Classification of very high resolution SAR images of urban areas

Aurélie Voisin, Vladimir Krylov, Gabriele Moser, Sebastiano B. Serpico, Josiane Zerubia

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Classification of very high resolution SAR images of urban areas

Aurélie Voisin — Vladimir Krylov — Gabriele Moser — Sebastiano B. Serpico — Josiane Zerubia

N° 7758

October 2011

Vision, Perception and Multimedia Understanding

ISSN 0249-6399 ISRN INRIA/RR--7758--FR+ENG

INSTITUT NATIONAL
DE RECHERCHE
EN INFORMATIQUE
ET EN AUTOMATIQUE

INRIA

centre de recherche
SOPHIA ANTIPOLIS - MÉDITERRANÉE

Classification of very high resolution SAR images of urban areas

Aurélie Voisin [*†], Vladimir Krylov [* ‡], Gabriele Moser [§¶],
Sebastiano B. Serpico [§ ‖], Josiane Zerubia [* **]

Theme : Vision, Perception and Multimedia Understanding
Perception, Cognition, Interaction
Équipe-Projet Ariana

Rapport de recherche n° 7758 — October 2011 — 35 pages

Abstract: In the framework of the assessment of environmental risks, we propose herein a new supervised Bayesian classification method. It combines statistical image modeling with a contextual approach via hierarchical Markov random fields. This research report aims to further focus on this kind of contextual classification approach by detailing both the quad-tree mathematical model and the statistics of the observations, obtained by wavelet transform. We therefore introduce modifications to a classical Markovian single-scale algorithm that lead to more accurate classification results.

Key-words: Synthetic Aperture Radar, urban areas, supervised classification, Bayesian method, finite mixture model, hierarchical Markov random fields, wavelets, copulas

[*] EPI Ariana, CR INRIA Sophia Antipolis Méditeranée, 2004, Route des Lucioles, B.P.93, F-06902, Sophia Antipolis Cedex (France)
[†] e-mail: Aurelie.Voisin@inria.fr
[‡] e-mail: Vladimir.Krylov@inria.fr
[§] Dept. of Biophysical and Electronic Engineering (DIBE), University of Genoa, Via Opera Pia 11a, I-16145, Genoa (Italy)
[¶] e-mail: gemini@unige.it
[‖] e-mail: sebastiano.serpico@unige.it
[**] e-mail: Josiane.Zerubia@inria.fr

Centre de recherche INRIA Sophia Antipolis – Méditerranée
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

Classification d'images RSO de zones urbaines à très haute résolution

Résumé : Dans le cadre d'une approche face aux risques environnementaux, nous proposons une nouvelle méthode de classification bayésienne supervisée. Celle-ci combine une modélisation statistique des images avec une prise en compte contextuelle via des champs de Markov hiérarchiques. Ce rapport de recherche vise à détailler plus amplement cette modélisation contextuelle, à savoir expliciter le modèle mathématique sur quad-arbre et l'obtention des observations par décomposition en ondelettes de l'image originale. Il met également en exergue certaines modifications apportées en vue d'améliorer la classification finale.

Mots-clés : Radar à Synthèse d'Ouverture, zones urbaines, classification supervisée, méthode Bayésienne, mélange fini, champs de Markov hiérarchiques, ondelettes, copules

Contents

1 Introduction

Synthetic aperture radar (SAR) is an active imagery system which allows day-and-night and all-weather acquisitions [1]. Indeed, the propagated wavelengths are outside the visible spectrum (X-band frequency range for the COSMO-SkyMed radar constellation). Such properties are relevant in the frame of risk management, allowing land-use and/or land-cover mapping, or the determination of areas damaged by natural disasters such as earthquakes or floodings [2]. In the framework of the assessment of environmental risks, we address herein the problem of classifying SAR images of urban areas, a specifically interesting typology given the fact that it is both strategic and critical for civil protection and damage assessment.

Several difficulties need to be considered to deal with such problems. The first one is related to speckle - backscattering of waves - which degrades the image. As a consequence, standard classification methods already validated for optical data, do not give satisfying results when applied to SAR images. A second difficulty is the heterogeneity of urban areas on very high resolution (VHR) images, leading to heterogenous statistical modeling, reflecting the different ground materials such as asphalt, concrete, metal, etc.

Lots of supervised methods have already been considered for SAR image classification, using various techniques such as, for instance, neural networks [3] [4], active contours [5], bags-of-features [6], support vector machines (SVM) [7], etc. Unsupervised methods, as well, have been studied leading to the consideration of techniques like contour tracing [8], Markov chains [9] or Markov triplets [10].

In this report, the proposed supervised Bayesian classification is statistical and can be divided into two steps. The first step deals with the SAR amplitude statistical modeling for each class considered in the classification (e.g. vegetation, urban...) by using a finite mixture model (FMM), estimated thanks to a dictionary-based stochastic expectation maximization (DSEM) algorithm [11] [12]. The SAR amplitude probability density functions (PDFs) are considered as mixtures of K PDFs automatically chosen among a predefined dictionary of SAR-specific distribution families. Such mixtures aim to model the previously mentioned heterogeneities. We take into account an additional knowledge by extracting a textural feature from the original SAR image in order to optimize the detection of urban land covers. The marginal PDFs of both the SAR image and the textural feature are combined via copulas, leading to a joint PDF for each class. The second step aims to generate the classification map, using the joint statistics as a learning information. To improve the robustness with respect to speckle noise, we considered a contextual model based on Markov random fields (MRF) [13], and more precisely a hierarchical MRF [14] [15], which also offers the possibilty to deal with multi-resolution imagery.

The first step has already been detailed in [11] [12] [16] and [17], thus we briefly sum up the method in Sec. 4.1. The main focus of this research report is the considered Bayesian classification, which uses hierarchical MRFs. First, we detail the considered hierarchical method by giving some relevant properties in Sec. 2. Then, we describe the observations (Sec. 3) and the mathematical model (Sec. 4) used in the quad-tree and we show preliminary results (Sec. 5). After discussing these results, we try to improve them by introducing further modifications into the developed algorithm in Sec. 6. Finally, in Sec. 7, we show classification results obtained on COSMO-SkyMed (©ASI) and TerraSAR-X (©DLR) acquisitions.

2 The hierarchical model

Hierarchical models can be divided into 2 distinct groups [14], both of which can be considered to address our classification problem:

- *Induced* hierarchical-based models, including the renormalization group [18] and the constrained configuration subspaces [19] techniques.
- *Explicit* hierarchical-based models [20], including the hierarchical graph-based model that we employ in this report.

The hierarchical approaches can be viewed along three complementary levels [14], which will be further detailed in the next sections of this research report:

- A hierarchical decomposition of data (Sec. 3.3) via wavelet tranforms (Sec. 3.2) in our case.
- A stochastic multiscale image modeling (Sec. 4.1).
- A hierarchical algorithm definition: in our case, we will use a hierarchical graph technique (quad-tree) (Sec. 4).

The first part of this research report aims to give general definitions and to make a general presentation of the considered explicit hierarchy (quad-tree).

2.1 Notations and definitions

We define here the notations that are used throughout this report. s defines a site and belongs to a finite set S. V_s describes its neighborhood and has the following properties:

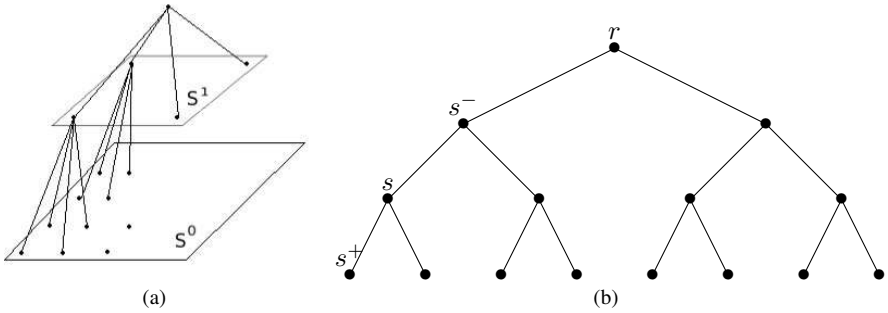
(i) $\forall s \in S, s \notin V_s$,

(ii) $\forall s_1, s_2 \in S, s_1 \in V_{s_2} \Leftrightarrow s_2 \in V_{s_1}$.

V is the set of neighborhoods. The couple (S, V) defines a graph G, which is both *simple* and *nondirected* thanks to the symmetry property (ii) of the neighborhood. A graph is defined *connected* when two random sites s and t, located in the graph, can be linked by a finite chain.

A *clique* is a non-empty subset c of neighboring sites of size equal to or higher than 1, such that two distinct sites of c are neighbor sites. C denotes the set of cliques of the graph G.

In the following, we shall focus on a specific graph, a *tree*, which is a connected graph without cycles. The set of sites is then hierarchically partitionned on $S = S^0 \bigcup S^1 \bigcup ... \bigcup S^R$ where R corresponds to the coarsest resolution, the *root*, and 0 corresponds to our reference level (finest resolution). In this tree structure, a parent-child relation can be defined: for each site s of any tree-level n a unique *parent* s^- and several *children* s^+ can be defined. $d(s)$ refers to the set including s and its descendants. For the specific case where s owns four children, the tree structure is named *quad-tree*. Such a structure is depicted in Fig. 1. It is straighforward that the sites at the root level S^R do not own any parent and that the sites at the reference level do not own any children ($d(s) = s$ for $s \in S^0$). The relationship between the different sites of the quad-tree is based on a parent-child link.



(a) (b)

Figure 1: (a): Hierarchical model structure: quad-tree; (b): Quad-tree notations.

2.2 Hierarchical graphs

This technique results from both graph theory and signal processing. Two well-known related graphs are the 3D-pyramidal graph [21] and the quad-tree [20], and the main difference lies in the respective existence or non-existence of neighbors in a given level n. Some pyramidal graphs may have a higher number of branches, due to the fact that more ascendants are considered in the hierarchy. In such cases, the neighborhood is better taken into account, but the graph obtained is not a tree anymore.

We want to estimate a set of hidden labels X given a set of observations Y attached to the sites. X and Y are random processes. The restriction of X (resp. Y) to the level n is $X^n = \{X_s, s \in S^n\}$ (resp. $Y^n = \{Y_s, s \in S^n\}$) where the realization x^n takes its values in Ω. Some extra hypotheses are needed to ensure that X is a *Markov random field* on the graph G: $\forall s \in S, \forall x \in \Omega$,

(i) $p(X = x) > 0$,

(ii) $p(X_s = x_s | X_t = x_t, t \in S - \{s\}) = p(X_s = x_s | X_t = x_t, t \in V_s)$.

