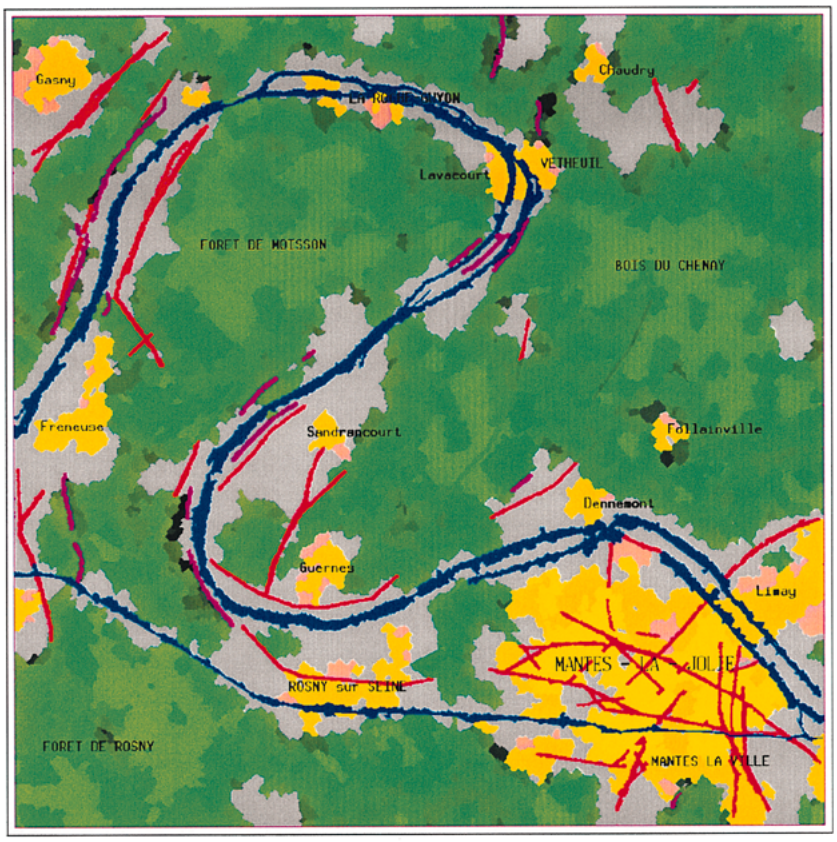
Further work could be the use of the whole interpretation result to improve the network detection using information like "there should be a road connecting two towns" or using some high-level information provided by a map (the interpretation result could then be used to register the image and the map, but relief distortions should be taken into account).

Moreover, there are very few works which aim at detecting the global shape of the network on SAR satellite images, and thus the comparison with concurrent methods is not easy.

As for the river detection, they are usually more easily seen on the SAR images and their detection is also easier; the main problem is then to distinguish between the road and hydrological networks. It is rather difficult without any new source of knowledge since they may have the same appearance on the radar image (for instance, channels on polder scenes).

• **Relief areas:** Relief areas are well detected and a new stage should permit the discrimination of isolated false alarms. The neighborhood we used in the Markovian field is, in fact, too "local" to introduce information of higher level, but a neighborhood based on a distance and not on adjacency, and relationships like parallelism should greatly improve the results. Of course, the aim of this class is to localize where the

(a)



(b)

