Multichannel SAR Image Classification by Finite Mixtures, Copula Theory and Markov Random Fields

Vladimir A. Krylov [1*,†], Gabriele Moser[**], Sebastiano B. Serpico[**] and Josiane Zerubia[†]

[*]*Dept. of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University; 1, Leninskie Gory, 119991, Moscow, Russia*
[†]*EPI Ariana, INRIA/CNRS/UNSA, CR INRIA Sophia Antipolis - Méditerranée; 2004, Route des Lucioles, B.P. 93, F-06902, Sophia Antipolis, France*
[**]*Dept. of Biophysical and Electronic Engineering (DIBE), University of Genoa; Via Opera Pia 11a, I-16145, Genoa, Italy*

Abstract. In this paper we develop a supervised classification approach for medium and high resolution multichannel synthetic aperture radar (SAR) amplitude images. The proposed technique combines finite mixture modeling for probability density function estimation, copulas for multivariate distribution modeling and a Markov random field (MRF) approach to Bayesian classification. The novelty of this research is in introduction of copulas to classification of D-channel SAR, with $D \geqslant 3$, within the mainframe of finite mixtures - MRF approach. This generalization results in a flexible and well performing multichannel SAR classification technique. Its accuracy is validated on several multichannel Quad-pol RADARSAT-2 images and compared to benchmark classification techniques.

INTRODUCTION

The last decades have witnessed an intensive development and a significant increase of interest to remote sensing, and, in particular, to synthetic aperture radar (SAR) image processing. This paper focuses on the supervised SAR image classification, which is one of the fundamental SAR image processing problems. Recently, various models have been proposed for modeling the single channel statistics of SAR data [1], however, neither is general and flexible enough to model the joint probability density function (pdf) in case of D-channel SAR, $D \geqslant 3$. In this paper we propose a joint pdf model for multichannel SAR, based on finite mixture modeling for marginal pdf estimation and copulas for multivariate distribution modeling. We apply this model to medium and high resolution multichannel SAR amplitude image classification by combining it with a contextual *Markov random field* (MRF) approach [2] that allows to take into account the contextual information and to gain robustness against the inherent noise-like phenomenon of SAR known as speckle. The finite mixture modeling is done via a recently proposed SAR-

---

[1] Email: vkrylov@cs.msu.ru.

specific *dictionary-based stochastic expectation maximization* (DSEM) approach [3, 4], that is applied to class-conditional amplitude probability density function estimation separately to all the SAR channels. For modeling the class-conditional joint distributions of multichannel data the statistical concept of *copulas* [5] is employed, and a dictionary-based copula selection method is proposed. The contribution of this study is the generalization of the recently considered DSEM - MRF classification approach [6] to *D*-channel SAR, $D \geqslant 3$, via copulas, and its experimental validation on several multichannel Quad-pol RADARSAT-2 images with a comparison to benchmark classification techniques.

METHODOLOGY

Method overview

We consider the supervised classification of a *D*-channel SAR image with *M* target classes and we assume training pixels to be available for all the classes.

Marginal-step. The marginal pdf of each channel for each class is estimated separately by applying DSEM to the training pixels. The mixture DSEM pdfs $p_{dm}(y_d|\omega_m)$ and the corresponding cumulative distribution functions (CDFs) $F_{dm}(y_d|\omega_m)$, where ω_m is the event of observing the *m*-th class, are as follows:

$$p_{dm}(y_d|\omega_m) = \sum_{i=1}^{K_{dm}} P_{dmi} p_{dmi}(y_d|\theta_{dmi}), \quad F_{dm}(y_d|\omega_m) = \sum_{i=1}^{K_{dm}} P_{dmi} F_{dmi}(y_d|\theta_{dmi}), \quad (1)$$

where $m = 1,\dots,M$, $d = 1,\dots,D$, y_d are the observed amplitudes and K_{dm} are the numbers of mixture components. F_{dmi} and p_{dmi} represent the *i*-th mixture component in the CDF and pdf domains, respectively, and P_{dmi} are the related mixture proportions.