The Markov property implies that the conditional local distribution of X_s is determined by its neighboorhood X_t, where $t \in V_s$.

The in-scale prior, further detailed in Sec. 4.2, is expressed as $p(x^n|x^{n+1}) = \prod_{s \in S^n} p(x_s|x_{s^-})$.

One last hypothesis is the pointwise dependance of Y with respect to X, thus implying that each couple (X, Y) is Markovian on the quad-tree:

$$p(y|x) = \prod_{n=0}^{R} p(y^n|x^n) = \prod_{n=0}^{R} \prod_{s \in S^n} p(y_s|x_s).$$

2.3 Advantages of this model

In this section, we list some advantages related to the use of hierarchical Markov random fields as compared to single scale models:

- Scale causality generated by the quad-tree structure, which allows to use of a non-iterative algorithm on the quad-tree, thus implying a computational time decrease.

- Possibility to model different kinds of statistics, and thus to use different kinds of images (different resolutions, different sensors) as observations [22].
- Impoved robustness to speckle noise.
- Weak sensitivity to initial conditions.
- Smaller probability to find a local minimum, hence more likely to converge to a global solution (multigrid properties).

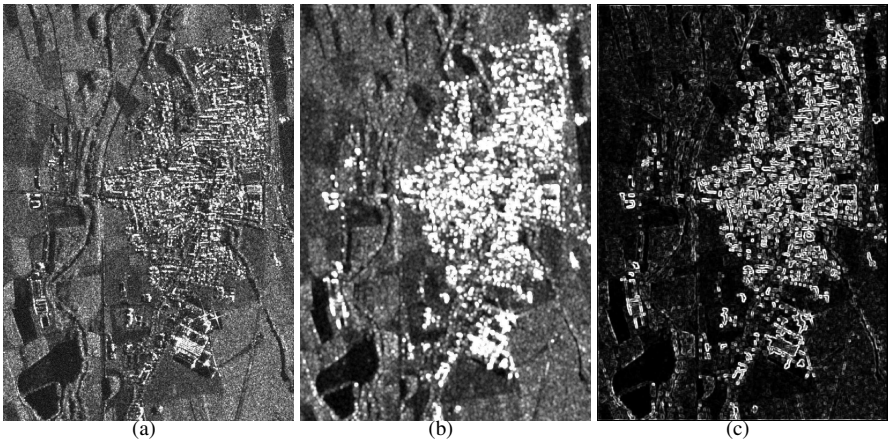
3 The observations

A difficulty in our work is that we only consider single-polarization (single-pol) SAR images. To deal with this problem, we extract some extra-information: a multi-scale decomposition via wavelets combined to a textural feature extraction at each scale.

In this section, the textural feature extraction and the wavelet decomposition are illustrated on the single-pol COSMO-SkyMed image (©ASI, 2008) acquired over the city of Cavallermaggiore in Italy (HH polarization, StripMap acquisition mode, geocoded, 950×700 pixels).

3.1 Textural features

We extract from the original image a textural feature, which allows to take into account an additional information. We employ some specific features to discriminate urban areas (cf. Fig. 2), thus allowing a better urban classification performance.



(a) (b) (c)

Figure 2: (a): Original SAR image of Cavallermaggiore (Italy) (COSMO-SkyMed, ©ASI); (b): Textural feature extracted by semivariogram approach; (c) : Textural feature extracted by GLCM approach. We notice that the urban areas are well discriminated.

Two well-known kinds of textural features have been considered, and are detailed in the next subsections:

- The greylevel co-occurrence matrices (GLCM) [23], which allow the second order image statistics estimation.
- The semivariograms [24] [25], which are related to the autocovariance function of the spatial random process modeling the input SAR image.

Greylevel co-occurrence matrix

Suggested by Haralick in 1973 [23], the GLCM is a commonly used tool. The greylevel co-occurrence matrix $Glcm$ is a square $Z \times Z$ matrix, Z being the number of greylevels in the image, which describes the joint statistics of the greylevels of different pixels as a function of their reciprocal locations. Typically, the element (i, j) of the matrix is the probability $p_{Glcm}(i, j)$ that a pixel with value i is adjacent to a pixel with value j. The adjacency can be defined in each of the possible directions (horizontal, vertical, left and right diagonals). In our case, we considered horizontal adjacency, with an offset equal to 1, i.e the matrix is filled by considering a reference pixel, and the pixel located to its right. Among various textures that can be extracted from the GLCM matrix, we use the variance, which usually in our case discriminates urban areas [16].

Semivariogram

The semivariogram describes the spatial properties of an image, in particular the degree of spatial dependence of a spatial random field assessed in terms of second-order statistics. If s and t are two adjacent pixels separated by a distance-offset h, the semivariogram is defined as the expected square increment of the values between s and t:

$$\gamma(s,t) = \frac{E[|z_s - z_t|^2]}{2}. \tag{1}$$

Empirically, the semivariogram is expressed as:

$$\hat{\gamma}_i(h) = \frac{\sum_{j:(i,j)\in N(h)} |z_i - z_j|^2}{N(h)}, \tag{2}$$

where i, j are adjacent pixels separated by an offset h, $N(h)$ denotes the set of pairs of observations and z_i, z_j their corresponding greylevels.

Textural feature approach

Both approaches to texture extraction are applied on a moving-window basis of size $w \times w$, meaning that each pixel of the image successively becomes a reference pixel, and its value is replaced by the variance or the semivariogram value estimated within the window (Fig. 3(a)). To estimate these parameters at the border pixels, we applied a pixel duplication so as to generate a textural feature image of the same size as the original image (Fig. 3). An alternative would have been zero-padding. According to preliminary experiments, the choice of a window of size $w = 5$ provided more accurate results on the considered data sets than other window sizes in the range $[3; 17]$.

In Sec. 4.1, we further explain how the statistics of this feature are taken into account in the mathematical model as an additional information. In our experiments, we only considered the GLCM features, motivated by the experimental textural feature extraction study performed in [16].

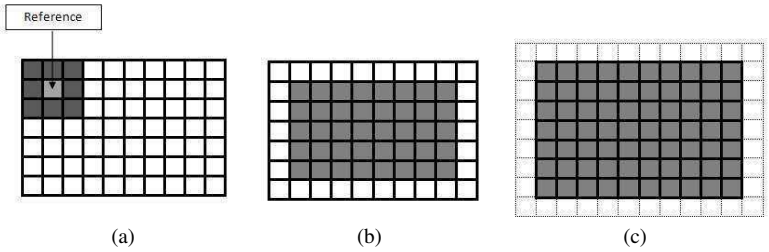


Figure 3: (a): Window (size $w = 3$) used to estimate the variance at the reference pixel; (b) and (c): In grey: pixels for which the variance is calculated ($w = 3$) respectively without any pixel duplication at the borders (b) and with duplication (c).

3.2 Wavelet decomposition

A natural image processing model to generate a hierarchical decomposition is the wavelet transform [26]. In this section, we will only recall some basic properties of the 2-D discrete wavelet transform that we use to decompose original SAR images (Fig. 4).

Figure 4: SAR image (©ASI) wavelet decomposition example using Daubechies wavelets.

The discrete 2-D wavelet transform is the decomposition of an image, mathematically modeled by the set $\{y(s,t)\}$ – where (s,t) are the coordinates – on an orthogonal basis $L^2(\mathbb{R}^2)$ generated by wavelet functions $\{\Psi^{LH}, \Psi^{HL}, \Psi^{HH}\}$ and by scaling functions Φ^{LL}. The considered space includes the translation and the dilatation with parameters $\tau = \{\tau_n\}$, $n \in \mathbb{Z}$ and $a = \{a_m\}$, $m \in \mathbb{Z}$. From now, we consider the special case where $\tau_n = 2^m n$ and $a_m = 2^m$.

We build a 2-D wavelet basis, for which the translation is linked to two parameters (horizontal and vertical descriptors, refered here as n and q). Thus, we can express the wavelet function basis as $\{\Psi^B_{n,q,m}(s,t)\} = \{2^{-m/2}\Psi^B(2^{-m}s - n, 2^{-m}t - q)\}$, where $B = \{LH, HL, HH\}$ and the scaling function basis as $\{\Phi^{LL}_{n,q,m}(s,t)\} = \{2^{-m/2}\Phi^{LL}(2^{-m}s - n, 2^{-m}t - q)\}$.

For a given scale J, the decomposition of the image for every couple (s,t) is :

$$y(s,t) = \sum_{n,q} a_{n,q,J} \Phi^{LL}_{n,q,J}(s,t) + \sum_{B} \sum_{j=1}^{J} \sum_{n,q} d_{n,q,j} \Psi^{B}_{n,q,j}(s,t), \tag{3}$$

where the coefficients $a_{n,q,J}$ and $d_{n,q,j}$ are respectively the approximation and the detail coefficients.

The approximation coefficient at the scale 0 corresponds to the original image we want to decompose. By filtering and decimating this image, we obtain for the scale 1:

- The approximation coefficients by applying a low-pass filter on the rows and the colums, which we refer as LL (Low-Low).
- The detail coefficients by low-pass filter on the rows and high-pass filter on the columns (LH), high-pass filter on the rows and low-pass filter on the columns (HL) or high-pass filter (HH) on both rows and columns.

In a same way, the approximation coefficients at scale j are decomposed by filtering and decimation so as to obtain the coefficients at scale $j+1$. A well-known scheme [27] to model this decomposition is given in Fig. 5.

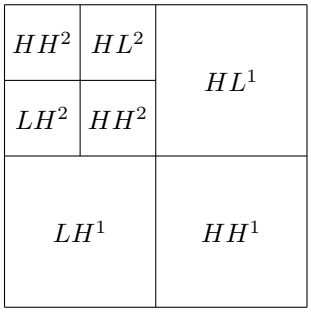


Figure 5: Wavelet transform for a scale 2.

For each scale, we only consider the approximation coefficients. The scale factor is in power of 2, thus leading to a quad-tree configuration in which one pixel at scale $j \neq 0$ owns one single parent and four children.