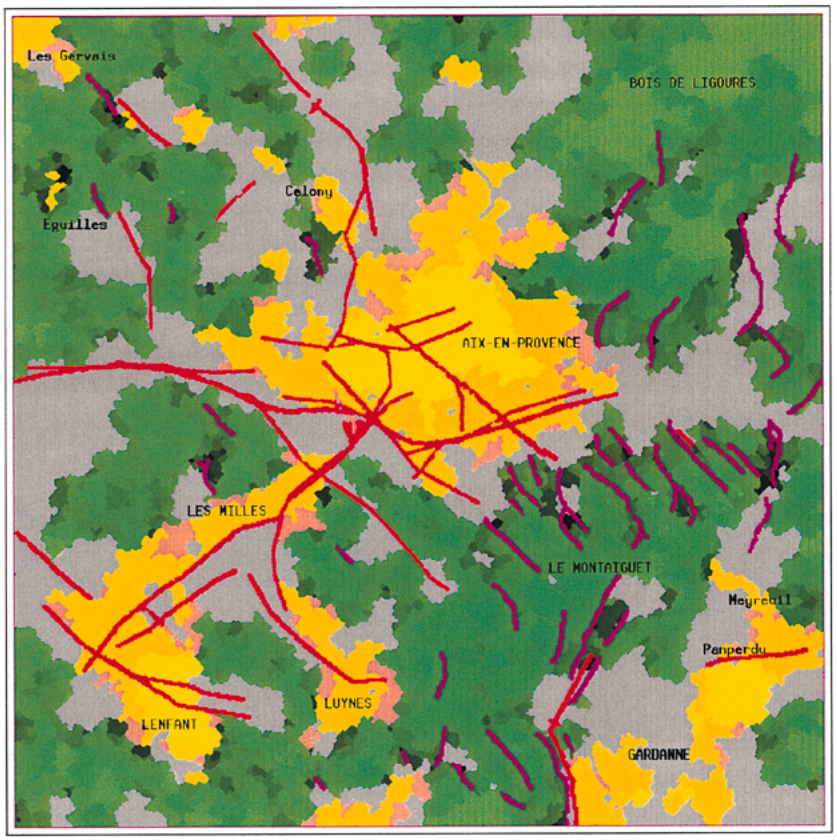


Fig. 8. Interpretation results. Caption: blue: rivers, red: roads, yellow: town (dense urban areas), orange: industrial areas, brown: suburb, pink: relief crests, green: forests. Interpretation result for (a) Mantes-La-Jolie and (b) Aix-en-Provence.