Copula-step. The joint pdfs $p_m(\mathbf{y}|\omega_m)$ for classes $m = 1,\dots,M$ are constructed from marginals (1) by means of automatically selected from a specific dictionary copulas C_m^*:

$$p_m(\mathbf{y}|\omega_m) = p_{1m}(y_1)\cdots p_{Dm}(y_D)\frac{\partial^D C_m^*}{\partial y_1 \cdots \partial y_D}(F_{1m}(y_1),\dots,F_{Dm}(y_D)). \quad (2)$$

Contextual-step. To take into account the contextual information disregarded by the pixel-based Copula-DSEM pdfs (2) and to gain robustness against the inherent noise-like phenomenon of SAR known as speckle, we adopt a contextual approach based on an MRF model. Following a classical definition of an MRF [2], we consider a two dimensional lattice $\mathcal{S}$ with N pixels $\{s_i\}$ and class labels $\mathbf{x} = \{x_i\}$, $x_i \in \{1,\dots,M\}$ and introduce an isotropic second-order neighborhood system $\mathcal{C}$ with cliques of size 2. The resulting MRF energy function for the class labels is parameterized by $\beta > 0$:

$$H(\mathbf{x}|\beta) = \sum_{\{s,s'\}\in\mathcal{C}} \left[-\beta\,\delta_{x_s=x_{s'}}\right], \quad \text{with } \delta_{x_s=x_{s'}} = \begin{cases} 1, & \text{if } x_s = x_{s'} \\ 0, & \text{otherwise} \end{cases}.$$

In a hidden MRF model, the unobserved class labels are modeled as an MRF and the observations are conditionally independent given labels [2]. Given (2) and denoting the

value of vector $\mathbf{y}$ at pixel i as $\mathbf{y}_i$, the hidden MRF energy function is the following:

$$U(\omega_m|\mathbf{y},\beta) = \sum_{i \in S} \left[-\log p_m(\mathbf{y}_i|\omega_m) - \beta \sum_{s:\{i,s\} \in C} \delta_{x_i=x_s} \right]. \quad (3)$$

In order to estimate β in (3), we suggest to use a simulated annealing procedure [7] with pseudo-likelihood function PL [2] of the following form:

$$\log PL(\mathbf{x}|\beta) = \log \left[\prod_{s \in S} p(x_s|\mathbf{x}_{S \setminus \{s\}},\beta) \right] = \log \left[\prod_{s \in S} \frac{\exp(-U(x_s|\mathbf{x}_{S \setminus \{s\}},\beta))}{\sum_{z_s \in X_S} \exp(-U(z_s|\mathbf{x}_{S \setminus \{s\}},\beta))} \right], \quad (4)$$

and $x_s = \{\omega_1, \ldots, \omega_M\}$. We employ $N(\beta_t, 1)$ as a proposal distribution [6] and exponential decrease $T_t = 0.95 \cdot T_{t-1}$ as a cooling schedule. The β-estimation is performed on maximum likelihood preclassification map, associating with every pixel of the image a label with the highest probability assigned by the context-free Copula-DSEM model (2).

Optimization-step. This step involves the minimization of the energy (3)-(4). For this optimization problem an iterative deterministic *Modified Metropolis Dynamics* (MMD) [8] algorithm is adopted. It is a compromise approach between fast but local deterministic Iterated Conditional Modes algorithm [2], and global yet slow Simulated Annealing [7]. As such, MMD is computationally feasible and provides reasonable results in real classification problems [8] and in our optimization problem. Structurally, MMD algorithm proceeds as follows:

1. sample a random initial configuration ω^0; set $k = 0$ and temperature $T_0 = T^0$;

2. increase k; using a uniform distribution pick a state η which differs exactly in one element from ω^k;

3. compute $\Delta U = U(\eta) - U(\omega)$ and accept η according to the rule:

$$\omega^{k+1} = \begin{cases} \eta, & \text{if } \Delta U \leqslant 0 \quad \text{or} \quad \ln(\alpha) \leqslant -\frac{\Delta U}{T_k} \leqslant 0, \\ \omega^k, & \text{otherwise.} \end{cases}$$

where $\alpha \in (0,1)$ is a **constant** threshold, chosen once at the start of the algorithm;

4. decrease $T_{k+1} = \tau \cdot T_k$, and goto Step 2 until convergence, i.e. $\Delta U/U < \gamma$ is fulfilled.