A number of wavelets functions exists, and we tried to decompose our original image using different wavelet families such as Daubechies [28], orthogonal, biorthogonal, and some others [27] [29]. We then applied our classification algorithm and compared the output results visually and by considering the overall accuracy. Visually, the results were very similar, and the main difference relied in the level of smoothness of the final classification map. In Tab. 6, we notice that the highest accuracy is reached by using Daubechies 10 wavelet. For this reason in the following we employ this type of wavelets.

3.3 Observations on the quad-tree

As mentioned above, the observations at each quad-tree scale n are generated by a wavelet decomposition of the original SAR image, leading to the consideration of a *multi-scale* context.

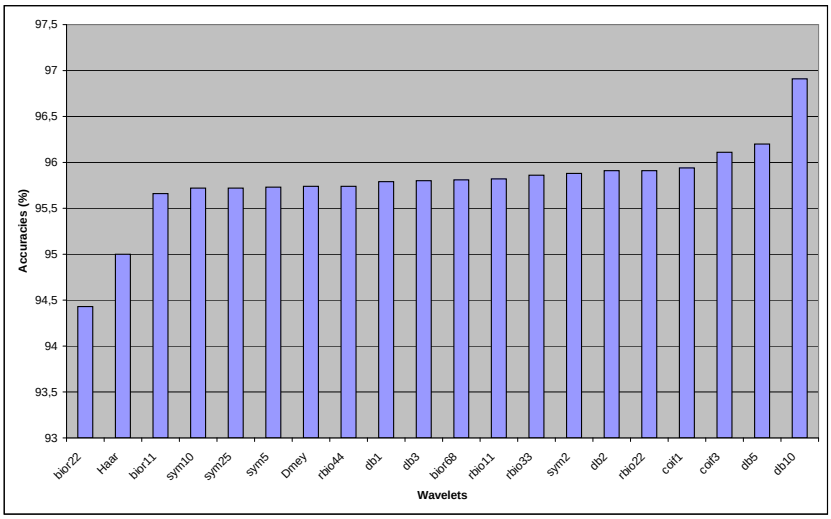


Figure 6: Overall classification accuracies obtained for different wavelet families obtained for the Cavallermaggiore image.

Another option would have been a *multi-resolution* approach, for which the observations are real data acquired at different resolution, not necessarily by the same sensors.

4 Mathematical model on the quad-tree

A variety of algorithms were proposed to estimate the labels on hierarchical graphs. Typically, a global energy minimization is done via iterative relaxation algorithms [30]. The consideration of a quad-tree allows to benefit from its good properties (e.g. causality) and to apply non iterative algorithms [31]. Among the different algorithms employed in the literature, a first option is to estimate exactly the Maximum A Posteriori (MAP) by using a linear Kalman filter [32] or a non-linear Viterbi algorithm [22] [33] [34]. However, such a criterion is known to generate underflow by the consideration of very small probabilities. For this reason, we take into account an exact estimator of the *marginal posterior mode* (MPM) [35] [20]. The cost function associated to this estimator offers the possibility to penalize the errors according to their number and the scale at which they occur: an error at the coarsest scale is more strongly penalized than an error at the finest scale, which is a desired property because a site located at the root corresponds to 4^R pixels at the finest scale.

To estimate the posterior probability, we need the following prior information: the likelihood, the prior probability and the transition probability at each site s of the quad-

tree. This section aims to present how this prior information is estimated. Secondly, we focus on the maximization of the posterior probability.

4.1 Likelihood term

The likelihood is the probability $p(z|\omega_m)$ that a greylevel pixel value is z given its label $x = \omega_m$. We want to model the statistical distributions of each class m considered for the classification, $m \in [1; M]$, given a training set. For each class, the PDF $p_m(z|\omega_m)$ is modeled via finite mixtures [36] of independent greylevel distributions:

$$p_m(z|\omega_m) = \sum_{i=1}^{K} P_{mi} p_{mi}(z|\theta_{mi}), \tag{4}$$

where z is a greylevel, $z \in [0; Z-1]$, and ω_m is the m^{th} class. P_{mi} are the mixing proportions such that for a given m, $\sum_{i=1}^{K} P_{mi} = 1$ with $0 \leq P_{mi} \leq 1$. θ_{mi} is the set of parameters of the i^{th} PDF mixture component of the m^{th} class. The use of finite mixtures instead of single PDFs offers the possibility to consider heterogeneous PDFs, usually reflecting the additive contributions of the different materials present in each class (for instance, different kinds of crops for the vegetation class). Such class heterogeneity is relevant since we deal with VHR images. Moreover, the use of finite mixtures can be seen as a generalization of the determination of a single PDF, and allows to estimate both the best finite mixture model and/or the best single PDF model.

The PDFs $p_{mi}(z|\theta_{mi})$ are automatically chosen in a predefined dictionary including the four following distributions: Log-Normal, Weibull, Nakagami and Generalized Gamma. The DSEM algorithm is used to estimate the best-fitting mixture model for each considered class. It combines a density parameter estimation via the method of Log-cumulants [37] and a stochastic expectation maximization (SEM) algorithm [38]. For more details concerning the proportion estimation, the selection of the best family and the estimation of its parameters for each component of the mixture and also the selection of the parameter K, see [11], [12], [16] and [17].

To improve the classification results, we model a joint PDF $p(y|\omega_m)$, where y contains observations of both original image and its corresponding textural feature. We start by extracting a textural feature from the original SAR image (Sec. 3.1) and then we apply the DSEM algorithm to each class and each input image (original and feature). This is possible thanks to the flexibility of the DSEM algorithm, granteed by its quasi-non parametric formulation. For each class, the marginal PDFs are combined into a joint PDF via copulas [39]. This joint PDF is plugged into the mathematical model of the quad-tree.

Copulas

The aim of using copulas is to compensate the fact that, contrary to marginal PDFs modeling, few models were proposed for modeling joint probabilities of SAR images (e.g. in dual-pol case) and to our best knowledge, no model is available for the joint statistics of SAR amplitudes and textures.

2-D copula is a bivariate joint distribution defined on $[0, 1]^2$ such that marginal distributions are uniform on $[0, 1]$. Specifically, a *bivariate copula* is a function $C : [0, 1]^2 \mapsto [0, 1]$, which satisfies the following properties:

(i) both marginals are uniformly distributed on $[0, 1]$;

(ii) for every u,v in $[0, 1]$: $C(u, 0) = C(0, v) = 0$, and $C(u, 1) = u$, $C(1, v) = v$;

(iii) for every $u_1 \leq u_2$, $v_1 \leq v_2$ in $[0, 1]$: $C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) + C(u_1, v_1) \geq 0$.

The importance of copulas in statistics is explained by Sklar's theorem [39], which states the existence of a copula C, that models the joint distribution function H of arbitrary random variables Z_1 and Z_2 with Cumulative Distribution Functions (CDFs) F and G:

$$H(z_1, z_2) = C(F(z_1), G(z_2)), \tag{5}$$

for all z_1, z_2 in $\mathbb{R}$. If F and G are continuous, then C is unique.

Taking the derivative in (5) over the two continuous random variables z_1 and z_2 with PDFs f and g, we obtain the joint PDF distribution:

$$h(z_1, z_2) = f(z_1)g(z_2)\frac{\partial^2 C}{\partial z_1 \partial z_2}(F(z_1), G(z_2)). \tag{6}$$

We wish to determine this specific joint PDF h given the marginal distributions f and g, corresponding to the marginal PDFs of both SAR image and its textural feature estimated by DSEM. The cumulative distribution functions F and G are easily estimated thanks to the knowledge of the parameters of f and g (Eq.(4)):

$$F_m(z|\omega_m) = \sum_{i=1}^{K} P_{mi} F_{mi}(z|\theta_{mi}). \tag{7}$$

Indeed, by definition, the CDF $F_{mi}(z|\theta_{mi})$ is the integral on $]-\infty; z]$ of its corresponding PDF $p_{mi}(z|\theta_{mi})$. According to Eq. (6), to determine h, we only have to determine the copula family C and its parameter α.

To find the best fitting copula C, we consider a dictionary of 5 copulas: Clayton, Ali-Mickhail-Haq, Franck, Marchal-Olkin and Farlie-Gumbel-Morgenstern. The mathematical expressions of such copulas involve only one parameter α. This choice of copulas is able to model a considerable variety of dependence structures [40]. To estimate α and the best fitting copula, we use the relation between α and Kendall's τ which is a ranking correlation coefficient [39]. By definition, Kendall's τ is a concordance-discordance measure between two independent realizations (Z_1, Z_2) and $(\hat{Z}_1, \hat{Z}_2)$: $\tau = P\{(Z_1 - \hat{Z}_1)(Z_2 - \hat{Z}_2) > 0\} - P\{(Z_1 - \hat{Z}_1)(Z_2 - \hat{Z}_2) < 0\}$. Given two realizations $z_{1,l}$ and $z_{2,l}$ ($l \in [1; N]$), the empirical estimator $\hat{\tau}$ is:

$$\hat{\tau} = \frac{\sum_{l=1}^{N-1} \sum_{k=l+1}^{N} z_{1,lk}\, z_{2,lk}}{\binom{N}{2}}, \qquad \text{where } z_{n,lk} = \begin{cases} 1, & \text{if } z_{n,l} \leq z_{n,k} \\ -1, & \text{otherwise} \end{cases}, \quad \text{for } n = 1, 2. \tag{8}$$

The general connection between Kendall τ and the copula C associated with $H(z_1, z_2)$ is obtained by integrating the definition of τ over the distribution of $(\hat{Z}_1, \hat{Z}_2)$ [39]:

$$\tau = 4 \int_0^1 \int_0^1 C(u, v) dC(u, v) - 1. \tag{9}$$

Once the estimate $\hat{\tau}$ (8) is computed, we plug it in place of τ in (9), so as to get parameter estimates $\hat{\alpha}$ for each of the copulas in our dictionary. Then, for each class m, we choose the best fitting copula according to the highest p-value in Pearson Chi-square test-of-fitness (PCS) [41]. The null hypothesis in PCS is that the sample frequencies $C_c(F_{1m}(u_1), F_{2m}(u_2))$ ($m = 1, \ldots, M$, $c = 1, \ldots, 5$), where (u_1, u_2) are the observed data, and F_{1m}, F_{2m} are the distribution estimates corresponding to the marginal DSEM density estimates, are consistent with the theoretical frequencies for the copula c: $C_c(v_1, v_2)$.

More details about this CoDSEM algorithm (copulas and DSEM) can be found in [12] and [17].

