



**HAL**  
open science

# Supervised Classification of Multi-sensor and Multi-resolution Remote Sensing Images with a Hierarchical Copula-based Approach

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

► **To cite this version:**

Aurélie Voisin, Vladimir Krylov, Gabriele Moser, Sebastiano B. Serpico, Josiane Zerubia. Supervised Classification of Multi-sensor and Multi-resolution Remote Sensing Images with a Hierarchical Copula-based Approach. *IEEE Transactions on Geoscience and Remote Sensing*, 2014, 52 (6), pp.3346-3358. 10.1109/TGRS.2013.2272581 . hal-00841234

**HAL Id: hal-00841234**

**<https://inria.hal.science/hal-00841234>**

Submitted on 4 Jul 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Supervised Classification of Multi-sensor and Multi-resolution Remote Sensing Images with a Hierarchical Copula-based Approach

Aur lie Voisin, Vladimir A. Krylov, Gabriele Moser, *Member, IEEE*,  
Sebastiano B. Serpico, *Fellow, IEEE*, and Josiane Zerubia, *Fellow, IEEE*

## Abstract

In this paper we develop a novel classification approach for multi-resolution, multi-sensor (optical and synthetic aperture radar, SAR) and/or multi-band images. This challenging image processing problem is of great importance for various remote sensing monitoring applications and has been scarcely addressed so far. To deal with this classification problem, we propose a two-step explicit statistical model. We first design a model for the multivariate joint class-conditional statistics of the co-registered input images at each resolution by resorting to multivariate copulas. Such copulas combine the class-conditional marginal probability density functions of each input channel that are estimated by finite mixtures of well-chosen parametric families. We consider different distribution families for the most common types of remote sensing imagery acquired by optical and SAR sensors. We then plug the estimated joint probability density functions into a hierarchical Markovian model based on a quad-tree structure, where each tree-scale corresponds to the different input image resolutions and to corresponding multi-scale decimated wavelet transforms, thus preventing a strong resampling of the initial images. To obtain the classification map, we resort to an exact estimator of the marginal posterior mode. We integrate a prior update in this model in order to improve the robustness of the developed classifier against noise and speckle. The resulting classification performance is illustrated on several remote sensing multi-resolution datasets including very high resolution and multi-sensor images acquired by COSMO-SkyMed and GeoEye-1.

## Index Terms

Multi-sensor remote sensing images, supervised classification, hierarchical Markov random fields, multi-resolution data.

## I. INTRODUCTION

A WIDE variety of remote sensing images are available nowadays, and the high diversity of camera/sensor properties explains the difficulty of developing generally applicable image processing algorithms. In this paper, we present a general classifier that can be applied to single-channel, as well as to multi-channel imagery, and that supports input images originated from different sensors and with distinct spatial resolutions, with the underlying assumption that the acquired images are co-registered [1]. In this paper, we focus on two of the most common classes of remote sensing imagery: optical and synthetic aperture radar [2] (SAR) acquisitions.

Two common ways are usually adopted to classify multi-sensor images acquired at different resolutions: the combination of classifiers and the classification of preliminary fused images. The first strategy is based on the hypothesis that a well-chosen combination of classifier outputs may lead to an improved global classification. A basic rule is the majority voting [3], but more specific methods have been proposed, such as those relying on neural networks [4] or the Dempster-Shafer theory [3], [5]. Wolpert [6] introduced the stacked generalization, in which the

This research was a collaborative effort between the National Institute for Research in Computer Science and Control, (INRIA), Sophia Antipolis, France, and the Dept. of Electrical, Electronic, Telecommunications Engineering and Naval Architecture (DITEN) of the University of Genoa, Italy.

A. Voisin and J. Zerubia are with the Ayn research team, INRIA, F-06902, Sophia Antipolis, France (e-mail: Aurelie.Voisin@inria.fr, Josiane.Zerubia@inria.fr). The work of A. Voisin was supported by the Direction G n rale de l'Armement (France) and by INRIA.

V. A. Krylov is with the Dept. of Statistical Science, University College London, WC1E 6BT, London, UK (email: v.krylov@ucl.ac.uk). His work was funded by an INRIA post-doctoral grant.

G. Moser, S. B. Serpico are with the DITEN, University of Genoa, I-16145, Genoa, Italy (e-mail: Gabriele.Moser@unige.it, Sebastiano.Serpico@unige.it).

classifier outputs are combined according to their individual performances. Fuzzy sets have also been studied [7]. In a Bayesian context, classifier outputs usually correspond to posterior probabilities, and various combinations are possible, from the simple posterior probability averaging [3] to a more complex weighted linear mixtures, e.g., Linear Opinion Pool (LOP) [8], or by resorting to the use of norms [9]. The second strategy relies on the direct fusion of the input images. In this case, the classification is obtained by applying to the fused images some widely used classification algorithms such as, for instance, methods based on Markov random fields (MRF) [10], on the Bayes decision rule for minimum error [11], on maximum of likelihood algorithms [12] or on machine-learning [13]. Some well-known fusion strategies are the Intensity-Hue-Saturation (IHS) [14], [15] or the Principal Component Analysis (PCA)-based methods [15]. High-pass filtering [16], Gram-Schmidt method [16] and projection to Gaussianity [17] have been investigated as well. Several sharpening techniques are presented in [18].

In the proposed method, the multi-sensor and multi-resolution fusion is based on explicit statistical modeling. Specifically, it consists in finding a joint probability distribution given the class-conditional marginal probability density functions (PDFs) related to each input channel. For each class and each channel in the chosen stacked-vector input dataset, the class-conditional PDFs are estimated by finite mixtures of specific parametric families [19]. For optical imagery, the normal distribution is a commonly accepted model, and the parameters of the employed normal PDF mixtures are estimated in the proposed method by the stochastic expectation maximization (SEM) [20] algorithm. For SAR amplitude data, the generalized gamma distribution [21] is an adequate choice, both thanks to its accurate results for this data typology and because it is a generalization of several earlier developed classical parametric SAR models, such as Nakagami or Rayleigh distributions. Nevertheless, log-normal, Nakagami and Weibull distributions may be other feasible options [19]. The mixture distribution parameters in the case of SAR data channels are determined by a modified SEM algorithm that employs the method of log-cumulants (MoLC) [22] for the estimation of the parameters of the generalized gamma distribution. Next, multivariate copulas [23] are applied to model the multivariate joint class-conditional statistics, merging the marginal PDF estimates of the coregistered individual input channels. The copula parameters are estimated by using the relationship with the Kendall's tau concordance-discordance coefficient [23]. The use of copulas enables to obtain flexible joint statistical models capable of describing various dependence structures, that are feasible and fast from the parameter estimation point-of-view [23], [24].

Most of the methods proposed in the literature have a tendency to degrade the initial image resolutions when resorting to a multi-input fusion. The novelty in this paper is in keeping the multi-resolution aspect by integrating the multi-sensor data in an explicit hierarchical Markovian model [25], [26] based on a quad-tree structure. To this end, the fusion is independently performed at each level of the tree, and aims at gathering the multiple bands at a given resolution into a joint statistical model. Multi-scale observations correspond to the initial multi-resolution acquisitions, and/or to their discrete wavelet decompositions. The consideration of a quad-tree allows to integrate an exact estimator of the marginal posterior mode (MPM) [27] that performs the unknown class labels estimation. The prior probabilities are iteratively updated at each level of the tree, leading to an algorithm more robust with respect to noise and speckle [2], when compared to constant (non-updated) priors [27].

The novel contribution of this paper is in the development of a statistical model integrated in a multi-resolution graph to generate a classifier based on a flexible and general statistical formulation that can be easily applied to multi-sensor and multi-resolution data. The multivariate PDFs are estimated by resorting to automatically generated multivariate copulas. The hierarchical MRF is adapted to integrate a robust-to-noise prior update, and flexibility in the amount of inputs at each tree resolution. Unlike the aforementioned previous techniques, which either focused on the classification of data acquired by multiple sensors at the same resolution or involved data-resampling to a common resolution and thus involved possible aliasing or artifacts, the resulting classification technique has an advantage of being applicable to a very wide and heterogeneous range of distinct resolution and multi-sensor (SAR and optical) imagery.

The paper is organized as follows. In Section II, we introduce the statistical multivariate copula-based model that combines marginal PDF models of the input images at each level. In Section III, we design a novel MPM model based on the quad-tree formulation. Section IV gives the developed method overview. In Section V, we present classification results obtained on two remote sensing multi-sensor and multi-resolution datasets involving very high resolution COSMO-SkyMed and GeoEye-1 imagery. Finally, Section VI presents conclusions of this study.

## II. MULTIVARIATE PDF MODEL

According to the statistical data fusion method we propose, we need to model the joint statistics of the input images at each resolution of the tree. In this section, we consider that we work at a given tree level  $n$ . Specifically, the input images may be, for instance, panchromatic, color optical images and single or multi-polarized SAR images. We propose, first, to model separately the statistics of each input band for each class, and then to estimate a multivariate PDF  $p(y|\omega_m)$ , where  $y = (y_1, \dots, y_d)$  represents the set of input images,  $d$  being the number of input bands (at a given level). To perform this multivariate modeling, we employ the flexible statistical modeling instrument of copulas [23].

It is well known that the reflectivity in SAR and optical bands are very different from each other [2]. To address the complicated problem of SAR+optical PDF modeling we, first, estimate the marginal class-conditional statistics of each SAR/optical channel separately via distinct finite mixtures and, then, we model the joint PDF through copulas. In the context of the supervised classification, the classes are defined a-priori and characterized through training pixels. This enables the choice of classes that demonstrate acceptable levels of inner-class statistical homogeneity on both SAR and optical acquisitions.