Dictionary-based Stochastic Expectation Maximization

To take into account the heterogeneity scenario, when several distinct land-cover typologies are present in the same SAR image, a finite mixture model (FMM) for the distribution of grey levels is assumed. Specifically, the dictionary-based stochastic expectation maximization (DSEM) approach developed in [3, 4] is applied to estimate marginal class-conditional statistics. Here, we follow the notations in [4] and refer to it for more details. Separately focusing on each class ω_m and each channel y_d, we assume the training samples to be i.i.d. random variables, drawn from a mixture

TABLE 1. DSEM dictionary $\mathcal{D}$ of parametric pdf families with MoLC equations

Family	Probability density function	MoLC equations
Log-normal	$p_1(r) = \frac{1}{\sigma r \sqrt{2\pi}} \exp\left[-\frac{(\ln r - m)^2}{2\sigma^2}\right],\ r > 0$	$\kappa_1 = m$ $\kappa_2 = \sigma^2$
Weibull *	$p_2(r) = \frac{\eta}{\mu^\eta} r^{\eta-1} \exp\left[-\left(\frac{r}{\mu}\right)^\eta\right],\ r \geqslant 0$	$\kappa_1 = \ln\mu + \Psi(1)\eta^{-1}$ $\kappa_2 = \Psi(1,1)\eta^{-2}$
Nakagami †	$p_3(r) = \frac{2}{\Gamma(L)}(\lambda L)^L r^{2L-1} \exp\left(-\lambda L r^2\right),\ r \geqslant 0$	$2\kappa_1 = \Psi(L) - \ln\lambda L$ $4\kappa_2 = \Psi(1,L)$
Generalized Gamma	$p_4(r) = \frac{\nu}{\sigma\Gamma(\kappa)}\left(\frac{r}{\sigma}\right)^{\kappa\nu-1} \exp\left\{-\left(\frac{r}{\sigma}\right)^\nu\right\},\ r \geqslant 0$	$\kappa_1 = \Psi(\kappa)/\nu + \ln\sigma$ $\kappa_j = \Psi(j-1,\kappa)/\nu^j,\ j = 2,3$

* $\Psi(\cdot)$ the digamma function and $\Psi(\nu,\cdot)$ the νth order polygamma function [9]
† $\Gamma(\cdot)$ is the gamma function [9]

pdf $p(r) = \sum_{i=1}^K P_i p_i(r),\ r \geqslant 0$, with K components, where $p_i(\cdot)$ is the i-th mixture component, $i = 1,\ldots,K$, and $\{P_i\}$ is a set of mixing proportions, i.e., $\sum_{i=1}^K P_i = 1$ and $0 < P_i < 1$. Each component $p_i(\cdot)$ is modeled by resorting to a finite dictionary $\mathcal{D} = \{p_1,\ldots,p_4\}$ (see Table 1) of four SAR specific distinct parametric pdfs $p_j(r|\theta_j)$, parameterized by $\theta_j \in A_j,\ j = 1,\ldots,4$.

For this mixture estimation problem we use the iterative stochastic expectation maximization (SEM) scheme [10]. The convergence properties of SEM have been proved under suitable assumptions [10], which do not hold strictly for all the pdfs in $\mathcal{D}$. However, we recall that SEM, compared to the classical EM or other deterministic variants for FMM estimation, was specifically designed to improve the exploratory properties of EM in case of multimodal likelihood function [10]. Instead of adopting the maximum likelihood estimates as the classical SEM scheme [10] suggests, DSEM employs the Method of Log-Cumulants (MoLC) [3] for component parameter estimation. Based on the Mellin transform [9], MoLC allows to formulate a set of equations relating the unknown parameters of a given parametric pdf with several sample log-cumulants (see Table 1). MoLC has been demonstrated to be an effective estimation tool for all the pdfs in $\mathcal{D}$ [4].

Each iteration of DSEM goes as follows (everywhere $i = 1,\ldots,K_t,\ j = 1,\ldots,M$):
• **E-step:** for each greylevel $z = 0,\ldots,Z-1$ and $\forall i$, compute the posterior probability estimates corresponding to the current pdf estimates:

$$\tau_i^t(z) = \frac{P_i^t p_i^t(z)}{\sum_{j=1}^{K_t} P_j^t p_j^t(z)},$$

with greylevel labels $s_z \in \{\sigma_1,\ldots,\sigma_{K_t}\}$, assigning each greylevel z to one of the mixture components, and $p_i^t(\cdot)$ being the σ_i-conditional pdf estimate on the t-th step.
• **S-step:** $\forall z$ sample the label $s^t(z)$ according to the posterior distribution $\{\tau_i^t(z)\}$.
• **MoLC-step:** for the i-th mixture component, compute the following histogram-based estimates of the mixture proportions and the first three log-cumulants:

$$P_i^{t+1} = \frac{\sum_{z \in Q_{it}} h(z)}{\sum_{z=0}^{Z-1} h(z)}, \qquad \kappa_{1i}^t = \frac{\sum_{z \in Q_{it}} h(z)\ln z}{\sum_{z \in Q_{it}} h(z)}, \qquad \kappa_{bi}^t = \frac{\sum_{z \in Q_{it}} h(z)(\ln z - \kappa_{1i}^t)^b}{\sum_{z \in Q_{it}} h(z)},$$

TABLE 2. Dictionary of copula families $C(\mathbf{u})$ with $\theta(\bar{\tau})$ dependencies.