4.2 Transition probabilities

The fundamental hypothesis in the application of the described hierarchical MRF is to consider the random process X Markovian on scale, i.e. $p(x^n|x^k, k > n) = p(x^n|x^{n+1})$, where n and k are scales. These transition probabilities between the scales, $p(x_s|x_{s^-})$, determine the hierarchical MRF because they represent the causality of the statistical interactions between the different levels of the tree. Thus, they need to be well defined. We decided to use the transition probability introduced by Bouman et al. [33]: for all sites $s \in S$ and all scale $n \in [0; R-1]$,

$$p(x_s = \omega_m | x_{s^-} = \omega_k) = \begin{cases} \theta_n, & \text{if } \omega_m = \omega_k \\ \frac{1-\theta_n}{M-1}, & \text{otherwise} \end{cases}, \tag{10}$$

where ω_m and ω_k represent respectively the classes m and k, $m, k \in [1; M]$ and M represents the number of considered classes for the final classification (chosen by the user). This model favours an identical parent-child labeling. Typically, we choose $\theta_n \approx 0,8$, meaning that a site s at scale n has a probability of about 80% to belong to the same class as its ascendent s^-.

4.3 Prior probability

The prior probability is the probability that the label of the site s is x_s, $x_s \in [1; M]$. The prior distribution at a level n in $[0; R-1]$ is given by:

$$p(x_s^n) = \sum_{x_{s^-}^n} p(x_s^n | x_{s^-}^n) p(x_{s^-}^n). \tag{11}$$

Thus, the prior information at the coarsest level R allows to determine the prior information at the other levels because the transition probabilities are known (Sec. 4.2).

A preliminary step is then to determine the prior information at the coarsest level. The hypothesis of equiprobability between classes often does not allow an accurate modeling. We suggest two alternatives which lead to similar results and which are based on the same principle: make a preliminary classification at the coarsest level R so as to be able to estimate the prior $p(x_s)$ for all the sites $s \in S^R$. The other priors are estimated by using (11).

The first alternative is to apply a K-nearest neighbors algorithm [42] at level R. A second alternative is to run the MPM criterion algorithm twice, first by considering an equiprobability of the priors, and then to update the priors with the preliminary results obtained for this classification.

To estimate the priors given a classification map, we use a Markovian model which takes into account the contextual information at each level, and thus leads to a better prior estimation. By the Hammersley-Clifford theorem [43], we can define a local characteristic for each site:

$$p(x_s) = \frac{1}{Z}\exp(-\beta \sum_{s:\{s,t\}\in C} \delta_{x_s=x_t}) \quad \text{with } \delta_{x_s=x_t} = \begin{cases} 1, & \text{if } x_s = x_t \\ 0, & \text{otherwise} \end{cases}. \tag{12}$$

where s is here a site at the root ($s \in S^R$), x_s its label and x_t such that t et s are in the same clique. We considered a second-order neighborhood, meaning that only cliques C of size 2 are taken into account.

4.4 Posterior probabilities and their estimation

Since the tree has no cycle, the labels are estimated exactly and not iteratively by MPM thanks to a forward-backward algorithm, similar to the classical Baum algorithm for Markov chains [44]. The aim is to maximize the posterior marginal at each site s:

$$\hat{x_s} = \underset{x_s}{\operatorname{argmax}}\, p(x_s|y). \tag{13}$$

.

The first step of this procedure is to estimate the posterior probability at the level 0 given the observations at each level $y = \{y_s, \forall s \in S, \forall n \in [0; R]\}$. This estimation is done in 2 passes, refered to as bottom-up ("forward") and top-down ("backward") passes.

Bottom-up pass

This pass aims to estimate for each site $s \in S$ the partial posterior marginals $p(x_s|y_{d(s)})$ and $p(x_s, x_{s^-}|y_{d(s)})$ that are needed for the complete posterior probabilities $p(x_s|y)$ estimation (top-down pass). By definition,

$$p(x_s, x_{s^-}|y_{d(s)}) = p(x_{s^-}|x_s)p(x_s|y_{d(s)}), \tag{14}$$

with

$$p(x_{s^-}|x_s) = \frac{p(x_s|x_{s^-})p(x_{s^-})}{p(x_s)}. \tag{15}$$

The prior and the transition probabilities for every site s are preliminary determined (see Secs. 4.2 and 4.3). First, we estimate $p(x_s|y_{d(s)})$ so as to determine $p(x_s, x_{s^-}|y_{d(s)})$ according to (14). Laferté et al. [20] showed that

$$p(x_s|y_{d(s)}) = \frac{1}{Z}p(y_s|x_s)p(x_s) \prod_{t\in s^+} \sum_{x_t} \left[\frac{p(x_t|y_{d(t)})}{p(x_t)} p(x_t|x_s)\right]. \tag{16}$$

Thus, we are going to proceed to a recursion, estimating first $p(x_s|y_{d(s)})$ then $p(x_s, x_{s^-}|y_{d(s)})$ for each level, starting from the leaves and proceeding until the root is reached. The probabilities $p(x_s|y_{d(s)})$ at a given level n are used to estimate $p(x_s|y_{d(s)})$ at the level $n+1$. The likelihoods $p(y_s|x_s)$ are preliminary computed using CoDSEM (Sec. 4.1).

The sites at the level 0 do not have any descendence, hence

$$p(x_s|y_{d(s)}) = p(x_s|y_s) = \frac{1}{Z}p(y_s|x_s)p(x_s)$$

and

$$p(x_s|y_s) = \frac{p(x_s|x_{s^-})p(x_s|y_s)p(x_{s^-})}{p(x_{s^-})}.$$

Top-down pass

For each level n, from the coarsest to the finest, we are going to estimate the complete posterior probabilities according to the partial posterior probabilities determined in the bottom-up pass. For a site $s \in \{S^0, ..., S^R - 1\}$, according to Bayes theorem,

$$\begin{aligned} p(x_s|y) &= \sum_{x_{s^-}} p(x_s|x_{s^-}, y)p(x_{s^-}|y) \\ &= \sum_{x_{s^-}} p(x_s|x_{s^-}, y_{d(s)})p(x_{s^-}|y) \\ &= \sum_{x_{s^-}} \frac{p(x_s, x_{s^-}|y_{d(s)})}{\sum\limits_{x_s} p(x_s, x_{s^-}|y_{d(s)})} p(x_{s^-}|y). \end{aligned} \tag{17}$$

At the coarsest level R, $p(x_s|y) = p(x_s|y_{d(s)})$. Hence if we want the classification map at this level, we can directly estimate it thanks to the relation $\hat{x_s} = \operatorname{argmax} p(x_s|y)$ for $s \in S^R$. Otherwise, at other levels, we plug the previous estimation of $p(x_s, x_{s^-}|y_{d(s)})$ into Eq. (17). The labels are determined by $\hat{x_s} = \underset{x_s}{\operatorname{argmax}}\, p(x_s|y)$ at each level.

The $p(x_s|y)$ maximization is done thanks to a modified Metropolis Dynamics algorithm (MMD) [45]. To apply this algorithm, we do not proceed to the maximization of $p(x_s|y)$, but we minimise the negative of its logarithm, which is possible because the logarithm is an increasing function. MMD algorithm has good properties for both its low computation time and the good precision of its results. It is a good compromise between the deterministic algorithm ICM (iterative conditional mode) [46] which is fast but can fall into a local minimum and the simulated annealing (SA) which in theory finds a global minimum but is often slow [47].

4.5 MPM estimation on the quad-tree

We sum up here the different steps of the algorithm, leading to the classification map:
Preliminary step

- Apply CoDSEM to each scale.
- Define the in-scale transition probabilities.
- Determine the prior probability at the root. The priors at the other scales are determined by (11).

Bottom-up pass

- *Initialization at the leaves of the tree ($s \in S^0$):* Estimation of $p(x_s|y_s)$ and $p(x_s, x_{s^-}|y_s)$.
- *For $s \in S^1 \cdots S^{R-1}$:* Estimation of $p(x_s|y_{d(s)})$ and $p(x_s, x_{s^-}|y_{d(s)})$ using (16) and (14).
- *For $s \in S^R$:* Estimation of $p(x_s|y_{d(s)}) = p(x_s|y)$.

Top-down pass

- *Initialization at the root ($s \in S^R$):* $\hat{x_s} = \underset{x_s}{\operatorname{argmax}}\, p(x_s|y)$.
- *For $s \in S^{R-1}...S^0$:* Estimation of $p(x_s|y)$ by (17) then $\hat{x_s} = \underset{x_s}{\operatorname{argmax}}\, p(x_s|y)$.

Our classification map is obtained thanks to the estimates of $\hat{x_s}$ for each site of the leaves.

This algorithm is also illustrated in Fig. 7.

5 Preliminary results about prior probability estimation

In this section, we show results obtained by applying the MPM algorithm to a quad-tree structure. First, we estimated the likelihoods thanks to the CoDSEM algorithm applied to the original SAR image, its detail coefficients obtained by DWT (discrete wavelet transform) [27] and their corresponding extracted textural features. The prior information at the root $p(x_s)$ is first considered as uniform, thus equals to $1/M$. The prior probabilities at other levels is given by equation (11).

We choose a decomposition on $R = 2$ levels, motivated by the fact that given the size of the input image (typically 800×800 pixels), higher levels would imply too small images at the root level. Besides, $R = 2$ and $R = 3$ lead to similar result qualities both visually and quantitatively.

As described in Sec. 4.3, we improved the prior information extracting it from a preliminary classification. We run the classification first with an uniform prior, and the obtained classification map is used as a prior (input) for a new MPM-based classification.

The comparison between using an uniform prior and using preliminary results is shown in Tab. 1 for a COSMO-SkyMed acquisition over the Port-au-Prince quay in Haiti (©ASI) (Fig. 8). It is a single-pol image of size 920×820 pixels whose characteristics are: HH polarization, StripMap acquisition mode (2.5 m pixel spacing), geocoded. We want to classify this image into 3 distinct classes: urban areas, water and vegetation. Only quantitative results (confusion matrix on a manually-predefined ground truth) are given here.

Table 1: Accuracy for each of the 3 classes and overall results for the test areas obtained by considering different priors for the acquisition over Port-au-Prince.

	Port-au-Prince			
	water	urban	vegetation	**overall**
Uniform prior	97.62%	97.25%	97.42%	**97.43%**
Estimated prior	97.62%	97.45%	97.59%	**97.55%**

The accuracies are not perceptibly improved here, but if we consider the acquisition of Cavallermaggiore (Fig. 20), and make the same comparison, this step becomes more relevant (see Tab. 2).