### A. Marginal PDF estimation

We want to model the distributions of each class  $\omega_m$  considered for the classification,  $m \in [1; M]$ , given a training set, for each input channel  $j$ . For each class, the PDF  $p_m(y_j|\omega_m)$  is modeled via finite mixtures [28] of independent greylevel distributions:

$$p_m(y_j|\omega_m) = \sum_{i=1}^{K_m} P_{mi} p_{mi}(y_j|\theta_{mi}), \quad (1)$$

$P_{mi}$  are the mixing proportions such that for a given  $m$ ,  $\sum_{i=1}^{K_m} P_{mi} = 1$  with  $0 \leq P_{mi} \leq 1$ .  $\theta_{mi}$  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 contributions of the components present in each class (for instance, different kinds of crops for the vegetation class). Moreover, the use of finite mixtures can be seen as a general method that allows to estimate both the best finite mixture model and/or the best single PDF model [28]. The mixture modeling is performed differently working with the following common types of remote sensing imagery:

1) *Optical images*: When the  $j^{th}$  input is an optical acquisition, we consider that the PDF  $p_m(y_j|\omega_m)$  related to each class can be modeled by a finite mixture of Gaussian distributions, hence

$$p_{mi}(y_j|\theta_{mi}) = \frac{1}{\sqrt{2\pi\sigma_{mi}^2}} \exp\left[-\frac{(y_j - \mu_{mi})^2}{2\sigma_{mi}^2}\right], \quad (2)$$

with  $\theta_{mi} = \{\mu_{mi}, \sigma_{mi}^2\}$ , where the mean  $\mu_{mi}$  and the variance  $\sigma_{mi}^2$  are estimated within a stochastic expectation maximization (SEM) [20] algorithm.

2) *SAR amplitude images*: The SAR acquisitions are known to be affected by speckle [2]. For this reason, we use appropriate SAR-specific models for such images, such as the generalized gamma distribution [21]. This model is a generalization of the Rayleigh, log-normal, Nakagami and Weibull distributions, widely used in SAR amplitude imagery modeling [2]. Each class-conditional PDF is then modeled by a mixture of generalized gamma distributions:

$$p_{mi}(y_j|\theta_{mi}) = \frac{\nu_{mi} y_j^{\kappa_{mi}\nu_{mi}-1}}{\sigma_{mi}^{\kappa_{mi}\nu_{mi}} \Gamma(\kappa_{mi})} \exp\left[-\left(\frac{y_j}{\sigma_{mi}}\right)^{\nu_{mi}}\right], \quad (3)$$

with parameters  $\theta_{mi} = \{\nu_{mi}, \sigma_{mi}, \kappa_{mi}\}$  and where  $\Gamma(\cdot)$  is the standard gamma function [29].

A modified SEM algorithm is used to estimate the best-fitting mixture model for each considered class [19]. It combines a density parameter estimation via a SEM algorithm [20] and the method of log-cumulants (MoLC) [22], [30] instead of maximum likelihood estimates that are difficult for the generalized gamma distribution parametric family [30]. When the MoLC applicability condition for generalized gamma distribution is not respected [19], the PDF  $p_{mi}(y_j|\theta_{mi})$  is chosen as the one reporting the highest likelihood value among the following distributions: log-normal, Weibull, Nakagami.

In both the SAR and optical cases, the estimation of the number of components  $K_m$  is integrated into the SEM procedure by using the algorithm proposed in [19]. Through this algorithm,  $K_m$  is initialized with an upper bound on the number of components and, at each iteration of the SEM algorithm, if the proportion of the  $i^{\text{th}}$  component  $P_{mi}$  is below under a given threshold, this component is then eliminated, and  $K_m$  is decremented. We note that this procedure was proposed and validated in [19] for the case of SAR data and is extended here to optical images. For more details concerning the mixture estimation, see [19].

### B. Joint-distribution modeling via multivariate copulas

We now briefly recall some relevant properties and definitions for copulas. For a comprehensive introduction see [23]. A multivariate copula is a  $d$ -dimensional joint distribution defined on  $[0, 1]^d$  such that marginal distributions are uniform on  $[0, 1]$ . The importance of copulas in statistics is explained by Sklar's theorem [23], which states the existence of a copula  $C_m$  that models the joint distribution function  $H_m$  of arbitrary random variables  $\{Y_1, \dots, Y_d\}$  with cumulative distribution functions (CDFs)  $\{F_{1m}(y_1|\omega_m), \dots, F_{dm}(y_d|\omega_m)\}$ :

$$H_m(\mathbf{y}) = C_m(F_m(y_1|\omega_m), \dots, F_m(y_d|\omega_m)), \quad (4)$$

for any  $\mathbf{y} = (y_1, \dots, y_d)$ , with  $y_1, \dots, y_d \in \mathbb{R}$ . Note that the first index  $z = 1, \dots, d$  is omitted in  $F_{zm}(y_z|\omega_m)$  for brevity.

Taking the derivative of (4) with respect to the continuous random variables vector  $\mathbf{y}$  with PDFs  $\{p_m(y_1|\omega_m), \dots, p_m(y_d|\omega_m)\}$ , we obtain the joint PDF distribution:

$$h_m(\mathbf{y}) = c_m(F_m(y_1|\omega_m), \dots, F_m(y_d|\omega_m)) \prod_{j=1}^d p_m(y_j|\omega_m), \quad (5)$$

where  $c_m$  is the copula density, i.e. the derivative of the copula  $C_m$  over  $\{y_1, \dots, y_d\}$ .

We wish to estimate, for each class  $m$ , the joint PDF  $h_m$  given the marginal distributions  $\{p_m(y_1|\omega_m), \dots, p_m(y_d|\omega_m)\}$  that correspond to the marginal PDFs of the  $d$  input images (Section II-A). The CDF  $F_m(y_j|\omega_m)$  can be found as an integral on  $]-\infty; z_j]$  of its corresponding PDF  $p_m(y_j|\omega_m)$ . Thus, the parameters of the cumulative distribution functions  $\{F_m(y_1|\omega_m), \dots, F_m(y_d|\omega_m)\}$  are the same as the parameters of the marginal estimates in Eq. (1). In order to determine  $h_m$ , we have to first determine the copula family  $C_m$ .

The bivariate copulas have been studied extensively in the literature, which is not necessarily the case for multivariate copulas. We extend here the work previously carried out in [24] to the multivariate case. To find the best fitting copula  $C_m^*$ , we consider a dictionary of Clayton, Ali-Mikhail-Haq (AMH) and Gumbel copulas. These three are able to model a considerable variety of dependence structures [31] and have empirically demonstrated good fit to SAR image data [24], [32]. Further copula families [23] can easily be adopted as well to arrive at further results. The analytical expressions of each of these copulas<sup>1</sup> involve only one parameter, that is estimated by using the relationship between copulas and the Kendall's  $\tau$  ranking correlation coefficient [23]. The latter is defined as

$$\tau = \frac{1}{2^{d-1} - 1} \left( 2^d \int_{[0,1]^d} C_m(u) dC_m(u) - 1 \right). \quad (6)$$

This allows to estimate the parameters in the considered copula-dictionary once we have an estimate of the Kendall's  $\tau$ . This coefficient is a ranking measure of concordance of two samples [23], and its consistent empirical estimator on samples of size  $N$  is given by [33]:

$$\hat{\tau} = \frac{4}{N(N-1)} \sum_{i \neq j} \left( I[Y_{1,i} \leq Y_{1,j}] I[Y_{2,i} \leq Y_{2,j}] \right) - 1, \quad (7)$$

where  $I[\cdot]$  is the indicator function,  $I[a \leq b] = 1$  if  $a \leq b$ , and  $I[a \leq b] = 0$  otherwise.  $(Y_1, Y_2)$  are different bivariate marginals corresponding to the observations of two distinct input channels and  $N$  is the sample size. When considering a higher number of input images, we estimate this measure by image pairs, and obtain the final  $\tau$  by

<sup>1</sup>The specific PDF expressions for these multivariate copulas can be easily obtained from the generator functions (see in [23]) by means of Matlab, and are omitted to avoid cumbersome notations.

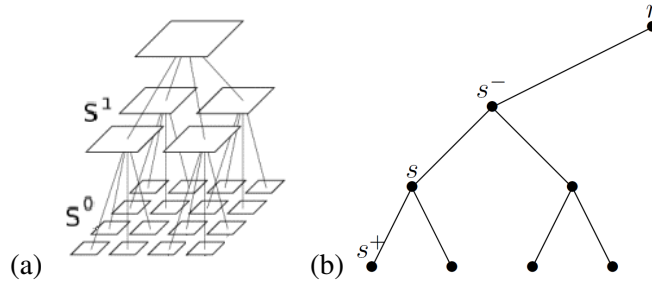


Fig. 1. (a) A quad-tree pixel structure and (b) the employed notations.

averaging [33]. It is worth mentioning, that we have observed stronger levels of dependence between SAR channels as witnessed by the Kendall's tau, i.e., typically,  $\tau_{\text{SAR-SAR}} \in [0.1, 0.7]$ , than between SAR and optical images with  $\tau_{\text{SAR-optical}} \in [0.1, 0.2]$ .

For each class  $m$ , we choose the best fitting copula among the three considered copula families according to the highest  $p$ -value reported by the Pearson chi-square goodness-of-fit test. Details on this procedure can be found in [24], [34].

As compared to using the product of the marginal PDFs, the gain in using copulas is that they represent a generalization that enables modeling the dependence structure of any type. For this reason, the case of mutually independent variables (i.e., product of marginal PDFs) corresponds to the copula density equal to 1. Furthermore, the *meta-Gaussian* model developed in [17] can be obtained within the developed copula-based approach by employing the Gaussian copula. The copula contribution in the joint PDF modeling becomes especially relevant when the dependence between the source images is strong, e.g., for the different polarization channels of the same SAR acquisition, see [24].

### III. THE HIERARCHICAL MODEL

In order to integrate directly the input data in a hierarchical graph [25], and, thus, reflect the multi-resolution aspect of the acquisitions, we employ an explicit hierarchical graph-based model [27], [35] to address the considered classification problem.