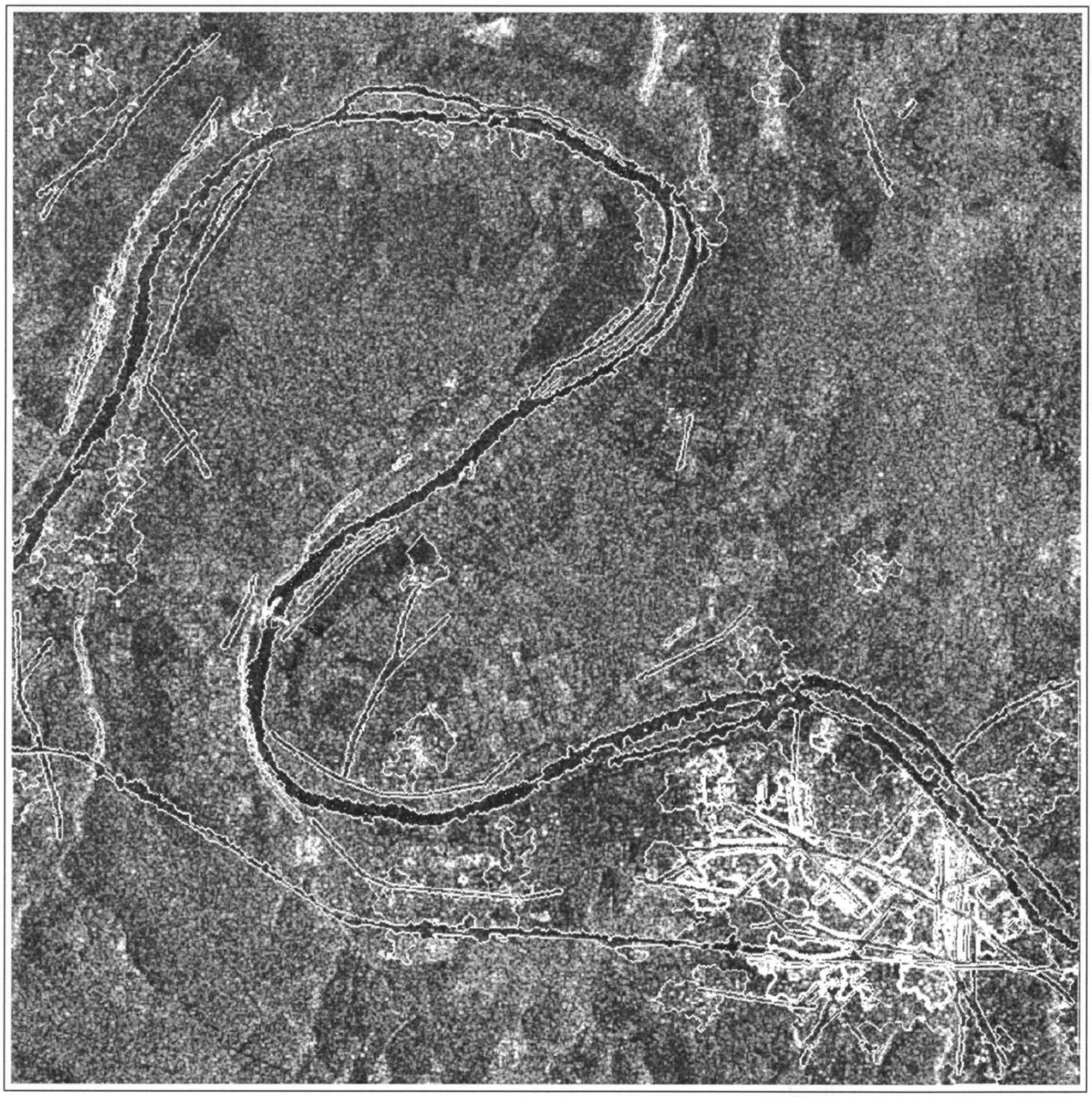


Fig. 9. Edges of the detected features (in white) superimposed on the SAR image.

TABLE VI
RESULTS OF THE URBAN CLASS DETECTION

Image	good detections	false alarms	omissions
Aix-en-Provence	10	2	2
Lorient	7	0	0
Roissy	13	0	3
Mantes-La-Jolie	14	0	3
Lelystad	4	5	0
Kourou	3	1	0

radar indicates the presence of relief, but an exhaustive detection of all the relief areas is impossible since the information we used depends on the incidence angle and relief orientation.

• **Forests:** Results for the *forest* class are usually satisfying, but they sometimes vary from one image to another, and the set of parameters we used for this class is not as stable as the other ones. Other textural operators should be introduced to improve the forest detection and localization. Furthermore, a more complete scheme should take other land-cover types into account.

To conclude about the provided results, let us make a remark here. Our aim is to automatically give a coarse map from the satellite SAR image (once the parameters have been defined for a sensor). Therefore, we are interested in the global organization and consistency of the detected features, and the method has been mostly evaluated by a visual comparison with the available maps and with the SAR images. The

accurate localization of the detected areas (urban boundaries, forest limits) is beyond the scope of this paper and should require further improvements since coarse areas provided by a segmentation are used in the interpretation scheme.

C. Practical Aspects

One of the advantages of the general scheme we proposed is its capability to take new operators into account without modification. For instance, more accurate textural discriminators could be added in a simple way, the only step to add being the fusion of the resulting mass function with the new one. Besides, the fusion with other sensor images could also be done in a simple way once the detectors are modeled (and supposing a perfect registration). Different degrees of reliability for each sensor or operator could be easily introduced using a non-null mass function for Θ, the problem being then the weighting of the reliability of each operator. New classes can also be introduced, but focal elements for each operator should be redefined.

The whole method is rather demanding in computing time, especially the road operator. It takes about 3 h for a 1024 $\times$ 1024 image on a SPARC 10 processor (this is a mean time since it depends on the information provided by the image, number of regions for the Markovian step, etc.). The first step (low-level step) takes about 2 to 2.5 h, the second step about 30 min (6 min to fuse two operators), and the third one (Markovian step) 10 min. Optimization of the code should reduce this time and faster minimization methods of the Markovian energy should also be investigated.

VI. Conclusion

This article presents an automatic method for radar image interpretation based on three main steps. The first step applies low- or intermediary-level operators taking SAR statistics into account. The second step fuses the responses of these operators in the evidence theory framework, which allows the modeling of the operator limits. The third and last step takes the final decision using contextual knowledge through a graph-based representation of the scene. The whole method gradually builds the image interpretation, first extracting raw information, then merging these information pieces in a well adapted framework, and finally introducing new *a priori* knowledge with contextual relationships between the different classes.

Once defined for a sensor, this approach is able to process a new SAR image in a fully automatic way, but many parameters have to be set beforehand. This is done using a learning step with manually detected samples and using our knowledge about the operator behavior. The general scheme is particularly modular and new operators can be easily added.