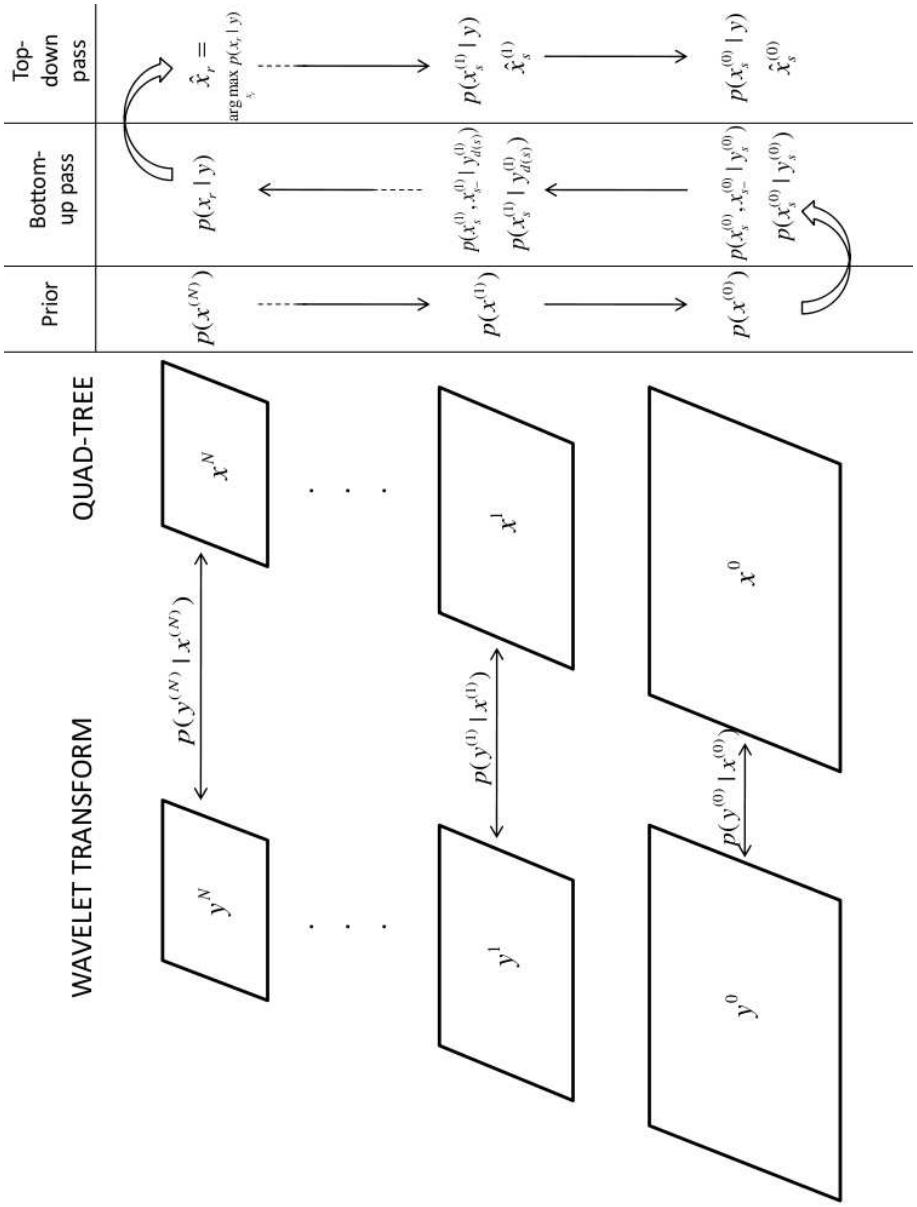


Figure 7: MPM estimation on the quad-tree.

6 Improvements

The first tests were satisfying, but we introduced some improvements to increase the classification results by making the hierarchical-based algorithm more robust with respect to speckle noise. The noise influence is probably due to the consideration of scale interactions, and not only a spatial context. We illustrate the improvement steps on an acquisition over Port-au-Prince in Haiti (Fig. 8).

Table 2: Accuracy for each of the 3 classes and overall results for the test areas obtained by considering different priors for the acquisition over Cavallermaggiore.

	Cavallermaggiore			
	water	urban	vegetation	**overall**
Uniform prior	93.85%	86.45%	94.53%	**91.61%**
Estimated prior	96.55%	87.65%	95.50%	**93.23%**

Figure 8: Original SAR image of the Port-au-Prince quay in Haiti (©ASI, 2009).

6.1 Prior estimation

First, we improve the classification results by better estimating the prior at the coarsest scale. That is why we suggest to update the prior, meaning that we proceed as previously, by applying a MPM estimation on a R-scale tree first choosing an uniform prior, and then by using the preliminary results as a new prior. Then, we consider a smaller tree of scale $R-1$ on which we apply the MPM algorithm to estimate a new prior. We proceed iteratively until scale 0 is reached. The scheme of this modification is given in Fig. 9.

This method allows a more robust classification, thus increasing the accuracy when compared to the basic MPM tree (Tab. 3).

Table 3: Accuracy for each of the 3 classes and overall results for the test areas obtained by comparing former results and the one obtained with the new tree scheme.

	Port-au-Prince quay			
	water	urban	vegetation	**overall**
Former method	97.62%	97.45%	97.59%	**97.55%**
With modification	97.74%	97.92%	98.15%	**97.94%**

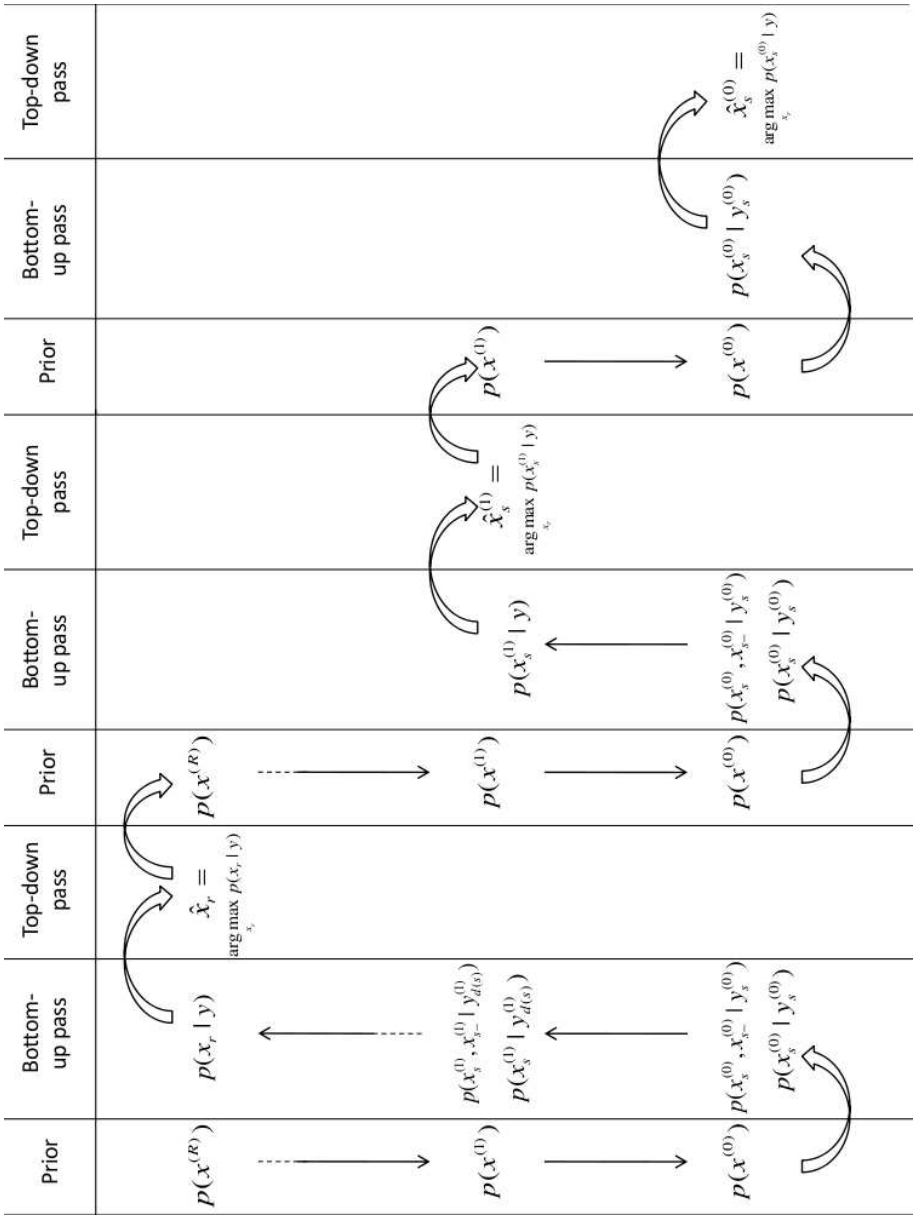


Figure 9: Modified MPM estimation on the quad-tree. In this representation, $R = 2$.

6.2 Adaptive neighborhood

When estimating the priors, we use a Markovian spatial model (see Sec. 4.3). Until now, we considered a second-order neighborhood based on the 8 pixels surrounding a considered pixel. We suggest to use an adaptive neighborhood, meaning that we consider different kinds of neighbor sets (Fig. 10) and select the one which leads to the smallest energy [48] [49].

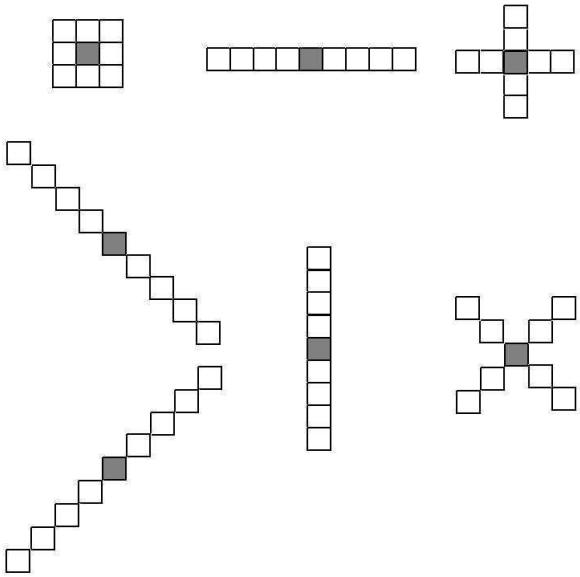


Figure 10: Considered neighbor sets.