In the following, we focus on specific simple and non-directed graphs, which have a quad-tree structure. The set of sites  $s \in S$  is, therefore, hierarchically partitioned on  $S = S^0 \cup S^1 \cup \dots \cup S^R$ , where  $R$  corresponds to the coarsest resolution, the root, and 0 corresponds to the reference level (finest resolution). In this structure, depicted in Fig. 1, parent-child relationships can be defined: for each site  $s$  of any tree-level  $n$  a unique parent  $s^-$  and four children  $s^+$  are well defined. The set of descendants of  $s$ , including itself, is denoted  $d(s)$ . The choice of a quad-tree structure relies on the good properties of this configuration such as the causality in scale, under Markovian assumption, which allows the use of a non-iterative algorithm with acceptable computational time [36]. Moreover, the multigrid properties reduce the chances to end up in a local minimum, hence the applied algorithm is more likely to converge to a global solution [27]. Finally, such a model is sufficiently flexible to be able to deal with different kinds of statistics, and, thus, to use different kinds of images (different resolutions, different sensors) [37]. It also enables to use a non-constant number of bands at each scale, which is critical when using distinct sensor acquisitions.

#### A. Multi-scale data modeling

As mentioned in the introduction, we use a multi-scale model to integrate the input multi-resolution data. To increase the amount of available information, and to potentially avoid “empty” tree levels, we integrate additional information by also employing a multi-scale decomposition of the original data. Specifically, if we consider, for instance, a 1-meter resolution optical image and a coregistered 4-meter resolution SAR acquisition, the finest resolution image is integrated at the bottom of the quad-tree, and the SAR image, at level  $R = 2$ . We, thus, decompose the finest resolution image along 2 levels (or more), and integrate the obtained images in the tree. At the bottom and at the first stage, we only have one observation, at the 4-meter resolution, we have two observations, the SAR image

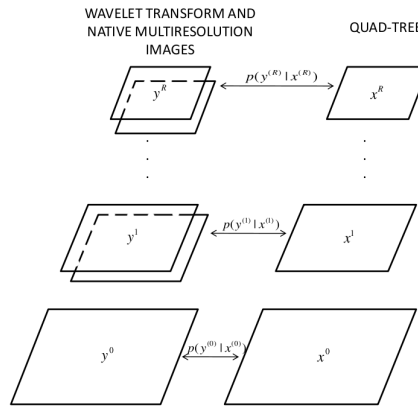


Fig. 2. Generic hierarchical graph-based model of an image quad-tree.

and the second-level decomposition of the optical one, that are combined by using the model presented in Section II. We can further increase the number of levels of the tree by decomposing both the SAR and the optical channels.

A natural image processing model to generate a hierarchical decomposition is the 2-D decimated discrete wavelet transform (DWT) [38]. We decompose the original image into multiple scales, corresponding to the levels of a tree. For each scale, we consider solely the approximation coefficients. The scale factor is always a power of 2, thus leading to a quad-tree configuration in which each pixel at scale  $j > 0$  owns four children, and each pixel at scale  $j < R$  exhibits one parent, where  $R$  is the depth of decomposition. The approximation coefficients at scale 0 correspond to the finest resolution original image. By filtering this image with a low-pass filter on the rows and columns, we obtain the approximation coefficients for scale 1. In the same way, the approximation coefficients at scale  $j$  are used to obtain the coefficients at scale  $j + 1$ , for  $j < R$  [38].

A wide choice of wavelets functions exists, and we have performed the decomposition of the original images using different wavelet families (see in [39]) such as Daubechies [40], orthogonal, bi-orthogonal, etc. [38], [41].

### B. Proposed MPM classification method

The aim of classification is to estimate a set of hidden labels  $X$  given a set of observations  $Y$  attached to the sites (pixels).  $X$  and  $Y$  are considered as random processes, and  $X$  is Markovian with respect to scale, i.e.  $p(x^n|x^k, k > n) = p(x^n|x^{n+1})$ , where  $n$  and  $k$  are scales and  $x^n$  is a realization of  $X$  at scale  $n$ .

Among the different classification algorithms employed on a quad-tree structure in the literature, two have been widely used. The first one aims to estimate exactly the maximum a posteriori (MAP) solution by using a linear Kalman filter [42] or a non-linear Viterbi algorithm [37], [43]. However, these criteria are 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) [27], [44]. The cost function associated with this estimator offers the possibility to penalize errors according to their number and the scale at which they occur: for example, an error at the coarsest scale is more strongly penalized than an error at the finest scale. This represents a desirable property because a site located at the root corresponds to  $4^R$  pixels at the finest scale.

The MPM algorithm runs in two passes, referred to as “bottom-up” and “top-down” passes, similar to the “forward-backward” algorithm applied to Markov chains [45]. A bottom-up pass estimates the partial posterior marginals  $p(x_s|y_{d(s)})$  by a recursion from the leaves to the root. A top-down pass successively estimates the labels at each tree level  $n$  by maximizing the posterior marginals  $p(x_s|y)$  from the root to the leaves. The  $p(x_s|y)$  maximization is done by employing a modified Metropolis Dynamics algorithm, that has good properties for both its relative low computation time and the good precision of its results [46].

In the proposed algorithm, we improve the robustness to noise [39] (e.g., speckle noise) by truncating the top-down pass and maximizing the posterior marginal solely at the root level (see Fig. 4). The obtained classification map is then used to update the prior probabilities at the top level (see below), whereas at other scales  $n \in [0; R - 1]$ ,

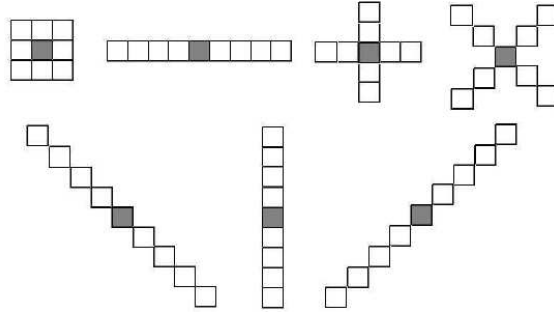
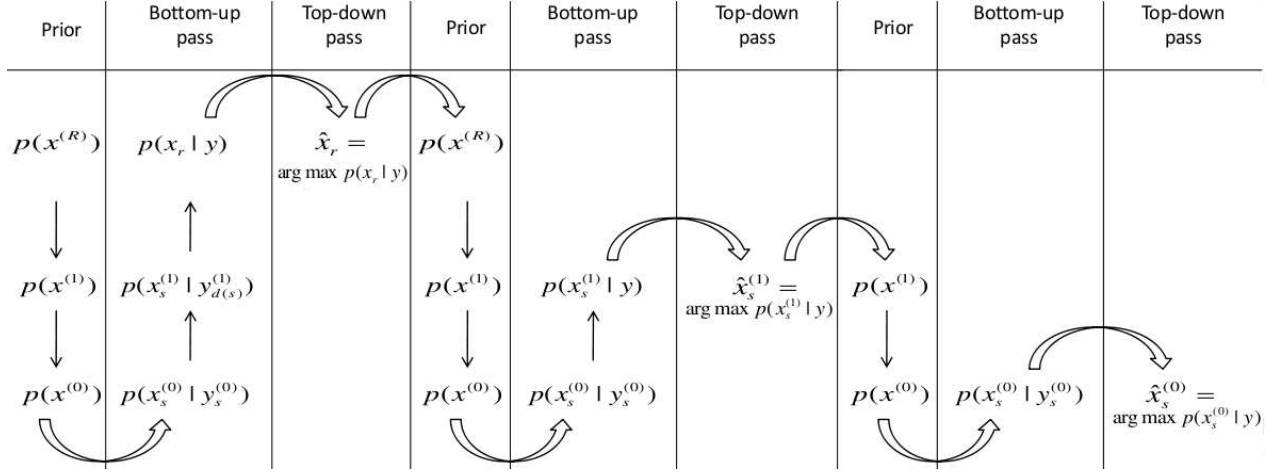


Fig. 3. The considered adaptive neighborhood structures.


 Fig. 4. Proposed marginal posterior mode (MPM) classifier on a quad-tree image pyramid. In this representation,  $R = 2$ .

the priors are estimated as:

$$p(x_s^n) = \sum_{x_{s^-}^n \in \Omega = \{\omega_i, i=1, \dots, M\}} p(x_s^n | x_{s^-}^n) p(x_{s^-}^n), \quad (8)$$

where  $p(x_s^n | x_{s^-}^n)$  are the transition probabilities, that are further studied below. We plug these new prior probabilities as updates in a novel “truncated” quad-tree of size  $R - 1$  to which we again apply the MPM algorithm (see Fig. 4). This procedure is repeated until scale 0 is reached.

For the bottom-up pass, Laferte et al. [27] showed that after the initialization at the bottom

$$p(x_s^{(0)} | y_s^{(0)}) \propto p(y_s^{(0)} | x_s^{(0)}) p(x_s^{(0)}), \quad (9)$$

the posterior updating at the remaining levels is performed recursively via

$$p(x_s | y_{d(s)}) \propto p(y_s | x_s) p(x_s) \prod_{t \in s^+} \sum_{x_t} \left[ \frac{p(x_t | y_{d(t)})}{p(x_t)} p(x_t | x_s) \right]. \quad (10)$$

Thus, to estimate the posterior probabilities, we need to determine for each site  $s$  of the quad-tree the following quantities: the likelihood  $p(y_s | x_s)$  (defined in Section II), the prior probability  $p(x_s)$ , and the transition probability  $p(x_t | x_s)$  for  $t \in s^+$  (i.e.,  $p(x_t | x_{t^-})$ ).

