Classification of Very High Resolution SAR Images of Urban Areas Using Copulas and Texture in a Hierarchical Markov Random Field Model

Aurélie Voisin, Vladimir A. Krylov, Gabriele Moser, Member, IEEE, Sebastiano B. Serpico, Fellow, IEEE, and Josiane Zerubia, Fellow, IEEE

Abstract—This letter addresses the problem of classifying synthetic aperture radar (SAR) images of urban areas by using a supervised Bayesian classification method via a contextual hierarchical approach. We develop a bivariate copula-based statistical model that combines amplitude SAR data and textural information, which is then plugged into a hierarchical Markov random field model. The contribution of this paper is thus the development of a novel hierarchical classification approach that uses a quad-tree model based on the wavelet decomposition and an innovative statistical model. The performance of the developed approach is illustrated on a high resolution satellite SAR image of urban areas.

Index Terms—Synthetic aperture radar, urban areas, supervised classification, hierarchical Markov random fields, wavelets, textural features.

I. INTRODUCTION

SYNTHETIC aperture radar (SAR) is an active image acquisition system which allows day-and-night and all-weather acquisitions [1]. Such properties are crucial, for example, for the risk management application, allowing land-use and land cover mapping, as well as the detection of areas damaged by natural disasters such as earthquakes or floodings. In this letter we address the problem of classifying SAR images of urban areas, that represents an especially interesting typology in important applications such as civil protection for disaster monitoring and damage assessment.

Several difficulties need to be considered to address the SAR classification problem. The first one is related to the inherent multiplicative noise known as speckle, which originates from the interference of the coherent wavefronts. This type of noise is specific to active imaging systems and degrades appreciably the registered imagery [1]. As a consequence, standard classification methods that have been validated for optical data, do not report satisfying results when applied directly to SAR images. Another difficulty is the heterogeneity of urban areas on very high resolution (VHR) images that leads to heterogeneous statistical modeling, reflecting the different ground materials such as asphalt, concrete, metal, etc.

A variety of supervised methods have been considered for SAR image classification in the literature, including various techniques such as active contours [2], bags-of-words [3], rule-based methods [4], etc.

In this paper we propose a statistical Bayesian supervised classification approach that consists of two steps. The first step deals with the SAR amplitude statistical modeling for each target class (e.g. vegetation, urban, etc.) by using finite mixtures of automatically chosen SAR-specific probability density functions (PDFs) [5], that are intended to take into account the above mentioned VHR SAR statistics heterogeneity. We further consider an additional source of information obtained by extracting a textural feature map from the original SAR image in order to optimize the detection of urban areas. The marginal PDFs of the original SAR image and the textural feature are combined via copulas [6], leading to a joint PDF for each class. On the second step the classification map is generated, using the joint copula-based statistics. To improve the robustness with respect to speckle noise, we consider a contextual model based on a hierarchical Markov random field (MRF) [7].

The novel contribution of this letter is twofold. First, we propose a joint amplitude-texture PDF modeling that aims to help discriminating the urban areas. Secondly, we integrate this joint PDF model into a hierarchical MRF in which we introduce a prior estimation update that experimentally leads to improved results, that are less affected by speckle noise when compared to a predefined prior [8].

The letter is organized as follows. In Sec. II, we introduce the statistical bivariate copula-based model that combines marginal PDFs model of SAR amplitude data and its textural feature. In Sec. III, we develop the Marginal Posterior Mode (MPM) model based on the quad-tree and integrating a prior update. In Sec. IV, we present classification results obtained on a high resolution COSMO-SkyMed image. Finally, in Sec. V, we draw conclusions.

II. JOINT AMPLITUDE-TEXTURE PDF MODEL

To construct an urban area-specific model, we propose to extract additional textural information from the original SAR image. To this end, the class-conditional marginal statistics of the SAR image and its textural feature are independently modeled and then gathered into a joint amplitude-texture PDF by using the statistical instrument of copulas.
A. Textural features

Well-chosen textural features turn out to be discriminative with respect to urban areas. Taking into account this information is helpful when classifying such areas. Motivated by the experimental textural feature extraction study performed in [9], we consider a feature given by the greylevel co-occurrence matrices (GLCM) [10]. It 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 that a pixel with value \(i\) is adjacent, in a given direction, to a pixel with value \(j\). In our case, we consider horizontal adjacency, with an offset equal to 1, i.e. the co-occurrence matrix is filled by considering a reference pixel and the pixel located to its right. Among various textures that can be extracted from the greylevel co-occurrence matrix, we use the variance, which discriminates well the urban areas (as confirmed by our experiments).