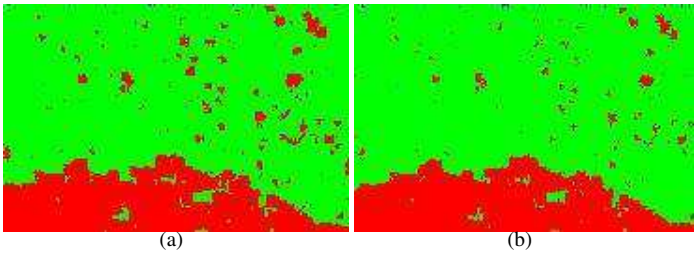
6.3 Beta parameter influence

By Eq. (12), we can notice that there is an unknown β parameter to determine. As in [16], this parameter can be estimated by minimizing a pseudo-likelihood over a training set. We stress here that this method brings to adequate estimates only when using an exhaustive ground truth, or at least by taking into account a sufficient amount of class borders, which is rarely the case in remote sensing. For this reason, we determine this parameter by trial-and-error. The results of different experiments shown in Fig. 11 have lead us to choose a quite high β parameter, with $\beta = 4.8$.

6.4 Optional smoothing

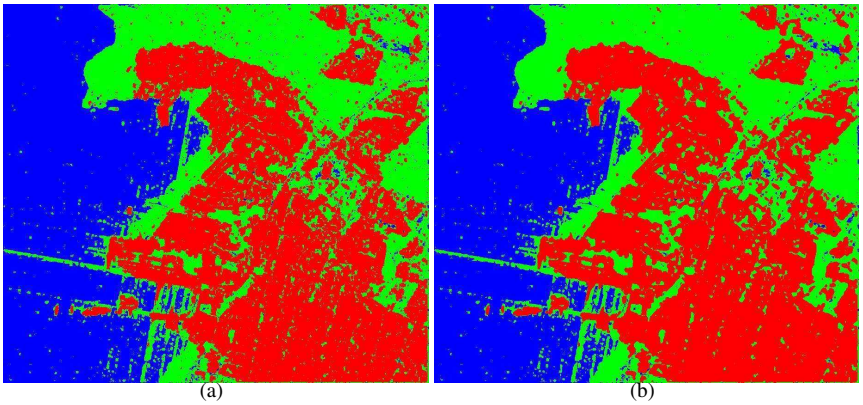
This final procedure is done outside the classification algorithm itself. We notice that applying a majority voting at each pixel of the final classification map can smooth the final classification map, thus allowing to smooth the remaining effects of speckle noise (Fig .12).

We zoomed in a portion of the classification maps so as to better compare the smoothing effects in Fig. 13.



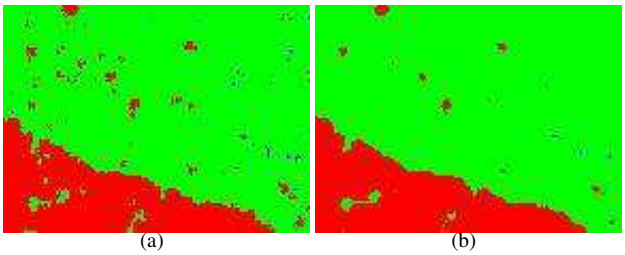
(a) (b)

Figure 11: Portion of classification maps obtained on Port-au-Prince acquisition using different β parameter values (a): $\beta = 1.2$ leading to an overall accuracy of 97.43%; (b) $\beta = 4.8$, leading to an overall accuracy of 97.94% .



(a) (b)

Figure 12: Classification maps obtained on Port-au-Prince acquisition (a): without smoothing, leading to an overall accuracy of 97.94%; (b) with smoothing, leading to an overall accuracy of 98.53%.



(a) (b)

Figure 13: Classification maps obtained on Port-au-Prince acquisition (a): without smoothing, leading to an overall accuracy of 97.94%; (b) with smoothing, leading to an overall accuracy of 98.53%.

This step is, of course, 'optional', and may be possibly applied after every image classifier that was used. In fact, this smoothing allows to remove some effects due to

the noise, but at the meantime, it can remove some details of the image. Thus its use depends on the application.

7 Final results

We compare in this section the results obtained for different SAR images by applying respectively:

- the CoDSEM-MRF algorithm [16]
- a K-nearest neighbor (K-NN) algorithm integrated in a hidden MRF scheme (K-NN-MRF)
- the CoDSEM-MPM algorithm detailed in Sec. 4
- the modified CoDSEM-MPM algorithm, according to Sec. 6 modifications.

The different subsections correspond to the different acquisition sites. For each considered method and each data set, we give the final classification map and the corresponding confusion matrix. The considered classes are water (in blue in the classification maps), urban areas (in red) and vegetation (in green).

Port-au-Prince quay

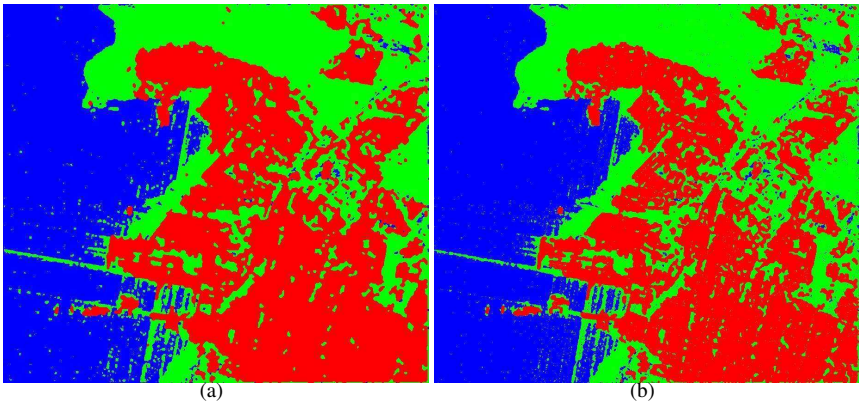
The considered image is a single-pol COSMO-SkyMed image of the quay of Port-au-Prince (Haiti) (©ASI, 2009), HH polarization, StripMap acquisition mode (2.5 m pixel spacing), geocoded, single-look image. 920×820 pixels.

Figure 14: Original SAR image of Port-au-Prince in Haiti (©ASI).

The results obtained for both CoDSEM-MRF and modified CoDSEM-MPM are very similar. The main difference lies in the visual maps, on which we can notice the smoothing effects of the first method (Fig. 15 and 16). The main water misclassifications come from the huge cross actifact, due to intrisic properties of SAR acquisitions (point spread function of the SAR sensor [1]). None of the algorithms employed here seems robust to such artifacts. The best results for this acquisitions are obtained with

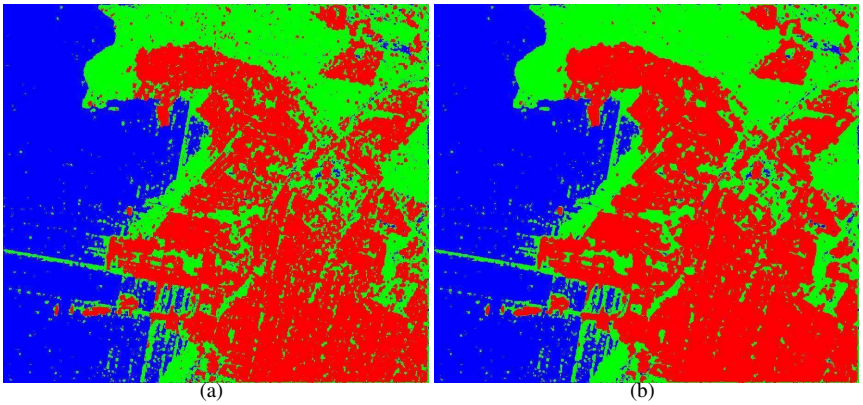
Table 4: Accuracy for each of the 3 classes and overall results for the test areas of the Port-au-Prince quay.

	Port-au-Prince quay			
	water	urban	vegetation	**overall**
CoDSEM-MRF	97.59%	99.03%	99.28%	**98.63%**
K-NN-MRF	99.23%	98.79%	99.64%	**99.22%**
CoDSEM-MPM	97.62%	97.45%	97.59%	**97.55%**
Modified CoDSEM-MPM	97.65%	98.99%	98.96%	**98.53%**



(a) (b)

Figure 15: Port-au-Prince quay classification maps (a): CoDSEM-MRF; (b) K-NN.



(a) (b)

Figure 16: Port-au-Prince quay classification maps (a): CoDSEM-MPM; (b) modified CoDSEM-MPM.

the K-NN method, but illustrations with other acquisitions show that it is not always the case.

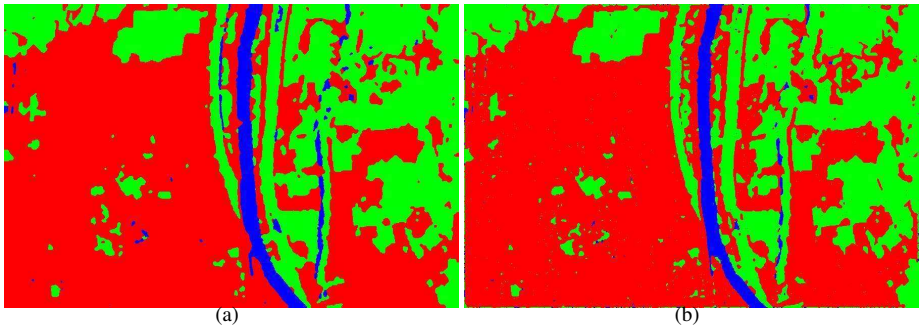
Rosenheim

The considered image is a single-pol TerraSAR-X image of the city of Rosenheim (Germany) (©DLR, 2008), HH polarization, SpotLight acquisition mode (8.2 m pixel spacing), geocoded, 4-look image. 900×600 pixels.

Figure 17: Original SAR image of Rosenheim in Germany (©DLR).

Table 5: Accuracy for each of the 3 classes and overall results for the test areas of Rosenheim.

	Rosenheim			
	water	urban	vegetation	**overall**
CoDSEM-MRF	97.11%	98.69%	99.72%	**98.51%**
K-NN-MRF	90.56%	98.49%	97.99%	**95.68%**
CoDSEM-MPM	84.46%	94.55%	95.63%	**91.55%**
Modified CoDSEM-MPM	94.35%	98.60%	99.18%	**97.38%**



(a) (b)

Figure 18: Rosenheim classification maps (a): CoDSEM-MRF; (b) K-NN.