Family	$C(\mathbf{u})$	$\theta(\bar{\tau})$ dependence
Clayton	$C^1(\mathbf{u}) = (u_1^{-\theta} + \cdots + u_D^{-\theta} - D + 1)^{-1/\theta}$	$\theta = \frac{2\bar{\tau}}{1-\bar{\tau}}$
Gumbel-Hougaard	$C^2(\mathbf{u}) = \exp\left(-\left[(-\log(u_1))^\theta + \cdots + (-\log(u_D))^\theta\right]^{1/\theta}\right)$	$\theta = \frac{1}{1-\bar{\tau}}$
Frank	$C^3(\mathbf{u}) = -\frac{1}{\theta}\ln\left(1 + \frac{(e^{-\theta u_1}-1)\cdots(e^{-\theta u_D}-1)}{(e^{-\theta}-1)^{D-1}}\right)$	$\bar{\tau} = 1 - \frac{4}{\theta^2}\int_0^\theta \frac{t}{e^{-t}-1}dt$

where $b = 2, 3$; $h(z)$ is the image histogram; $Q_{it} = \{z : s^t(z) = \sigma_i\}$ is the set of grey levels assigned to the i-th component; then, solve the corresponding MoLC equations (Table 1) for each parametric family $f_j(\cdot|\theta_j)$ in $\mathcal{D}$, resulting in MoLC estimates θ_{ij}^t.

• **K-step:** $\forall i$: if $P_i^{t+1} < \gamma$, eliminate the i-th component; update K_{t+1}. The choice of threshold γ does not appreciably affect DSEM, provided it is small, e.g. 0.005.

• **Model Selection-step:** $\forall i = 1, \ldots, K_{t+1}$, compute the log-likelihood of each estimated pdf $f_j(\cdot|\theta_{ij}^t)$ according to the data assigned to the i-th component: $L_{ij}^t = \sum_{z \in Q_{it}} h(z) \ln f_j(z|\theta_{ij}^t)$, and define $p_i^{t+1}(\cdot)$ as the estimated pdf $f_j(\cdot|\theta_{ij}^t)$ yielding the highest value of L_{ij}^t.

Copulas

For the purpose of modeling the joint distribution of a D-channel SAR image, given the estimates of the related marginal distributions, we employ copulas [5]. In the literature, various models have been proposed for the marginal statistics of SAR amplitudes [1], however, none of them is general and flexible enough to model a joint pdf in case of multichannel SAR data for each class. In order to overcome this limitation, we merge the marginal pdfs provided by DSEM into a joint pdf by means of copulas.

A *D-dimensional copula* is a function $C : [0,1]^D \to [0,1]$, which satisfies:

1. $C(\mathbf{u}) = 0$ for any $\mathbf{u}$: $\exists i \in \{1, \ldots, D\}$ so that $u_i = 0$.
2. $C(\mathbf{u}) = u_d$ for any $\mathbf{u}$: $u_i = 1$, for all $i \neq d$.
3. the *D-increasing* condition: for any $0 \leqslant u_{i,1} \leqslant u_{i,2} \leqslant 1$, where $i = 1, \ldots, D$,

$$V_C \equiv \sum_{i_1=1}^{2} \cdots \sum_{i_D=1}^{2} (-1)^{(i_1+\cdots+i_D)} C(u_{1,i_1}, \ldots, u_{D,i_D}) \geqslant 0.$$

The important property of copulas is given by the *Sklar's theorem* [5], which states the existence of a copula C, that models the joint distribution H of arbitrary random variables $Y_1, \ldots, Y_D$ with CDFs $F_1, \ldots, F_D$: $H(\mathbf{y}) = C(F_1(y_1), \ldots, F_D(y_D))$, $\forall \mathbf{y} \in \mathbb{R}^D$. In this paper we consider *one-parameter archimedean* copulas [5]. This class provides a representative panel of joint pdf forms and has easy parameter estimation techniques [5, 11]. For parameter estimation we use the copula's connection with *Kendall's τ ranking coefficient*:

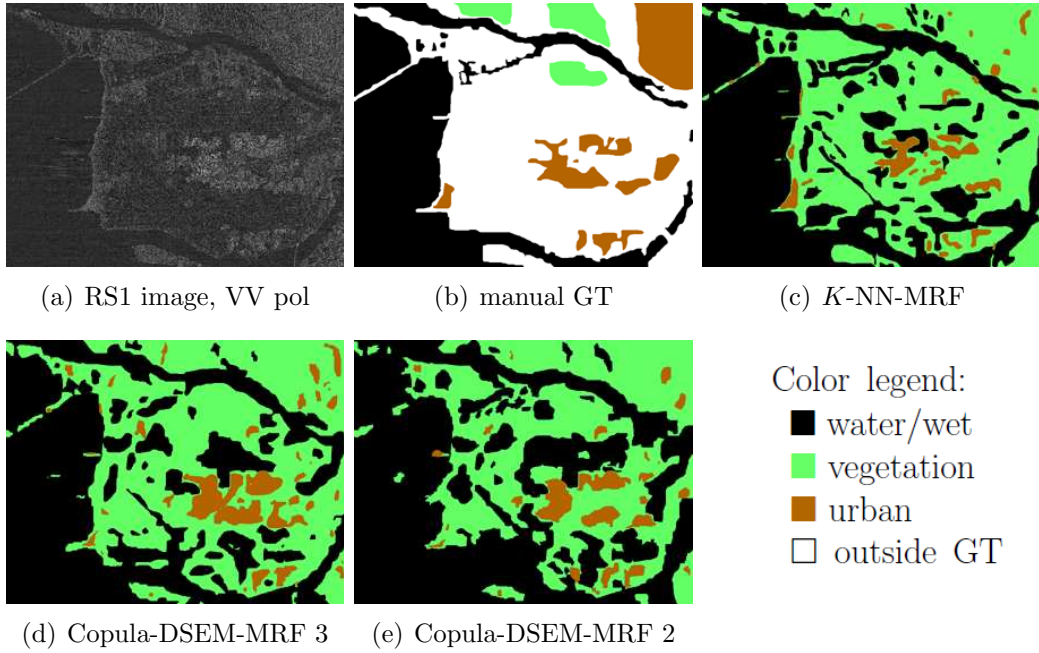$$\bar{\tau} + 1 = 4 \int_{[0,1]^2} C(u_1, u_2) dC(u_1, u_2),$$

(a) RS1 image, VV pol   (b) manual GT   (c) K-NN-MRF

(d) Copula-DSEM-MRF 3   (e) Copula-DSEM-MRF 2

Color legend:
- ■ water/wet
- ■ vegetation
- ■ urban
- □ outside GT

FIGURE 1. (a) RS1 image (1000×700 pixels) in VV pol, RADARSAT-2 Data and Products ©MacDonald, Dettwiler and Associates Ltd., 2008 - All Rights Reserved. (b) Manually created ground truth (GT) and classification maps obtained by: (c) K-NN-MRF on HH/HV/VV, (d) Copula-DSEM-MRF on HH/HV/VV and (e) Copula-DSEM-MRF on HH/VV. Automatically selected copulas: Gumbel for "water/wet" and "vegetation", Frank for "urban".

where $C(u,v)$ is a 2-copula of the same type and $\overline{\tau}$ is the average of $D(D-1)/2$ consistent Kendall's τ estimates corresponding to different bivariate marginals (Y_{d_1}, Y_{d_2}): $\tau = \frac{4}{N(N-1)} \sum_{i \neq j} I[Y_{i,d_1} \leqslant Y_{j,d_1}] I[Y_{i,d_2} \leqslant Y_{j,d_2}] - 1$. The use of Kendall's tau based parameter estimation is explained by a relatively small number of samples N employed [11].