Such approach to texture extraction is applied on a moving-window basis, meaning that each pixel of the image successively becomes the central pixel of the window, and its value is replaced by the variance value estimated within the window. According to preliminary experiments, the choice of a \(5 \times 5\) local window provides more accurate results on the considered data sets than any other window of size in the range \([3; 17]\) . Higher values of window size have not been considered so as to avoid a smearing effect on object edges.

B. Combined amplitude-texture model

We now present the marginal statistics estimation and the construction of a copula-based joint amplitude-texture PDF \(p(y | \omega_m)\). Here, \(y = (y_1, y_2)\) represents the original image \(y_1\) and its corresponding textual feature \(y_2\); \(\omega_m\) is the \(m^{th}\) class considered for the classification, \(m \in [1; M]\) .

1) Marginal PDF estimation: Given a training set, for each input image \(j\) (SAR and textural feature) and for each class \(m\), the PDF \(p_m(y_j | \omega_m)\) is modeled by finite mixtures of independent greylevel distributions:

\[
p_m(y_j | \omega_m) = \sum_{i=1}^{K} P_{mi} p_{mi}(y_j | \theta_{mi}),
\]

(1)

where \(P_{mi}\) are the mixing proportions such that for a given \(m\), \(\sum_{i=1}^{K} P_{mi} = 1\) with \(0 \leq P_{mi} \leq 1\); \(\theta_{mi}\) 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 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 [5].

The PDFs \(p_{mi}(y_j | \theta_{mi})\) are automatically chosen in a predefined dictionary including the four following distributions: Log-Normal, Weibull, Nakagami and Generalized Gamma. To estimate the best-fitting mixture model for each considered class, we combine a density parameter estimation via the method of Log-cumulants and a stochastic expectation maximization (SEM) algorithm (see in [5] and [11]).

The presented estimation approach was initially developed for SAR amplitude data, but the high flexibility of this model and the lack of accurate statistical model for the textural data motivated us to extend its use to the extracted textural features. We empirically observed the very high estimation quality of this approach.

2) Copula-based joint PDF modeling: The use of the copula instrument allows us to model joint PDFs for amplitude and textural SAR channels by fitting several different dependence structures and selecting the best fit. A bivariate copula is a 2-dimensional joint distribution defined on \([0, 1]^2\) such that marginal distributions are uniform on \([0, 1]\). According to Sklar’s theorem [6], we can obtain a unique class-conditional joint PDF \(p(y | \omega_m)\) for any continuous random variables \(y_1\) and \(y_2\) in \(\mathbb{R}^+\) as:

\[
p_m(y_1 | \omega_m)p_m(y_2 | \omega_m) \times \frac{\partial^2 C_m}{\partial y_1 \partial y_2} (F_m(y_1 | \omega_m), F_m(y_2 | \omega_m)),
\]

(2)

where \(p_m(y_1 | \omega_m)\) and \(p_m(y_2 | \omega_m)\) are the marginal PDFs estimated in Sec. II-B1, and \(\{F_m\}\) their corresponding cumulative distribution functions. For each class \(m\), the computation of the joint PDF \(p(y | \omega_m)\) boils down to the determination of the copula family \(C_m\), according to Eq. (2).

To find the best fitting copula \(C_m\), we consider a dictionary of five copulas: Clayton, Ali-Mikhail-Haq, Frank, Marshall-Olkin and Farlie-Gumbel-Morgenstern [6]. This choice of copulas is somewhat arbitrary, but is capable of modeling a wide variety of dependence structures [6], [12]. This set of copulas can be adjusted to fit specific classes for a given application. The analytical expressions of these copulas involve only one parameter \(\alpha_m\). To estimate this parameter, we use the relationship between copulas and Kendall’s \(\tau\), a correlation coefficient [6] that can be empirically estimated over a training set. Then, for each class \(m\), we choose the best fitting copula of the dictionary according to the highest \(p\)-value reported by the Pearson chi-square goodness-of-fit test [11].

III. DEVELOPED HIERARCHICAL CLASSIFICATION APPROACH

We aim at estimating a set of hidden labels \(X\) given a set of observations \(Y\). \(X\) and \(Y\) are considered to be random processes and \(X\) is Markovian with respect to scale. We propose to employ an explicit hierarchical graph-based model [8] to address our classification problem. In the following, we shall focus on specific graphs, which have a quad-tree structure (see Fig. 1). The set of sites \(S\) is, therefore, hierarchically partitioned on \(S = S^0 \cup S^1 \cup ... \cup S^R\), where \(R\) corresponds to the coarsest resolution (the root), and \(0\) is the reference level (finest resolution).