In this specific case, most of the trees (top-center of the original image, Fig. 17) are misclassified as urban area. This is due to the fact that the statistics of the trees are very

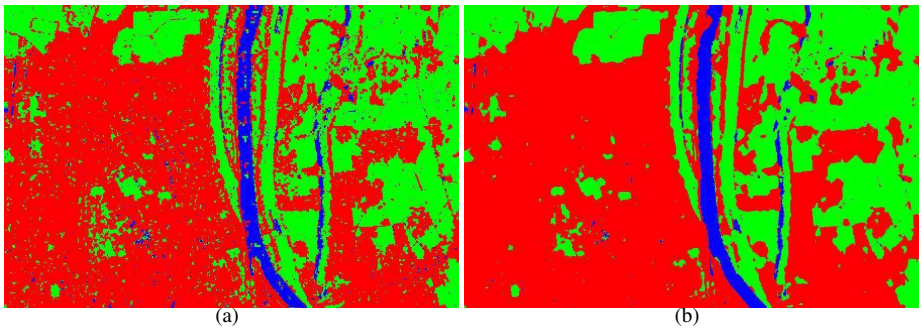


Figure 19: Rosenheim classification maps (a): CoDSEM-MPM; (b) modified CoDSEM-MPM.

similar to the urban ones. The additional information (Haralick textural feature) does not help for such discrimination; we can suppose that a more sophisticated textural feature, or the availabily of a differently polarized acquistion (e.g. VV polarization) would be helpful. Similarly, the shadows of the trees misclassified as water could be avoided. The introduction of a shadow class would not be very useful because the statistics are very similar to the water statistics, and it would generate much more misclassifications.

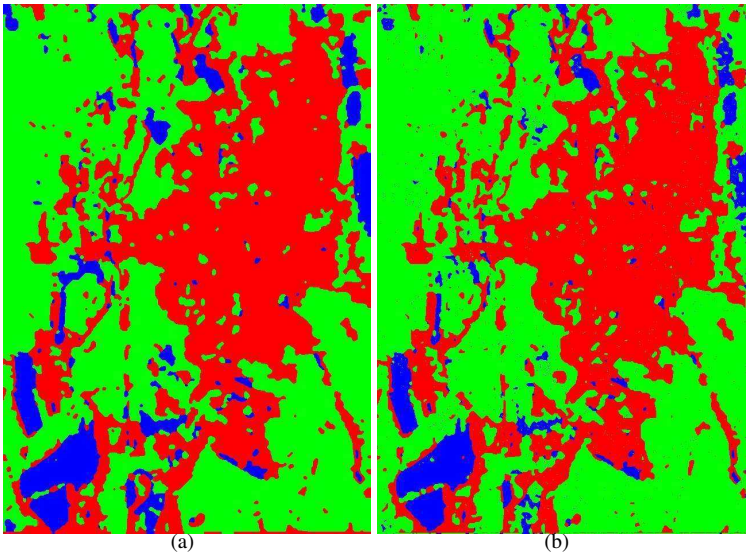
Cavallermaggiore

The considered image is a single-pol COSMO-SkyMed image of the city of Cavallermaggiore (Italy) (©ASI, 2008), HH polarization, StripMap acquisition mode (2.5 m pixel spacing), geocoded, single-look image. 650×950 pixels.

Figure 20: Original SAR image of Cavallermaggiore in Italy (©ASI).

Table 6: Accuracy for each of the 3 classes and overall results for the test areas of Cavallermaggiore.

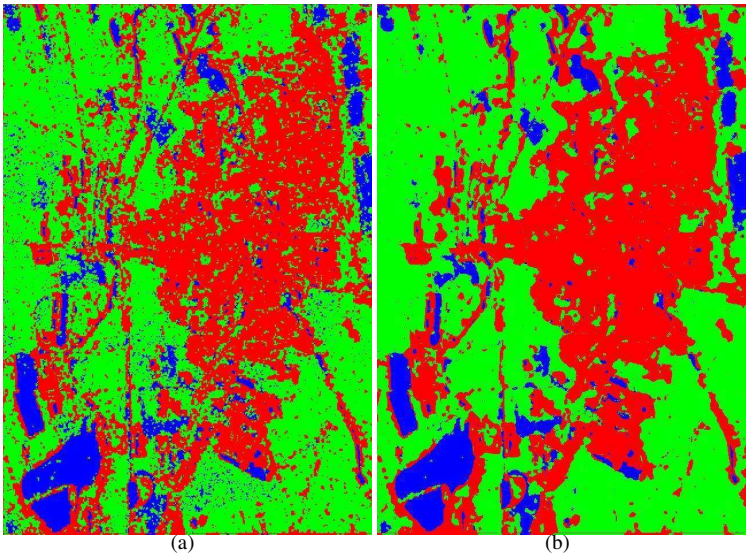
	Cavallermaggiore			
	water	urban	vegetation	**overall**
CoDSEM-MRF	98.50%	96.24%	99.87%	**98.20%**
K-NN-MRF	95.92%	94.62%	99.86%	**96.80%**
CoDSEM-MPM	96.53%	87.63%	95.50%	**93.22%**
Modified CoDSEM-MPM	98.79%	95.59%	99.82%	**98.07%**



(a) (b)

Figure 21: Cavallermaggiore classification maps (a): CoDSEM-MRF; (b) K-NN.

This example highlights the classification improvements when comparing CoDSEM-MRF and modified CoDSEM-MPM (Fig. 21 and 22). In fact, the method introduced in this paper leads to more detailed results: vegetation inside the urban area and large roads are better taken into account.



(a) (b)

Figure 22: Cavallermaggiore classification maps (a): CoDSEM-MPM; (b) modified CoDSEM-MPM.

Amiens

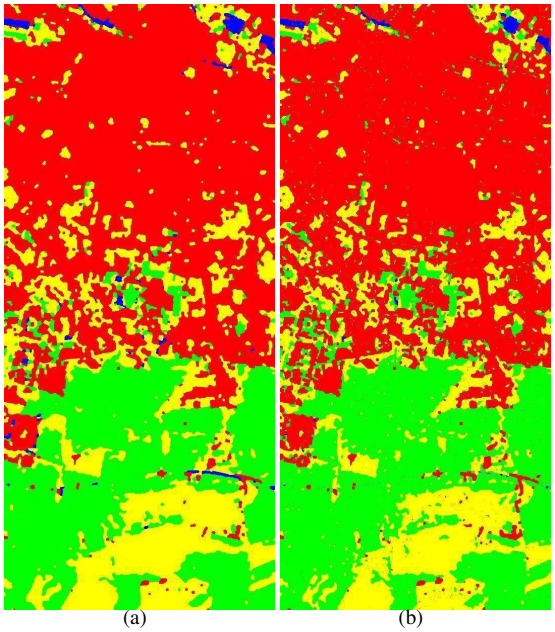
The considered image is a single-pol COSMO-SkyMed image of the city of Amiens (France) (©ASI, 2011), HH polarization, StripMap acquisition mode (2.5 m pixel spacing), geocoded, single-look image. 530×1200 pixels. In this specific case, we deal with 4 classes: urban (in red), water (in blue), vegetation (in green) and trees (in yellow).

Table 7: Accuracy for each of the 4 classes and overall results for the test areas of Amiens.

	Amiens				
	water	urban	vegetation	tree	**overall**
CoDSEM-MRF	98.92%	98.46%	87.38%	95.80%	**95.14%**
K-NN-MRF	90.69%	98.62%	88.45%	93.53%	**92.82%**
CoDSEM-MPM	95.01%	98.23%	83.03%	94.66%	**92.73%**
Modified CoDSEM-MPM	98.40%	98.85%	87.78%	97.57%	**95.65%**

In this example, we show that our method is more accurate on more classes. Similar comments to the previous ones (see other experiments) can be done: CoDSEM-MRF and modified CoDSEM-MPM lead to similar results, but the second one highlights more details (Fig. 24 and 25). The loss of accuracy in the vegetation class for all the methods (Tab. 7) is due to misclassifications of this area as forest at the bottom of the image. By only looking at the SAR image, we cannot really see a difference between these crop and forest areas.

Figure 23: Original SAR image of Amiens in France (©ASI).



(a) (b)

Figure 24: Amiens classification maps (a): CoDSEM-MRF; (b) K-NN.

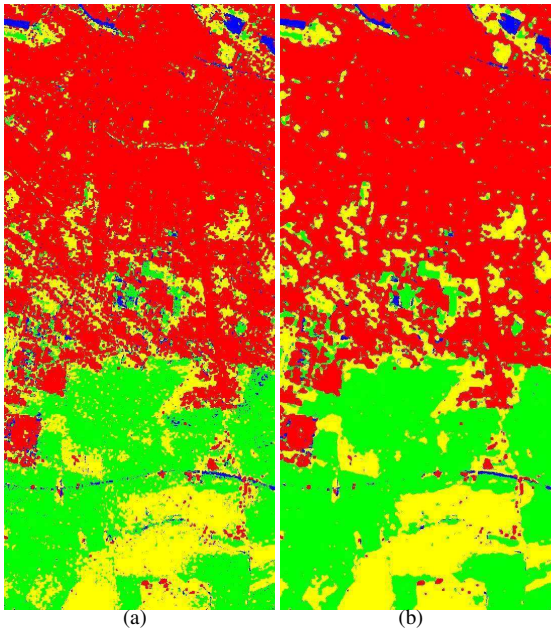


Figure 25: Amiens classification maps (a): CoDSEM-MPM; (b) modified CoDSEM-MPM.

8 Conclusion

In this report, we have combined a SAR amplitude statistics modeling via finite mixtures and copulas, with a hierarchical Markov random field, leading to a supervised Bayesian classification. We compared this hierarchical model to a spatial MRF, and introduced some modifications so as to improve the final classification accuracy. The results are assessed both qualitatively (classification maps) and quantitatively (confusion matrices). We could notice that they lead to similar results. The hierarchical MRF considered here has two advantages: it is quite robust to speckle noise, and we can apply a non-iterative optimization algorithm (MPM estimation). Besides, one main advantage when compared to a basic spatial MRF is that it can be extended to the use of not only single-pol images, as presented in this report, but also to multi-resolution or multi-sensor data. The study of the classification of multi-sensor data is considered as major direction of further research.