### C. Transition probabilities

The transition probabilities between scales  $p(x_s | x_{s^-})$ , determine the hierarchical MRF since they formalize the causality of the statistical interactions between the different levels of the tree. As such, they need to be carefully



defined. We use the transition probability in the form introduced by Bouman et al. [43]. Specifically, for all sites  $s \in S$  and all scales  $n \in [0; R - 1]$ ,

$$p(x_s^n = \omega_m | x_{s^-}^n = \omega_k) = \begin{cases} \theta_n, & \text{if } \omega_m = \omega_k \\ \frac{1-\theta_n}{M-1}, & \text{otherwise} \end{cases}, \quad (11)$$

with the parameter  $\theta_n > 1/M$  and where  $\omega_m$  and  $\omega_k$  represent the  $m$ -th and  $k$ -th classes,  $m, k \in [1; M]$ , respectively, and  $M$  is the number of considered classes. This model favors an identical parent-child labeling.

#### D. From classification maps to prior probabilities

To estimate the priors given a classification map, we use a Markovian model which takes into account the contextual information at each scale  $n$ , and therefore, leads to a better prior estimation. By employing the Hammersley-Clifford theorem [47], we can define a local characteristic for each site as a well-known Potts MRF model:

$$p(x_s | x_t, t \in S^n, t \sim s) \propto \exp(-\beta \sum_{s \sim t} I[x_s = x_t]) \quad (12)$$

with  $I[\cdot]$  the indicator function previously defined.  $s \sim t$  denotes that sites  $s$  and  $t$  ( $s, t \in S^n$ ) are neighbors with respect to the neighborhood system on  $S^n$ .

In the proposed method, to estimate the priors given the current classification map, we replace the prior probability  $P(x_s)$  by the local conditional prior probability in Eq. (12). This choice generally provides a biased prior-probability estimate, but favours spatial adaptivity, a desired property when working with high resolution images where spatial details are common. Furthermore, for the same reason, instead of considering a commonly used second-order neighborhood based on the 8 pixels surrounding the considered pixel [48], we suggest to use of pairwise cliques in an adaptive neighborhood, which involves the use of oriented neighbor configurations, see Fig. 3, as in [39], [49], and at each site we select the one that leads to the smallest energy [39], [49], [50]. Such adaptivity aims to take into account the geometrical properties of areas in the original image. Such spatial selectivity is very important, especially in high-resolution imagery, and is an appropriate tool that helps improving the classification results in geometrically rich areas such as urban.

In Eq. (12), we notice the presence of a parameter  $\beta$  to estimate. This parameter can be found by minimizing the pseudo-likelihood function over the training set, by resorting to a simulated annealing algorithm [48]. However, we stress that this method brings to adequate estimates only when using an exhaustive ground truth, or, at least, when the training map contains a sufficient amount of class borders, which is rare in remote sensing. For this reason, we determine this parameter empirically by a trial-and-error procedure.

## IV. OVERVIEW OF THE CLASSIFICATION METHOD

The proposed hierarchical classification consists of a learning and a classification stages. A step-by-step presentation of the approach is given in Algorithm 1.

The developed hierarchical classifier is applied on an image pyramid and, thus, the algorithm begins with input data decomposition into a multi-resolution representation (line 1). First, the finest resolution of the pyramid (level 0) is set equal to the finest resolution of the input images. Then, given an a priori value of decomposition depth  $R$ , each input image is downsampled to the closest pyramid level  $R_{img}$  and its wavelet decompositions (approximations) obtained at all coarser levels  $(R_{img} + 1), \dots, R$ .

The learning step consists of estimating the class-conditional joint PDFs. First (lines 2-3), given a training set, at each scale of the tree  $n$ , for each input channel  $j$  and class  $m$ , the marginal PDFs are estimated as mixtures of appropriate types, see Section II-A. Then (lines 4-7), the joint class-conditional PDFs are estimated by using the copula approach, see Section II-B.

On the classification stage, the pyramid-decomposed test images (line 8) are handled by a MPM hierarchical classifier (lines 9-20), see Section III-B, graphically presented in Fig. 4.

## V. EXPERIMENTAL RESULTS

In this section, we discuss the results of the experimental validation of the developed model on two datasets: a multi-resolution SAR acquisition of the city of Amiens (France), and a multi-sensor acquisition of Port-au-Prince (Haiti). The proposed approach is compared to other methods, detailed below. For the considered methods and

**Algorithm 1:** Proposed hierarchical classifier

---

PART I: Learning

- 1 perform DWT decomposition of the learning dataset;
  - for all images in pyramid and all classes  $m$  do**
    - estimate marginal densities as mixtures (1) of
    - 2 **if optical data then** Gaussian components (2);
    - 3 **if SAR data then** generalized gamma components (3);
  - for all pyramid levels and all classes  $m$  do**
    - estimate class densities (*likelihoods*), i.e.:
    - if multiple channels are present ( $d \geq 2$ ) then**
      - 4 estimate Kendall's tau via (7);
      - 5 select the best fitting copula, see Section II-B;
      - 6 merge marginals via (5);
    - 7 **else** set equal to marginal densities.

PART II: Hierarchical classification

- 8 perform DWT decomposition of the test dataset;
- 9 obtain maximum likelihood classification with the likelihoods at level  $R$ ;
- 10 set a MRF-based prior  $p(x^R)$  via (12);
  - for levels  $0, \dots, (R-1)$  do**
    - 11 set transition probabilities (11);
- 12 set  $R_{cur} = R$ ;
  - while  $R_{cur} > 0$  do**
    - 13 recursive top-down pass at levels  $(R_{cur} - 1) \rightarrow 0$  to update priors via (8);
    - 14 initialize the posterior at level 0 via (9);
    - 15 recursive bottom-up pass at levels  $1 \rightarrow R_{cur}$  to update posteriors via (10);
    - 16 obtain classification at level  $R_{cur}$  by maximizing the posterior at level  $R_{cur}$ ;
    - 17 estimate prior  $p(x^{R_{cur}-1})$  via (8);
    - 18 set  $R_{cur} = R_{cur} - 1$ ;
- 19 obtain posterior at level 0 via (9);
- 20 obtain final classification by maximizing the posterior at level 0.

---

datasets, we present the final classification map and the corresponding classification accuracies, obtained on test sets not overlapping with the training sets. In the considered supervised context, the number of classes is determined by the user. The training and testing is performed on the same sets of images by using the non-overlapping subsets of the available ground truth.

Empirically, we perform the experiments with the following parameters:  $\beta = 4.8$  in (12), and  $\theta_n = 0.8$  in (11), which means that a site  $s$  at scale  $n$  has a probability of 80% to belong to the same class as its ascendant  $s^-$ . We could also notice that the highest overall accuracy is reached by using Daubechies 10 wavelet for SAR images, and Haar wavelets for optical images. For this reason, in the following we employ these wavelet types.

#### A. Amiens dataset

We consider two single-polarization COSMO-SkyMed (CSK) SAR images of the city of Amiens (France) (©ASI, 2011):

- a StripMap acquisition (2.5 m pixel spacing), HH polarized, geocoded, single-look,  $510 \times 1200$  pixels, shown in Fig. 5(a).
- a PingPong acquisition (5 m pixel spacing), HH polarized, geocoded, single-look,  $255 \times 600$  pixels, Fig. 5(b).

We consider four classes: urban, water, low vegetation and trees. The available ground truth is presented in Fig. 5(d).

We compare the classification results of the following methods:

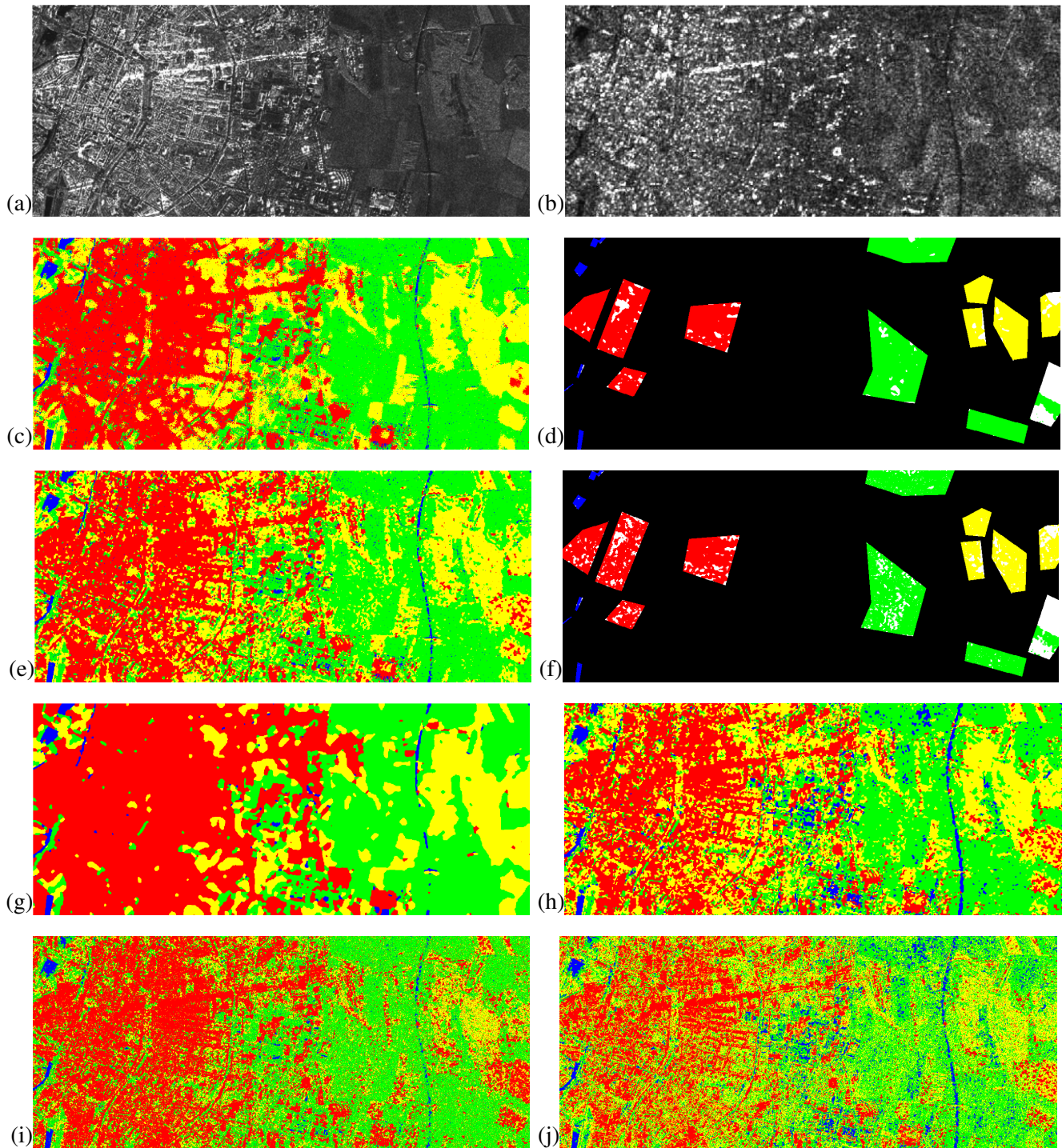


Fig. 5. Amiens dataset: (a) StripMap SAR and (b) Ping-Pong SAR acquisitions (CSK, ©ASI, 2011), and classification results obtained with multi-resolution methods: (c) the proposed approach, (e) the Laferte's method with (d) and (f) their respective projections on the available ground truth. Single-scale classification results obtained after rescaling to a common resolution: (g) the single-scale MRF-based method, (h) multi-sensor Storvik's method with MRF regularization, (i) GML with MRF and (j) grouped k-means. Legend: urban (red), water (blue), low vegetation (green) and trees (yellow); on (d),(f) black pixels represent absent ground truth, white - misclassification.

1. The proposed hierarchical MRF-based approach, see Fig. 5(c). The two initial multi-resolution acquisitions are integrated in the tree, and decomposed by discrete wavelet transform (see Section III-A) in order to work on a quad-tree of size  $R = 2$ . Such a decomposition aims to improve the classification by integrating additional scale information in the quad-tree. The PDFs of the SAR images and of their decompositions are independently

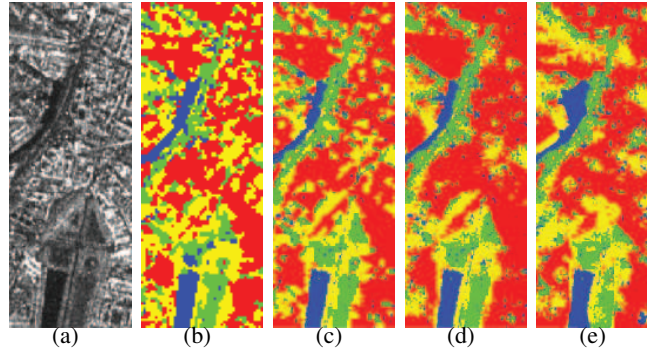


Fig. 6. Zoom of the lower left corner of the Amiens images: (a) StripMap SAR, (b) Storvik's model with MRF, (c) the Laferte's method, (d) the proposed method, and (e) its modification with independence copula.

TABLE I  
CLASSIFICATION ACCURACIES FOR THE AMIENS DATASET

Amiens, France					
	water	urban	vegetation	tree	overall
Proposed model $R = 2$	96.99%	92.41%	84.91%	91.21%	<b>91.38%</b>
excl. DWT	96.62%	91.10%	85.16%	91.13%	<b>91.00%</b>
excl. mixtures	97.12%	83.66%	82.17%	93.07%	<b>88.98%</b>
with independence	97.05%	89.40%	84.01%	89.32%	<b>89.94%</b>
with isotropic MRF	97.10%	92.25%	84.56%	90.92%	<b>91.20%</b>
excl. multi-pass	97.08%	87.71%	78.20%	88.42%	<b>87.85%</b>
Proposed model $R = 1$	96.78%	88.77%	81.63%	87.15%	<b>88.59%</b>
Laferte's model	96.52%	84.06%	79.13%	85.55%	<b>86.32%</b>
Single-scale MRF	98.31%	99.07%	83.95%	92.88%	<b>93.55%</b>
Storvik's model + MRF	93.54%	78.24%	84.12%	78.75%	<b>83.69%</b>
GML + MRF	92.72%	77.20%	79.88%	79.15%	<b>82.26%</b>
K-means	75.09%	62.53%	62.45%	59.90%	<b>65.02%</b>

modeled by using generalized gamma mixtures (see Section II-A2), and then combined via multivariate copulas.

2. The hierarchical MRF-based Laferte's approach [27], see Fig. 5(e), in which the two considered acquisitions are plugged, without the additional information derived from the wavelet decomposition.
3. The single-scale isotropic 3-by-3 Potts-model MRF with optimization by the modified Metropolis Dynamics algorithm [46], see Fig. 5(g), applied to the finest resolution image (StripMap) with the MRF-weight parameter  $\beta = 3.60$  set to maximize the overall classification accuracy on the available ground truth data. The likelihood term is modeled by using generalized gamma mixtures (see Section II-A2).
4. The single-scale multi-sensor Storvik's approach [17], see Fig. 5(h), where the StripMap image is downsampled and the likelihood term is constructed by merging the two generalized gamma marginals into a meta-Gaussian distribution [17]. The classification is obtained by maximum likelihood with an isotropic 3-by-3 Potts-model MRF with optimization by ICM [51] with  $\beta = 1.70$ .
5. The Gaussian maximum likelihood (GML) [52], see Fig. 5(i), where the coarser image is upsampled. The classification is obtained by maximum likelihood with an isotropic 3-by-3 Potts-model MRF with optimization by ICM [51], with  $\beta = 1.48$  estimated by the Ho-Kashyap method [53].
6. The single-scale grouped k-means approach [52], see Fig. 5(j), where the coarser image is upsampled. The detected clusters are associated with classes and the optimal number of clusters is determined ( $k = 20$ ) via a maximum likelihood criterion, see [54].

Visually, the proposed hierarchical method leads to satisfying results, and is a good compromise between the noise-affected map of the Laferte's hierarchical MRF model, see Fig. 5(e), and the spatially smoothed classification map resulting from the single-scale MRF-based method, see Fig. 5(g). In Table I, we present the quantitative results obtained by all methods. We observe that the developed method reports the highest overall accuracy with the exception of that obtained with the single-scale MRF-based approach. Visual inspection shows that this result is obtained at the expense of a strong smearing effect on object borders that is demonstrated by the single-scale MRF method. This does not decrease the classification accuracy in Table I because the reference ground truth does not include transition

zones, see Fig. 5(d). Moreover, the hierarchical approach is visually preferable for the urban areas classification since it allows to extract more details than the single-scale model. The available ground truth does not take into account details in urban areas (such as parks, trees, etc), hence the hierarchical classification results may be considered erroneous, as compared with the ground truth, whereas they might actually better describe the observed area.

The comparison with the Storvik's method, see Fig. 5(h), suggests the expected improvement that can be obtained by operating at multiple scales. Note that the construction of the likelihood terms in this approach can also be considered as copula-based, since the resulting meta-Gaussian distribution is similar to the Gaussian-copula based reconstruction from the marginals [17], except for parameter estimation routines. Moreover, although the multivariate Gaussian structure is adequate for the optical imagery scenario, it becomes inaccurate on radar data, where more appropriate dependence structures can be modeled with copulas. The GML with MRF and k-means performed worse due to only two images available and their results were strongly affected by speckle.

In Table I, we further report a panel of results obtained with the developed technique by switching off its methodological components one at a time or replacing them with standard alternatives. These comparisons serve to demonstrate what kind of improvement each of the algorithmical components is able to bring forth. Specifically, we consider: 1) the proposed model with DWT decomposition replaced by direct subsampling; 2) the proposed model with mixtures replaced by single distribution estimates or, in other words, by setting the number of mixture components equal to one; 3) the proposed model under channel independence assumption, i.e., with an independence copula or, simply, a product of marginals; 4) the proposed model with isotropic 3-by-3 Potts-model MRF instead of adaptive neighborhoods; and, finally, 5) the proposed model with no truncated paths. Mind that the last setting allows a more direct comparison with the Laferte's model by highlighting the contribution of the copula-based likelihoods. We further compare the accuracies obtained by the proposed method based on  $R = 2$  pyramid size, with the same method based on  $R = 1$ . This allows to highlight the accuracy improvement linked to the introduction of higher-level quad-tree levels. The consideration of these results shows a negligible quantitative improvement brought by the introduction of DWT, as compared to simple subsampling, and adaptive neighborhoods in the MRF structure. While the former is consistent with our experience of a limited improvement reported by the use of DWT-based subsampling, the latter can be related to the coarse disjoint ground truth employed. Indeed, the adaptive structure of neighborhoods has been explicitly designed to improve classification in geometrically rich areas, such as urban [49]. The next dataset (next Section) allows to better visually appreciate the impact of this component. The other algorithm components brought to an overall accuracy boost of 2 – 3%. We consider these observations as an experimental justification of the structure of the algorithm with, perhaps, exception for the wavelet component, whose impact is limited.

In order to appreciate the spatial details of the classification, we separately present a zoom in the classification maps, see Fig. 6. Note that all the results are at the finest resolution except that of the Storvik's approach. It is immediate that, although the proposed method brings to a higher level of salt and pepper noise (which can be dealt with by morphological manipulations [52], e.g., dilate-erode operation), the classification is less fragmented than both Laferte's and Storvik's. We also demonstrate the result with the independence copula (see its numerical performance in Table I): it shows that the class-conditional independence assumption, which is inaccurate in multiple SAR channels case, brings to a big amount of urban-tree confusion (on the right side). The visual inspection of the image suggests that the results reported by the independence-copula are less accurate and a lot of urban area is misclassified. More direct comparisons of the copula-based modeling improvement have been performed in [24] in the mono-scale context.