The consideration of a quad-tree allows to benefit from its good properties (e.g. causality) and to apply non iterative algorithms, resulting in a computational time decrease. Among the different algorithms employed in the literature, a first option is to estimate exactly the maximum a posteriori (MAP), but 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
successively the labels at each tree level. We use the transition probability in the
posterior mode (MPM) [8]. 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: for example, an error at the coarsest scale is more strongly penalized than an error at the finest scale.

A. Posterior probabilities and their estimation using MPM

The aim here is to maximize the posterior marginal at each site $s$. A classical MPM algorithm [8] is generally run in two passes, referred to as bottom-up (“forward”) and top-down (“backward”) passes. The former estimates the partial posterior marginals $p(x_s | y_{d(s)})$, and the latter estimates successively the labels at each tree level $n$ by maximizing the posterior marginals $p(x_s | y)$. Laferte et al. [8] showed that

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

Thus, the bottom-up pass is a recursion that estimates $p(x_s | y_{d(s)})$ starting from the leaves and proceeding until the root is reached. The needed prior information are the likelihood (see Sec. II), the prior (see Sec. III-C) and the transition probabilities (see Sec. III-B).

In the proposed algorithm, the top-down pass is truncated, and we only maximize the posterior marginal at the root (see Figs. 2 and 3). The obtained classification map is thus used to update the prior probabilities at this level (Sec. III-C). The prior distribution at levels $n \in [0; R - 1]$ is given by:

$$p(x^n_s) = \sum_{x^n_s \in \Omega} \omega^n(p^n(x^n_s) | x^n_{s^-}) p(x^n_{s^-}).$$

Then, a novel MPM algorithm is run on a smaller quad-tree until scale 0 is reached. Such procedure aims to improve the robustness with respect to the speckle noise [13]. The $p(x_s | y)$ maximization is done by employing a modified Metropolis Dynamics algorithm (MMD), that has good properties for both its relative low computation time and the good precision of its results [14].

B. Transition probabilities

The transition probabilities between the scales, $p(x_s | x_{s^-})$, determine the hierarchical MRF since they represent the causality of the statistical interactions between the different levels of the tree. We use the transition probability in the form introduced by Bouman et al., used in [8]: for all sites $s \in S$ and all scales $n \in [0; R - 1]$,

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

where $\omega_m$ and $\omega_k$ represent the classes $m$ and $k$, $m, k \in [1; M]$, respectively, and $M$ represents the number of considered classes for the final classification. $\theta_n$ is the prior parameter with $\theta_n > 1/M$, here chosen as independent of the tree level $n$. This model favors an identical parent-child labeling.

C. Prior probabilities

To estimate the priors given the obtained classification map, we resort to a single-scale Markovian model which takes into account the spatial contextual information, and therefore leads to a better prior estimation. By employing the Hammersley-Clifford theorem, we can define a local characteristic for each site:

$$p(x_s) = \frac{1}{Z} \exp(-\beta) \sum_{s \in C} \delta_{x_s = x_t},$$

with $\delta_{x_s = x_t} = 1$ if $x_s = x_t$, $\delta_{x_s = x_t} = 0$ otherwise, and where $Z$ is the normalization constant, $s, t$ denote the sites in the same clique, and $x_s, x_t$ their labels. Instead of using a second-order neighborhood set based on the 8 pixels surrounding a given pixel, we consider different kinds of set geometries (adaptive neighborhood), and we select the one which leads to the smallest energy at a given site [13]. The adaptivity of the neighborhood aims to take into account the geometrical properties of the different areas in our original image.

IV. EXPERIMENTAL RESULTS

We present results obtained on 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 x 820 pixels, shown in Fig. 4(a). In the considered explicit hierarchical model, we need multi-scale input data. When, as in our case, a multi-resolution input is not available, we suggest to decompose our image by resorting to a 2-D discrete wavelet transform. The approximation coefficients at each scale are considered as a multi-resolution input. At each level, the
textural feature map is obtained from the corresponding image in the decomposition stack. Finally, at each level, the wavelet image is combined with the textural image via a copula-based model presented in Sec. II. Empirically, we selected the Daubechies-10 wavelet [15] that reported the highest classification accuracy when compared to the other families [13].