The method has been intensively tested on many radar images taken with different sensors. The results show a real progress toward SAR image automatic interpretation, although some of them could be improved and completed. They can be considered as preliminary results to prepare some higher level steps, such as refining the interpretation using high-level knowledge on landscape organization, or to perform a registration step with other sources of knowledge. Further work includes both these two approaches, possibly merging them.

References

[1] R. Azencott and C. Graffigne, "Non supervised segmentation using multi-level Markov random fields," in *ICPR*, The Hague, Netherlands, 1992, pp. 201–204.

[2] I. Bloch, "Information combination operators for data fusion: A comparative review with classification," *IEEE Trans. Syst., Man, Cybern. A*, vol. 26, pp. 52–67, Jan. 1996.

[3] ——, "Some aspects of Dempster–Shafer evidence theory for classification of multi-modality medical images taking partial volume effect into account," *Pattern Recognit. Lett.*, vol. 17, no. 8, pp. 905–919, 1996.

[4] I. Bloch, L. Aurdal, D. Bijno, and J. Müller, "Estimation of class membership functions for grey-level based image fusion," in *ICIP'97*, Santa Barbara, CA, vol. III, Oct. 1997, pp. 268–271.

[5] W. J. Christmas, J. Kittler, and M. Petrou, "Structural matching in computer vision using probabilistic relaxation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 17, pp. 749–764, Aug. 1995.

[6] V. Clement, G. Giraudon, S. Houzelle, and F. Sandakly, "Interpretation of remotely sensed images in a context of multisensor fusion using a multispecialist architecture," *IEEE Trans. Geosci. Remote Sensing*, vol. 31, pp. 779–791, July 1993.

[7] S. R. Cloude and E. Pottier, "An entropy based classification scheme for land applications of polarimetric SAR," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, no. 1, pp. 68–78, 1997.

[8] A. P. Dempster, "A generalization of Bayesian inference," *J. Royal Stat. Soc.*, vol. 30, no. B, 1968.

[9] J. Desachy, "ICARE: An expert system for automatic mapping from satellite imagery," in *Mapping and Spatial Modeling for Navigation, Volume F65 of NATO-ASI*, L. F. Pau, Ed. Berlin, Heidelberg: Springer Verlag, 1990.

[10] D. Dubois and H. Prade, *Fuzzy Sets and Systems: Theory and Applications*. New York: Academic, 1980.

[11] D. Dubois and H. Prade, *Possibility Theory*. New York: Plenum, 1988.

[12] M. A. Eshera and K.-S. Fu, "An image understanding system using attributed symbolic representation and inexact graph matching," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-8, no. 5, pp. 604–618, 1986.

[13] O. Faugeras and K. Price, "Semantic description of aerial images using stochastic labeling," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 3, pp. 633–642, Nov. 1981.

[14] A. M. Finch, R. C. Wilson, and E. R. Hancock, "Matching Delaunay graphs," *Pattern Recognit.*, vol. 30, no. 1, pp. 123–140, 1997.

[15] R. T. Frankot and R. Chellappa, "Lognormal random-field models and their applications to radar image synthesis," *IEEE Trans. Geosci. Remote Sensing*, vol. 25, pp. 195–207, Mar. 1987.

[16] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distribution, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, pp. 721–741, Nov. 1984.

[17] C. Gouinaud, F. Tupin, and H. Maître, "Potential and use of radar images for characterization and detection of urban areas," in *IGARSS'96*, Nebraska, vol. 1, May 1996, pp. 474–476.

[18] O. Hellwich, H. Mayer, and G. Winkler, "Detection of lines in synthetic aperture radar (SAR) scenes," *ISPRS, Int. Archives of Photogrammetry and Remote Sensing*, Vienna, vol. 31, pp. 312–320, 1996.

[19] F. M. Henderson and Z. G. Xia, "SAR applications in human settlement detection, population estimation and urban land use pattern analysis: A status report," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 79–85, Jan. 1997.

[20] R. Horaud and T. Skordas, "Stereo correspondence through feature grouping and maximal cliques," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-11, pp. 1168–1180, Nov. 1989.

[21] P. A. Kelly, H. Derin, and K. D. Hartt, "Adaptive segmentation of speckled images using a hierarchical random field model," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-36, pp. 1628–1641, Oct. 1988.