The loss of accuracy in the vegetation class for all the methods is due to confusion of this class with the tree class at the bottom of the image. By visual inspection of the SAR image, it is very hard to detect a difference between this vegetation area and a tree area from radar images. The ground truth, on the other hand, was constructed by interpretation of optical images. If, instead of using a single HH-band of the PingPong acquisition, we employ both the HH and VV bands, we do not obtain at any appreciable improvement. This is due to the fact that these two bands contain highly correlated observations, and the considered copulas intrinsically assume the same kind and degree of dependence between any two channels. If multi-source and multi-band images are processed at the same time, hierarchical copula-structures [23] are preferable.

Experiments were conducted on an Intel Xeon quad-core (2.40 GHz and 18-GB-RAM) Linux system. The proposed algorithm ran in around 150 sec, which is comparable to the hierarchical Laferte's method computation time of 120 sec and led to a visual refinement of the classification map. The single-scale MRF method took 50 sec, Storvik's method with MRF 20 sec, GML with MRF 20 sec and k-means 10 sec. The last three are fast due

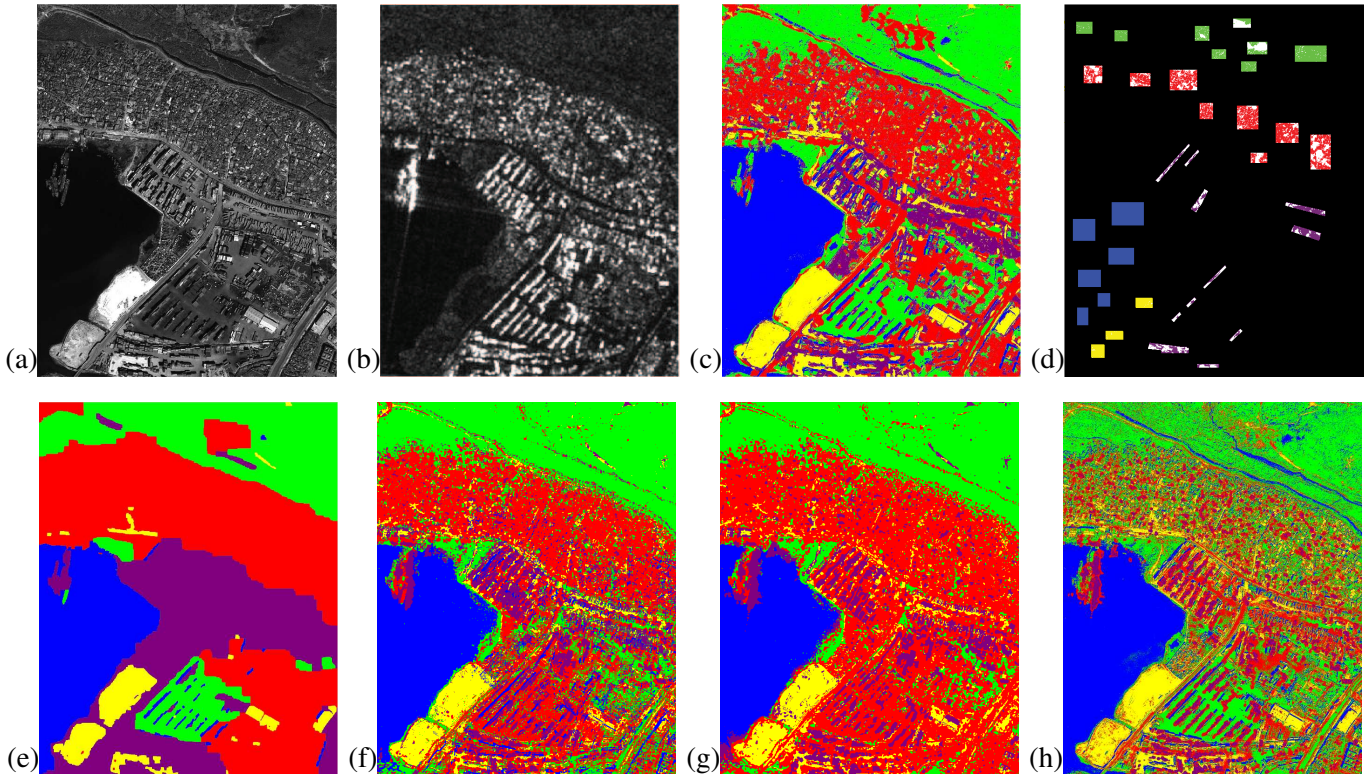


Fig. 7. Port-au-Prince dataset: (a) optical (©GeoEye, 2010) and (b) SAR images (CSK, ©ASI, 2010), and classification results obtained with: (c) the proposed method applied to the multi-sensor dataset with (d) its projection on the available ground truth, (e) the single-scale MRF applied to the optical image, (f) the multi-sensor Storvik's model, (g) GML with MRF, (h) grouped k-means. Legend: urban (red), water (blue), low vegetation (green), sand (yellow), containers (purple); on (d) black pixels represent absent ground truth, white - misclassification.

to their mono-scale structure and the MRF optimization (for the first two of the three) performed with a fast ICM method. Note that a proper parallelization and optimization of the algorithms can reduce their computation times.

### B. Port-au-Prince dataset

This multi-sensor dataset consists of two images of the port area of Port-au-Prince (Haiti):

- a pan-sharpened GeoEye acquisition (©GeoEye, 2010),  $1280 \times 1600$  pixels, shown in Fig. 7(a).
- a single-polarization COSMO-SkyMed SAR image (©ASI, 2010), HH polarization, StripMap acquisition mode (2.5 m pixel spacing), geocoded, single-look,  $320 \times 400$  pixels, shown in Fig. 7(b).

Five classes are considered: urban, water, low vegetation, sand and containers. The available ground truth is presented in Fig. 7(d).

The GeoEye image resolution comes at the 0.5 meters resolution. To fit with the dyadic decomposition imposed by the quad-tree, we slightly resampled the data to obtain the  $0.625 = 2.5/4$  m resolution. The impact of downsampling [55] from 0.5 m to 0.625 m is expected to have a minor impact on the resulting map, because the resampling ratio is quite close to unity.

We compare classification results obtained with the following methods:

1. The proposed hierarchical MRF-based approach, see Fig. 7(c). The two acquisitions are decomposed over a quad-tree with  $R = 3$  scales, by resorting to discrete wavelet transforms.
2. The single-scale isotropic 3-by-3 Potts-model MRF with optimization by the modified Metropolis Dynamics algorithm [46], see Fig. 7(e), applied only to the optical image with the MRF-weight parameter  $\beta = 3.7$  set to maximize the overall classification accuracy on the available ground truth data. The likelihood is modeled by using Gaussian mixtures (see Section II-A1).
3. The multi-sensor single-scale Storvik's approach [17], see Fig. 7(f), where the images are rescaled to the SAR resolution and the likelihood term is constructed by merging generalized gamma (for SAR) and Gaussian (for

TABLE II  
CLASSIFICATION ACCURACIES FOR THE PORT-AU-PRINCE DATASET

	Port-au-Prince, Haiti					
	water	urban	vegetation	sand	containers	overall
Proposed SAR+opt.	100%	75.24%	87.16%	98.89%	49.31%	<b>82.12%</b>
Proposed opt.	100%	67.12%	86.89%	98.83%	41.90%	<b>78.95%</b>
Single-scale MRF	100%	100%	81.42%	99.94%	59.62%	<b>88.20%</b>
Storvik's model	99.95%	79.32%	90.81%	96.22%	37.25%	<b>79.44%</b>
GML + MRF	99.99%	84.51%	93.24%	99.68%	58.64%	<b>86.28%</b>
K-means	100%	39.01%	91.19%	82.32%	6.54%	<b>57.30%</b>

- optical) marginals into a meta-Gaussian distribution [17]. The classification is obtained by maximum likelihood.
4. The Gaussian maximum likelihood (GML) [52], see Fig. 7(g), on four images: upscaled SAR and optical R, G, and B channels. The classification is obtained by maximum likelihood with an isotropic 3-by-3 Potts-model MRF with optimization by ICM [51], with  $\beta = 1.10$  estimated by the Ho-Kashyap method [53].
  5. The single-scale grouped k-means approach [52], see Fig. 7(h), where the SAR image is upscaled. The detected clusters are associated with classes and the optimal number of clusters is determined ( $k = 10$ ) via a maximum likelihood criterion, see [54].

To the best of our knowledge, none of the previously developed application-specific methods (note, of course, the existence of generic approaches, such as k-means, maximum likelihood, etc.) allows a direct integration of both multi-resolution and multi-source data without major pre- or postprocessing.

The results obtained by the proposed hierarchical MPM-based method, see Fig. 7(c), lead to a detailed classification with an adequate level of classification map regularity. The main source of misclassification is the container area, where the asphalt is classified as vegetation. In Table II we compare numerically the results obtained with the proposed hierarchical method when considering either only optical, or both SAR and optical images. We observe an improvement related to the combination of the two images, in particular in the urban areas for which the SAR acquisition represents a significant source of discriminative information. More specifically, we have observed that the optical image has a relevant effect in the sand discrimination, and the SAR acquisition is very helpful to detect the containers.