We built non-overlapping training and test sets, and compared the following classification methods:

1. The proposed amplitude texture MPM-based approach.
2. The proposed MPM-based approach based solely on the amplitude statistical model (no textures).
3. Laferte’s hierarchical MRF approach based solely on the amplitude statistical model (no textures).
4. The amplitude-texture statistical model (Sec. II) combined with a single-scale MRF (parameter set to $\beta = 1.2$).

The Laferte’s algorithm is applied to a quad-tree [8], and the prior probabilities at the root are estimated by using a $K$-nearest neighbor ($K$-NN) [16] preliminary classification, with $K$ obtained by cross-validation. The considered classes are water, urban areas and vegetation. Empirically, we have concluded that good results are obtained with the decomposition on $R = 2$ levels, $\beta = 4.8$ (Eq. 6) and $\theta_n = 0.8$ (Eq. 5).

The results obtained by our proposed method are satisfying: the classification map (Fig. 4 (c)) is quite detailed and the accuracy table (Tab. I) indicates that the overall accuracy is 98.07%. As expected, the hierarchical model allows to take into account spatial details and is also relatively robust to speckle noise. Moreover, the urban areas are well discriminated by the use of the textural information in the appropriate copula-based model. Such improvement is shown when comparing the results obtained by the proposed method applied to only amplitude SAR image, and applied to the amplitude-texture combination (see Figs. 4 (c), (d)).

To isolate the improvements brought by our proposed hierarchical model, we compare it with the Laferte’s method, both based only on the amplitude statistical model. The former model is less affected by speckle noise (Fig. 4) thanks to the prior update, especially for the vegetation class (in the top center of the maps). Moreover, the urban classification accuracy is higher with the developed method. Laferte’s one classifies almost all roads as vegetation (e.g., in the bottom center of Fig. 4 (e)).

By looking at Tab. I, we notice that slightly better results for this dataset are obtained with the single-scale MRF-based approach (Fig. 4 (f)). We stress though that this result is obtained at the expense of a quite strong smearing effect on object borders (this does not affect the classification accuracy in Tab. I because the test set has been extracted from the internal parts of objects). Thus, the hierarchical approach is preferable for the urban areas classification since it allows to extract more details than the single-scale model. In general, the main water misclassifications comes from the cross-like artifact of the SAR acquisition (point spread function of the SAR sensor [1]). We observe that none of the algorithms employed in this comparison is robust to such artifacts.

Experiments were conducted on Intel Xeon quad-core (2.4GHz, 12MB cache), 18Gb RAM, 64-bits Linux system. Our proposed algorithm runs in 6 minutes and 45 seconds, which is acceptable when compared to the Laferte’s method computation time of 5 minutes and 30 seconds, given the visual refinement of the classification map.

<table>
<thead>
<tr>
<th></th>
<th>Port-au-Prince quay</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>water</td>
<td>urban</td>
</tr>
<tr>
<td>Proposed method</td>
<td>97.65%</td>
<td>97.99%</td>
</tr>
<tr>
<td>Proposed method (no texture)</td>
<td>97.67%</td>
<td>97.27%</td>
</tr>
<tr>
<td>Laferte (no texture)</td>
<td>97.61%</td>
<td>88.52%</td>
</tr>
<tr>
<td>MRF-based classif.</td>
<td>97.59%</td>
<td>99.03%</td>
</tr>
</tbody>
</table>

V. CONCLUSION

The supervised classification method proposed in this paper combines a joint copula-based statistical modeling of amplitude SAR images and extracted textural features, with a hierarchical Markov random field that integrates a prior probability update, thus improving the robustness of the classifier with respect to speckle noise. The labels are determined by resorting to a non-iterative optimization algorithm (MPM estimation). The results of the proposed hierarchical classification algorithm appear to be a good trade-off between the quite blurred
results produced by the single-scale MRF-based method and the rather noisy ones obtained by using the Laferte standard model. Furthermore, the developed approach can be extended to the use of multi-pol, multi-resolution and/or multi-sensor data, which represents a crucial advantage given the considerable amount of remotely sensed data acquired on a daily basis. This is considered as the main direction of further research.

ACKNOWLEDGMENT

The authors would like to thank Dr. M. De Martino (University of Genoa, Italy) for preparing the ground truth maps, and the Italian Space Agency for providing the COSMO-SkyMed (CSK®) image (COSMO-SkyMed Product - ©ASI - Agenzia Spaziale Italiana - 2009. 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)”.

REFERENCES