[22] H. Kim and P. H. Swain, "Evidential reasoning approach to multisource data classification in remote sensing," *IEEE Trans. Syst., Man, Cybern.*, vol. 25, pp. 1257–1265, Aug. 1995.

[23] I. Y. Kim and H. S. Yang, "Efficient image labeling based on Markov random field and error backpropagation network," *Pattern Recognit.*, vol. 26, no. 11, pp. 1695–1707, 1993.

[24] S. Le Hégarat-Mascle, I. Bloch, and D. Vidal-Madjar, "Application of Dempster–Shafer evidence theory to unsupervised classification in

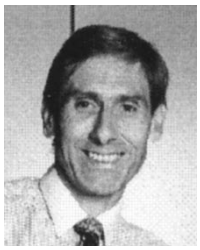
multi-source remote sensing," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 1018–1032, July 1997.
[25] T. Matsuyama, "Knowledge-based aerial image understanding systems and expert systems for image processing," in *Int. Geoscience Remote Sensing Symp.*, Zurich, 1986, pp. 1026–1038.
[26] D. M. McKeown, W. A. Harvey, and J. McDermott, "Rule-based interpretation of aerial imagery," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-7, pp. 570–585, May 1985.
[27] R. Samadani and J. F. Vesecky, "Finding curvilinear features in speckled images," *IEEE Trans. Geosci. Remote Sensing*, vol. 28, pp. 669–673, July 1990.
[28] F. Sery, D. Ducrot-Gambart, A. Lopes, R. Fjortoft, E. Cubero-Castan, and P. Marthon, "Multisource classification of SAR images with the use of segmentation, polarimetry, texture and multitemporal data," in *Proc. SPIE Europto'96*, Taormina, Italy, 1996, pp. 186–197.
[29] G. Shafer, *A Mathematical Theory of Evidence*. Princeton, NJ: Princeton Univ. Press, 1976.
[30] P. Smets, "The combination of evidence in the transferable belief model," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 12, pp. 447–458, May 1990.
[31] ———, "The transferable belief model for uncertainty representation," Tech. Rep., IRIDIA, Université Libre de Bruxelles, Belgium, 1995.
[32] 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 Sensing*, vol. 36, pp. 434–453, Mar. 1998.
[33] F. Tupin, J. F. Mangin, E. Pechersky, J. M. Nicolas, and H. Maitre, "A graph-based representation to detect linear features," *Comput. Arch. Inform. Numer. Comput.*, supp. 12, pp. 21–31, 1998.
[34] R. R. Yager, "Hedging in the combination of evidence," *J. Inform. Optim. Sci.*, vol. 4, no. 1, pp. 73–81, 1983.
[35] ———, "On the relationship of methods of aggregation evidence in expert systems," *Cybern. Syst.*, vol. 16, pp. 1–21, 1985.
[36] L. A. Zadeh, "Fuzzy sets," *Inform. Contr.*, vol. 8, pp. 338–353, 1965.
[37] ———, "Fuzzy sets as a basis for a theory of possibility," *Fuzzy Sets Syst.*, vol. 1, pp. 3–28, 1978.

Florence Tupin received the engineering degree from École Nationale Supérieure des Télécommunications (ENST), Paris, France, in 1994 and the Ph.D. degree from ENST in 1997.

She is currently working at ENST in the TSI (Image and Signal Processing) Department. Her main research interests are image analysis and interpretation, Markovian random field techniques, and SAR remote sensing.

Isabelle Bloch graduated from Ecole des Mines de Paris, Paris, France, in 1986 and received the Ph.D. degree from Ecole Nationale Supérieure des Télécommunications (ENST), Paris, in 1990 and the "Habilitation à Diriger des Recherches" from University Paris 5 in 1995.

She is a Professor in the TSI Department, ENST. Her research interests include 3-D image and object processing, 3-D and fuzzy mathematical morphology, decision theory, data fusion in image processing, fuzzy set theory, evidence theory, medical imaging, as well as aerial and satellite imaging.



Henri Maître received the engineering degree from the Ecole Centrale de Lyon, Lyon, France, in 1971 and the Docteur ès Sciences degree in physics from the University of Paris VI in 1982.

He has taught digital picture processing at École Nationale Supérieure des Télécommunications, Paris, since 1973. He is a Professor in the TSI Department. His research includes works on image analysis, image understanding, and computer vision.