Numerical results suggest that the single-scale MRF-based method, see Fig. 7(e), leads to the highest accuracy. However, as with the previous dataset, this comes at the price of a severe oversmoothing. This is a result of selecting the  $\beta$  parameter maximizing the classification accuracies. Consequently, the localization of the ground truth within homogeneous regions brought to an excessive oversmoothing of the edges. A lower value of  $\beta$  brings to a lower level of smoothing at the price of an accuracy decrease. Indeed, the proposed method classifies more accurately the border areas, which, however, does not impact the final numerical accuracies. The second best result is demonstrated by the GML+MRF method. The visual inspection suggests that the resulting classification map, see Fig. 7(g), is highly fragmented, especially as compared with the map reported by the proposed method. A closer look into the container area (lower-right part of the images) reveals a better structured and (visually) more accurate classification with a spatially more precise characterization of the geometrical structure reported by the proposed approach thanks to the use of adaptive neighborhoods. The Storvik's and k-means approaches perform considerably worse with the container class.

## VI. CONCLUSION

The novel method proposed in this paper allows to perform classification on multi-sensor, multi-band, and/or multi-resolution imagery. It combines a joint statistical modeling of considered input images (optical and/or SAR imagery), with a hierarchical Markov random field, leading to a statistical supervised classification approach. The central advantage of the proposed method is the applicability to a very general setting in terms of admissible input data. More specifically, we have proposed an explicit copula-based multivariate statistical model that enables to fuse multi-sensor acquisitions. We developed a novel MPM-based hierarchical Markov random field model that iteratively updates the prior probabilities and, thus, leads to the improved robustness of the classifier. The hierarchical MRF developed in this paper has two advantages: it is robust to speckle noise, and employs a non-iterative optimization algorithm (MPM estimation). The flexible number of input images at each scale allows to integrate additional information that may not be available at each tree level. This may occur when processing SAR and optical acquisitions acquired at different resolutions. To our best knowledge, so far, no classification methods that preserve both multi-resolution and multi-source aspects of the acquisitions were proposed in the literature.

We performed an experimental validation of the proposed method and compared it to various other methods, including a basic MPM-based hierarchical approach, meta-Gaussian technique for multi-source imagery classification, single-scale MRF model, and several standard classifiers. According to the quantitative (classification accuracy on test samples) and qualitative (visual photointerpretation of the maps) analysis of the classification maps, we may consider that the developed model outperforms the considered benchmark hierarchical and multi-source techniques. The multi-scale Markovian model allows to better keep details when compared to the smoothing effects generated by the single-scale MRF-based method. The iterative prior update increases the robustness to speckle noise and the regularity of the final map.

The consideration of a quad-tree tends to somewhat restrict the applicability of the developed multi-resolution approach by requiring a dyadic decomposition and image rescaling. A possible further improvement is, therefore, to find a new hierarchical Markovian-based algorithm that may overcome such limits. For this purpose, strategies based on irregular hierarchical trees (not necessarily aligned with the quad-tree topology) could be integrated within the proposed method. Further experimental validation with data from recent missions such as Pléiades (operated by CNES) would also be a relevant extension of the present work.

#### ACKNOWLEDGMENT

The authors would like to thank the Italian Space Agency for providing the COSMO-SkyMed (CSK®) images (COSMO-SkyMed Product - ©ASI - Agenzia Spaziale Italiana. All Rights Reserved) in the framework of the project “Development and validation of multitemporal image analysis methodologies for multirisk monitoring of critical structures and infrastructures (2010-2012)”. The authors would also like to thank GeoEye Inc. and Google crisis response for providing the GeoEye image available on the website <http://www.google.com/relief/haitiearth> and Dr. Michaela De Martino (University of Genoa) for her help with data preparation. Finally, the authors thank the associate editor and the anonymous reviewers for their helpful and constructive comments.

#### REFERENCES

- [1] A. A. Goshtasby, *2-D and 3-D image registration: for medical, remote sensing, and industrial applications*. Wiley-Interscience, 2005.
- [2] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar images*. SciTech Publishing, 2004.
- [3] L. Xu, A. Krzyzak, and C. Y. Suen, “Methods of combining multiple classifiers and their applications to handwriting recognition,” *IEEE Trans. Syst., Man, Cybern.*, vol. 22, no. 3, pp. 418–435, 1992.
- [4] J. A. Benediktsson, P. H. Swain, and O. K. Ersoy, “Neural network approaches versus statistical methods in classification of multisource remote sensing data,” *IEEE Trans. Geosci. Remote Sens.*, vol. 28, no. 4, pp. 540–552, 1990.
- [5] S. Foucher, M. Germain, J.-M. Boucher, and G. B. Benie, “Multisource classification using ICM and Dempster-Shafer theory,” *IEEE Trans. on Instrumentation and Measurement*, vol. 51, no. 2, pp. 277–281, 2002.
- [6] D. H. Wolpert, “Stacked generalization,” *Neural Networks*, vol. 5, no. 2, pp. 241–259, 1992.
- [7] M. Fauvel, J. Chanussot, and J. A. Benediktsson, “Decision fusion for the classification of urban remote sensing images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2828–2838, 2006.
- [8] G. J. Briem, J. A. Benediktsson, and J. R. Sveinsson, “Multiple classifiers applied to multisource remote sensing data,” *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 10, pp. 2291–2299, 2002.
- [9] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas, “On combining classifiers,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 3, pp. 226–239, 1998.
- [10] R. Gaetano, G. Moser, G. Scarpa, and S. B. Serpico, “Region-based classification of multisensor optical-SAR images,” in *IEEE International Geoscience and Remote Sensing Symposium (IGARSS’08)*, vol. 4, 2008, pp. 81–84.
- [11] M. Datcu, F. Melgani, A. Piardi, and S. B. Serpico, “Multisource data classification with dependence trees,” *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 3, pp. 609–617, 2002.
- [12] D. B. Hester and H. I. Cakir and A. C. Nelson and S. Khorram, “Per-pixel classification of high spatial resolution satellite imagery for urban land-cover mapping,” *Photogramm. Eng. Rem. Sensing*, vol. 74, no. 4, pp. 463–471, 2008.
- [13] F. Yuan, “Land-cover change and environmental impact analysis in the greater Mankato area of Minnesota using remote sensing and GIS modelling,” *International Journal of Remote Sensing*, vol. 29, no. 4, pp. 1169–1184, 2008.
- [14] P. M. Mather and M. Koch, *Computer Processing of Remotely-Sensed Images: An Introduction*, 4th ed. Wiley, 2011.
- [15] D. Amarsaikhan and T. Douglas, “Data fusion and multisource image classification,” *International Journal of Remote Sensing*, vol. 25, no. 17, pp. 3529–3539, 2004.
- [16] M. R. Metwalli, A. H. . Nasr, O. S. F. Allah, S. El-Rahaie, and F. E. A. El-Samie, “Satellite image fusion based on principal component analysis and high-pass filtering,” *Journal of the Optical Society of America*, vol. 27, no. 6, pp. 1385–1394, 2010.
- [17] B. Storvik and G. Storvik and R. Fjortoft, “On the combination of multisensor data using meta-Gaussian distributions,” *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 7, pp. 2372–2379, 2009.
- [18] B. Zhukov and D. Oertel and F. Lanzl and G. Reinhackel, “Unmixing-based multisensor multiresolution image fusion,” *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1212–1226, 1999.