Acknowledgements

The authors would like to thank:

- The Direction Générale de l'Armement (DGA, France) and Institut National de Recherche en Informatique et Automatique (INRIA, France) for the partial financial support.
- The Italian Space Agency (ASI) for providing the COSMO-SkyMed images.
- The German Space Agency (DLR) for providing the TerraSAR-X image.

• Dr. Michaela DeMartino from the University of Genoa (Italy) for preparing the ground truth maps.

References

[1] C. Oliver and S. Quegan. *Understanding Synthetic Aperture Radar images*. SciTech Publishing, 2004.

[2] G. Boni, F. Castelli, L. Ferraris, N. Pierdicca, S. Serpico, and F. Siccardi. High resolution COSMO-SkyMed SAR data analysis for civil protection from flooding events. In *IEEE International Geoscience and Remote Sensing Symposium*, IGARSS'07, pages 6–9, 2007.

[3] A. M. Jacob, E. M. Hemmerly, and D. Fernandes. SAR image classification using a neural classifier based on Fisher criterion. In *Proceedings of the VII Brazilian Symposium on Neural Networks*, SBRN '02, pages 168–172, 2002.

[4] X. Xue, Y. Zhang, R. Zhao, F. Duan, and Y. Chen. A new method of SAR image segmentation based on neural network. In *Proceedings of the 5th International Conference on Computational Intelligence and Multimedia Applications*, ICCIMA '03, pages 149–153, 2003.

[5] M. Silveira and S. Heleno. Separation between water and land in SAR images using region-based level sets. *IEEE Geosci. Remote Sens. Lett.*, 6(3):471–475, 2009.

[6] W. Yang, D. Dai, B. Triggs, and G. Xia. Semantic labeling of SAR images with hierarchical Markov aspect models. Research report hal-00433600, HAL, University of Wuhan, 2009.

[7] Su Fulin, Ni Liang, Li Dafang, and Sun Huadong. Classification of SAR image based on gray cooccurrence matrix and support vector machine. In *Proceedings of the 7th International Conference on Signal Processing*, volume 2 of *ICSP'04*, pages 1385–1388, 2004.

[8] V. V. Chamundeeswari, D. Singh, and K. Singh. Unsupervised land cover classification of SAR images by contour tracing. In *IEEE International Geoscience and Remote Sensing Symposium*, IGARSS'07, pages 547–550, 2007.

[9] R. Fjortoft, Y. Delignon, W. Pieczynski, M. Sigelle, and F. Tupin. Unsupervised classification of radar images using hidden Markov chains and hidden Markov random fields. *IEEE Trans. Geosci. Remote Sens.*, 41(3):675–686, 2003.

[10] D. Benboudjema and W. Pieczynski. Unsupervised image segmentation using triplet Markov fields. *Computer Vision and Image Understanding*, 99(3):476–498, 2005.

[11] G. Moser, S. B. Serpico, and J. Zerubia. Dictionary-based Stochastic Expectation Maximization for SAR amplitude probability density function estimation. *IEEE Trans. Geosci. Remote Sens.*, 44(1):188–199, 2006.

[12] V. Krylov, G. Moser, S. B. Serpico, and J. Zerubia. Enhanced dictionary-based SAR amplitude distribution estimation and its validation with very high-resolution data. *IEEE Geosci. Remote Sens. Lett.*, 8(1):148–152, 2011.

[13] R. C. Dubes and A. K. Jain. Random field models in image analysis. *Journal of Applied Statistics*, 16(2):131–164, 1989.

[14] C. Graffigne, F. Heitz, P. Pérez, F. Prêteux, M. Sigelle, and J. Zerubia. Hierarchical Markov random field models applied to image analysis: a review. In *Proc. of the Conf. on Neural, Morphological and Stochastic Methods in Image Processing*, volume 2568 of *SPIE's International Symposium on Optical Science, Engineering and Instrumentation*, 1995.

[15] P. Fieguth. *Statistical image processing and multidimensional modeling*. Springer, 2011.

[16] A. Voisin, G. Moser, V. Krylov, S. B. Serpico, and J. Zerubia. Classification of very high resolution SAR images of urban areas by dictionary-based mixture models, copulas and Markov random fields using textural features. In *Proc. of SPIE*, volume 7830 of *SPIE Symposium on Remote Sensing 2010*, page 78300O, 2010.

[17] V. Krylov, G. Moser, S. B. Serpico, and J. Zerubia. Supervised high resolution dual polarization SAR image classification by finite mixtures and copulas. *IEEE Journal of Selected Topics in Signal Processing*, 5(3):554–566, 2011.

[18] B. Gidas. A renormalization group approach to image processing problems. *IEEE Trans. Patt. Anal. Mach. Intell.*, 11(2):164–180, 1989.

[19] F. Heitz, P. Perez, and P. Bouthemy. Multiscale minimization of global energy functions in some visual recovery problems. *CVGIP: image understanding*, 59(1):125–134, 1994.

[20] J.-M. Laferte, P. Perez, and F. Heitz. Discrete Markov modeling and inference on the quad-tree. *IEEE Trans. Image Processing*, 9(3):390–404, 2000.

[21] Z. Kato, M. Berthod, and J. Zerubia. A hierarchical markov random field model and multitemperature annealing for parallel image classification. *Graphical models and image processing*, 58(1):18–37, 1996.

[22] J.-M. Laferte, F. Heitz, P. Perez, and E. Fabre. Hierarchical statistical models for the fusion of multiresolution image data. In *Proceedings of the 5th International Conference on Computer Vision*, ICCV'95, pages 908–913, 1995.

[23] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.*, 3(6):610–621, 1973.

[24] P. Curran. The semivariogram in remote sensing: an introduction. *Remote Sensing Of Environment*, 24(3):493–507, 1988.

[25] Q. Chen and P. Gong. Automatic variogram parameter extraction for textural classification of the panchromatic IKONOS imagery. *IEEE Trans. Geosci. Remote Sens.*, 42(5):1106–1115, 2004.

[26] S. G. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Patt. Anal. Mach. Intell.*, 11(7):674–693, 1989.

[27] S. G. Mallat. *A wavelet tour of signal processing*. Academic Press, 3rd edition, 2008.

[28] I. Daubechies. Orthonormal bases of compactly supported wavelets. *Communications on Pure and Applied Mathematics*, 41:909–996, 1988.

[29] M.Vetterli, J. Kovacevic, and V. K. Goyal. *Fourier and wavelet signal processing*. Protected by copyright under the Attribution-NonCommercial-NoDerivs 3.0 Unported License from Creative Commons, 2010. http://www.fourierandwavelets.org.

[30] J. Zerubia, Z. Kato, and M. Berthod. Multi-temperature annealing: a new approach for the energy-minimization of hierarchical Markov random field models. In *Proceedings of the 12th IAPR International Conference on Pattern Recognition*, volume 1, pages 520–522, 1994.

[31] E. Fabre. *Traitement du signal multi-résolution : conception de lisseurs rapides pour une famille de modèles*. Thèse de doctorat, Université de Rennes 1, France, 1994.

[32] M. R. Luettgen, W. Karl, and A. S. Willsky. Efficient multiscale regularization with applications to the computation of optical flow. *IEEE Trans. Image Processing*, 3(1):41–64, 1992.

[33] C. Bouman and M. Shapiro. A multiscale random field model for Bayesian image segmentation. *IEEE Trans. Image Processing*, 3(2):162–177, 1994.

[34] J.-M. Laferte. *Contribution à l'analyse d'images par modèles markoviens sur des graphes hiérarchiques. Application à la fusion de données multirésolutions*. Thèse de doctorat, Université de Rennes 1, 1996.

[35] J. Marroquin, S. Mitter, and T. Poggio. Probabilistic solution of ill-posed problems in computational vision. *Journal of the American Statistical Association*, 82(397):76–89, 1987.

[36] M. A. T. Figueiredo and A. K. Jain. Unsupervised learning of finite mixture models. *IEEE Trans. Patt. Anal. Mach. Intell.*, 24(3):381–396, 2000.

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

[38] G. Celeux, D. Cheveau, and J. Diebolt. On stochastic versions of the EM algorithm. Research report 2514, INRIA, 1995.

[39] R. B. Nelsen. *An introduction to copulas*. Springer, New York, 2nd edition, 2006.

[40] D. Huard, G. Evin, and A.-C. Fabre. Bayesian copula selection. *Computational statistics and data analysis*, 51(2):809–822, 2006.

[41] E. Lehmann and J. Romano. *Testing statistical hypotheses*. Springer, 3rd revised edition, 2007.

[42] T. Cover and P. Hart. Nearest neighbor pattern classification. *IEEE Trans. on Information Theory*, 13(1):21–27, 1967.

[43] J. Besag. Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society*, 36(2):192–236, 1974.

[44] L. E. Baum, T. Petrie, G. Soules, and N. Weiss. A maximization technique occuring in the statistical analysis of probabilistic functions of Markov chains. *IEEE Ann. Math. Stats*, 41(1):164–171, 1970.

[45] Z. Kato, J. Zerubia, and M. Berthod. Satellite image classification using a Modified Metropolis Dynamics. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 3 of *ICASSP'92*, pages 573–576, 1992.

[46] J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society*, 48(3):259–302, 1986.

[47] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Patt. Anal. Mach. Intell.*, 6(6):721–741, 1984.

[48] P.C.Smits and S.G. Dellepiane. Synthetic aperture radar image segmentation by a detail preserving Markov random field approach. *IEEE Trans. Geosci. Remote Sens.*, (4):844–857, 1997.

[49] P. Zhong, F. Liu, and R. Wang. A new MRF framework with dual adaptive contexts for image segmentation. In *International Conference on Computational Intelligence and Security*, pages 351–355, 2007.

Centre de recherche INRIA Sophia Antipolis – Méditerranée
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Centre de recherche INRIA Bordeaux – Sud Ouest : Domaine Universitaire - 351, cours de la Libération - 33405 Talence Cedex
Centre de recherche INRIA Grenoble – Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier
Centre de recherche INRIA Lille – Nord Europe : Parc Scientifique de la Haute Borne - 40, avenue Halley - 59650 Villeneuve d'Ascq
Centre de recherche INRIA Nancy – Grand Est : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex
Centre de recherche INRIA Paris – Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex
Centre de recherche INRIA Rennes – Bretagne Atlantique : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex
Centre de recherche INRIA Saclay – Île-de-France : Parc Orsay Université - ZAC des Vignes : 4, rue Jacques Monod - 91893 Orsay Cedex

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399