- [19] 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.*, vol. 8, no. 1, pp. 148–152, 2011.
- [20] G. Celeux, D. Chauveau, and J. Diebolt, "Stochastic versions of the EM algorithm: an experimental study in the mixture case," *Journal of Statistical Computation and Simulation*, vol. 55, no. 4, pp. 287–314, 1996.
- [21] H.-C. Li, W. Hong, Y.-R. Wu, and P.-Z. Fan, "On the empirical-statistical modeling of SAR images with generalized gamma distribution," *IEEE J. Sel. Top. Signal Process.*, vol. 5, no. 3, pp. 386–397, 2011.
- [22] 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.*, vol. 42, no. 10, pp. 2046–2057, 2004.
- [23] R. B. Nelsen, *An introduction to copulas*, 2nd ed. New York: Springer, 2006.
- [24] V. Krylov, G. Moser, S. B. Serpico, and J. Zerubia, "Supervised high resolution dual polarization SAR image classification by finite mixtures and copulas," *IEEE J. Sel. Top. Signal Process.*, vol. 5, no. 3, pp. 554–566, 2011.
- [25] C. Graffigne, F. Heitz, P. Perez, M. Sigelle, and J. Zerubia, "Hierarchical Markov random field models applied to image analysis: a review," in *Proc. SPIE Conf. on Neural, Morphological and Stochastic Methods in Image Proc.*, vol. 2568, 1995, pp. 2–17.
- [26] P. Fieguth, *Statistical image processing and multidimensional modeling*. Springer, 2011.
- [27] J.-M. Laferte, P. Perez, and F. Heitz, "Discrete Markov modeling and inference on the quad-tree," *IEEE Trans. Image Process.*, vol. 9, no. 3, pp. 390–404, 2000.
- [28] M. A. T. Figueiredo and A. K. Jain, "Unsupervised learning of finite mixture models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 3, pp. 381–396, 2002.
- [29] I. Sneddon, *The use of integral transforms*. New York: McGraw-Hill, 1972.
- [30] V. A. Krylov, G. Moser, S. B. Serpico, and J. Zerubia, "On the method of logarithmic cumulants for parametric probability density function estimation," *IEEE Trans. Image Process.*, 2013, accepted. DOI: 10.1109/TIP.2013.2262285.
- [31] D. Huard, G. Evin, and A.-C. Fabre, "Bayesian copula selection," *Computational statistics and data analysis*, vol. 51, no. 2, pp. 809–822, 2006.
- [32] G. Mercier, G. Moser, and S. B. Serpico, "Conditional copulas for change detection in heterogeneous remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 5, pp. 1428–1441, 2008.
- [33] I. Kojadinovic and J. Yan, "Comparison of three semiparametric methods for estimating dependence parameters in copula models," *Insurance: Mathematics and Economics*, vol. 47, no. 1, pp. 52–63, 2010.
- [34] E. Lehmann and J. Romano, *Testing statistical hypotheses*, 3rd ed. Springer, 2007.
- [35] 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*, vol. 58, no. 1, pp. 18–37, 1996.
- [36] E. Fabre, "New fast smoothers for multiscale systems," *IEEE Trans. Signal Process.*, vol. 44, no. 8, pp. 1893–1911, 1996.
- [37] 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)*, Cambridge, USA, Jun. 1995, pp. 908–913.
- [38] S. G. Mallat, *A wavelet tour of signal processing*, 3rd ed. Academic Press, 2008.
- [39] A. Voisin, V. Krylov, G. Moser, S. B. Serpico, and J. Zerubia, "Classification of very high resolution SAR images of urban areas," INRIA, France, Research report 7758, oct 2011.
- [40] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Communications on Pure and Applied Mathematics*, vol. 41, no. 7, pp. 909–996, 1988.
- [41] M. Vetterli, J. Kovacevic, and V. K. Goyal, *Fourier and wavelet signal processing*. Cambridge University Press, 2013, <http://www.fourierandwavelets.org>.
- [42] M. R. Luetggen, W. Karl, and A. S. Willsky, "Efficient multiscale regularization with applications to the computation of optical flow," *IEEE Trans. Image Process.*, vol. 3, no. 1, pp. 41–64, 1994.
- [43] C. Bouman and M. Shapiro, "A multiscale random field model for Bayesian image segmentation," *IEEE Trans. Image Process.*, vol. 3, no. 2, pp. 162–177, 1994.
- [44] J. Marroquin, S. Mitter, and T. Poggio, "Probabilistic solution of ill-posed problems in computational vision," *Journal of the American Statistical Association*, vol. 82, no. 397, pp. 76–89, 1987.
- [45] L. E. Baum, T. Petrie, G. Soules, and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains," *IEEE Ann. Math. Stats*, vol. 41, no. 1, pp. 164–171, 1970.
- [46] M. Berthod, Z. Kato, S. Yu, and J. Zerubia, "Bayesian image classification using Markov random fields," *Image and Vision Computing*, vol. 14, no. 4, pp. 285–295, 1996.
- [47] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *J. Royal Stat. Soc. Ser. B*, vol. 36, no. 2, pp. 192–236, 1974.
- [48] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6, no. 6, pp. 721–741, 1984.
- [49] 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.*, vol. 35, no. 4, pp. 844–857, 1997.
- [50] 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*, Dec. 2007, pp. 351–355.
- [51] J. Besag, "On the statistical analysis of dirty pictures," *J. Royal Stat. Soc. Ser. B*, vol. 48, no. 3, pp. 259–302, 1986.
- [52] C. M. Bishop, *Pattern Recognition and Machine Learning*. New-York: Springer-Verlag, 2006.
- [53] S. B. Serpico and G. Moser, "Weight parameter optimization by the Ho-Kashyap algorithm in MRF models for supervised image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 12, pp. 3695–3705, 2006.
- [54] G. Moser, S. B. Serpico, M. De Martino, and D. Coppolino, "Automatic partially supervised classification of multitemporal remotely sensed images," in *Proc. of SPIE*, L. Bruzzone, Ed., vol. 5573, Nov. 2004, pp. 126–137.
- [55] A. K. Jain, *Fundamentals of digital image processing*. Prentice Hall, 1988.



**Aurélie Voisin** obtained an engineering degree and master's degree in signal and image processing both from Ecole Nationale Supérieure d'Electrotechnique, d'Electronique, d'Informatique, d'Hydraulique et des Telecommunications (ENSEEIH) in Toulouse, France, in 2009. She pursued her Ph.D. degree in signal and image processing at INRIA Sophia Antipolis, and defended in 2012. She is currently working as an engineer in the French satellite industry.

Her main interests include SAR and optical image processing and Markov random field techniques.



**Vladimir A. Krylov** received the "specialist" (M.Sc.) degree in applied mathematics and computer science, and the "candidate of physico-mathematical sciences" (Ph.D.) degree in statistics both from the Lomonosov Moscow State University, Moscow, Russia, in 2007 and 2011, respectively. In 2008–2010 he was collaborating with the Ariana team at INRIA as a visiting student, and in 2011–2012 he was a postdoctoral fellow with Ariana and Ayin teams at INRIA, Sophia Antipolis, France. Since 2012 he is a research associate with the Dept. of Statistical Science at the University College London, London, UK.

His research interests are in the field of statistical signal processing and pattern recognition applied to medical and remote sensing imagery.



**Gabriele Moser** (S'03–M'05) received the "laurea" (M.S.) degree (*summa cum laude*) in telecommunications engineering and the Ph.D. degree in space sciences and engineering from the University of Genoa, Genoa, Italy, in 2001 and 2005, respectively. Since 2010, he has been an Assistant Professor of telecommunications at the University of Genoa. Since 2001, he has cooperated with the Image Processing and Pattern Recognition for Remote Sensing (IPRS) laboratory of the University of Genoa. From January to March 2004, he was a visiting student at the Institut National de Recherche en Informatique et en Automatique (INRIA), Sophia Antipolis, France, working with the Ariana research group on the problem of SAR data modeling.

Dr. Moser has been an Associate Editor of IEEE Geoscience and Remote Sensing Letters and Pattern Recognition Letters since 2008 and 2011, respectively. He was the recipient of the Best Paper Award at the 2010 IEEE Workshop on Hyperspectral Image and Signal Processing. He has been a reviewer for several international journals.

His current research interests include contextual classification, multitemporal image classification and change detection, SAR data analysis, kernel-based methods, and geo/biophysical parameter estimation.



**Sebastiano B. Serpico** (M'87–SM'00–F'08) received the Laurea degree in electronic engineering and the Doctorate from the University of Genoa, Italy, in 1982 and 1989, respectively. Full Professor of telecommunications at the Polytechnic School of the University of Genoa, he is the Head of the Image Processing and Pattern Recognition for Remote Sensing (IPRS) laboratory of the Department of Electrical, Electronic, Telecommunications Engineering and Naval Architecture (DITEN) of the University of Genoa.

Prof. Serpico is the Chairman of the Institute of Advanced Studies in Information and Communication Technologies (ISICT). He has been the project manager of numerous research contracts and an evaluator of project proposals for various programmes of the European Union. He is author (or coauthor) of more than 200 scientific articles published in journals and conference proceedings. Since 2001, he has been an Associate Editor of the international journal IEEE Transactions on Geoscience and Remote Sensing (TGRS). He was a guest editor of two Special Issues of TGRS on "Analysis of hyperspectral image data" (July 2001) and "Advances in techniques for the analysis of remote sensing data" (March 2005). From 1998 to 2002 he was the chairman of a SPIE/EUROPTO

series of conferences on Signal and Image Processing for Remote Sensing. He was the recipient of the Best Paper Award at the 2010 IEEE Workshop on Hyperspectral Image and Signal Processing.

His current research interests include pattern recognition for remote sensing images and for biomedical images.



**Josiane Zerubia** (S'78–M'81–SM'99–F'03) received the M.Sc. degree from the Department of Electrical Engineering at Ecole Nationale Supérieure d'Ingénieurs Electriciens de Grenoble, Grenoble, France, in 1981, and the Dr.Eng. degree, the Ph.D. degree, and the "Habilitation," all from the University of Nice, Sophia-Antipolis, France, in 1986, 1988, and 1994, respectively. She has been a permanent research scientist at INRIA since 1989, and director of research since July 1995. She was head of the PASTIS remote sensing laboratory (INRIA Sophia-Antipolis) from mid-1995 to 1997 and of the Ariana research group (INRIA/CNRS/University of Nice), which worked on inverse problems in remote sensing, from 1998 to 2011. Since January 2012, she has been head of Ayin research group (INRIA-SAM) dedicated to hierarchical and stochastic models for remote sensing and skincare imaging. She has been professor at SUPAERO (ISAE) in Toulouse since 1999. Before that, she was with the Signal and Image Processing Institute of the University of Southern California (USC) in Los-Angeles as a postdoc. She also worked as a researcher for the LASSY (University of Nice/CNRS) from 1984 to 1988 and in the Research Laboratory of Hewlett Packard in France and in Palo-Alto (CA) from 1982 to 1984.

Dr. Zerubia is currently a member of the IEEE IVMSP TC and was a member of the IEEE BISP TC (SP Society) from 2004 till 2012. She was associate editor of IEEE Trans. on IP from 1998 to 2002; area editor of IEEE Trans. on IP from 2003 to 2006; guest co-editor of a special issue of IEEE Trans. on PAMI in 2003; and member-at-large of the Board of Governors of the IEEE SP Society from 2002 to 2004. She has also been a member of the editorial board of the French Society for Photogrammetry and Remote Sensing (SFPT) since 1998, of the International Journal of Computer Vision since 2004, and of the Foundation and Trends in Signal Processing since 2007. She has been associate editor of the on-line resource Earthzine (IEEE CEO and GEOSS). She was co-chair of two workshops on Energy Minimization Methods in Computer Vision and Pattern Recognition (EMMCVPR'01, Sophia Antipolis, France, and EMMCVPR'03, Lisbon, Portugal); co-chair of a workshop on Image Processing and Related Mathematical Fields (IPRM'02, Moscow, Russia); technical program chair of a workshop on Photogrammetry and Remote Sensing for Urban Areas, Marne La Vallee, France, 2003 ; co-chair of the special sessions at IEEE ICASSP 2006 (Toulouse, France) and IEEE ISBI 2008 (Paris, France); and publicity chair of IEEE ICIP 2011 (Brussels, Belgium). She is currently a member of the organizing committee of a Workshop on Mathematics of the Planet Earth (MPE) 2013 (Mumbai, India) and tutorial co-chair of IEEE ICIP 2014 (Paris, France).

Her current research interests are in image processing using probabilistic models and variational methods. She also works on parameter estimation and optimization techniques.