The dictionary of employed archimedean copulas include Clayton, Gumbel-Hougaard and Frank copula families (Table 2). This is a very basic set, yet it allowed to achieve good results and it will be studied further. The choice of a best fitting copula C_m^* in the dictionary for every class $m = 1,\dots,M$, is based on a dedicated criterion: we choose the copula with the highest p-value provided by a *Pearson Chi-square test-of-fitness* (PCS) [12]. PCS tests the null hypothesis that the sample frequencies $(F_{1m}(y_1),\dots,F_{Dm}(y_D))$, $m = 1,\dots,M$, are consistent with the theoretical probabilities for the copula $C^c(\mathbf{y})$, $c = 1,2,3$. Its statistics $\sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$ is asymptotically χ_{n-r-1}^2 distributed [12], with O_i and E_i the observed and the hypothetical frequencies, respectively, n the number of clusters and r the number of parameters for Copula-DSEM pdfs (2).

EXPERIMENTS

The developed multichannel supervised classification approach has been tested on Single-look, Fine Quad-Pol HH/HV/VH/VV, 7.5 m resolution RADARSAT-2 amplitude
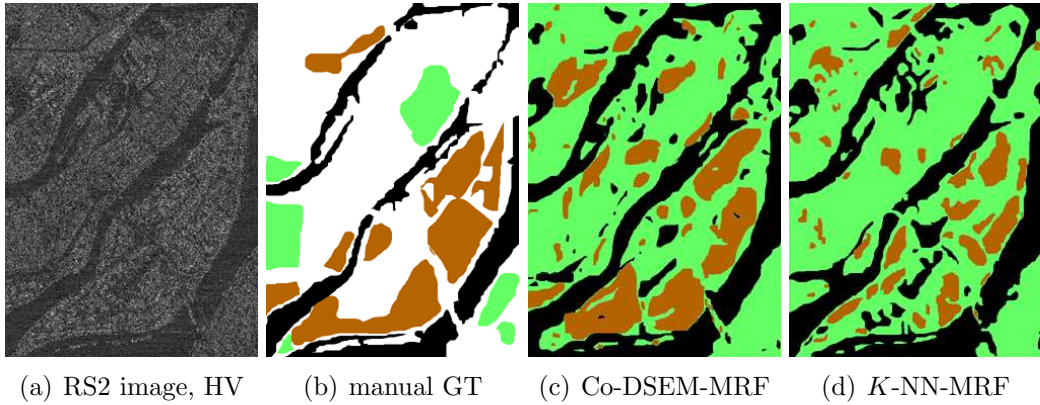
(a) RS2 image, HV   (b) manual GT   (c) Co-DSEM-MRF   (d) K-NN-MRF

FIGURE 2. (a) RS2 image (500×700 pixels) in HV pol, RADARSAT-2 Data and Products ©MacDonald, Dettwiler and Associates Ltd., 2008 - All Rights Reserved. (b) Manually created ground truth (GT) and classification maps obtained by: (c) Copula-DSEM-MRF on HH/HV/VV and (d) K-NN-MRF on HH/HV/VV. See color legend on Fig. 1. Automatically selected copulas: Gumbel for "water/wet", Frank for "vegetation" and "urban".

TABLE 3. Classification accuracies on the considered test images: by classes, average and overall

Image	Method	"Water/wet"	"Vegetation"	"Urban"	Average	Overall
RS1	Copula-DSEM-MRF on HH/HV/VV	98.04%	90.33%	71.49%	86.61%	**88.69%**
	K-NN-MRF on HH/HV/VV	98.85%	89.02%	62.89%	83.59%	**86.42%**
	Copula-DSEM-MRF on HH/VV	98.97%	87.16%	58.74%	81.62%	**84.78%**
RS2	Copula-DSEM-MRF on HH/HV/VV	97.30%	95.12%	86.12%	92.85%	**94.18%**
	K-NN-MRF on HH/HV/VV	98.10%	96.06%	81.33%	91.83%	**93.51%**

images of Vancouver RS1 (Fig. 1) and RS2 (Fig. 2). It is well known that HV and VH polarizations provide very similar information [1], thus, $D = 3$ channels HH/HV/VV without VH were employed in experiments. The experiments involved the classification into $M = 3$ classes: "water/wet", "vegetation" and "urban". In Marginal-step, DSEM was initialized with $K_0 = 3$ mixture components. For PCS test in Copula-step, $n = 25$ of equal-sized square clusters was selected. In Optimization-step, parameters were set to $T^0 = 5.0$, $\alpha = 0.3$, $\tau = 0.97$, $\gamma = 10^{-4}$. The training was performed on a 100×100 subimage endowed with a manually annotated nonexhaustive ground truth, not overlapping with the test areas.

Table 3 reports the accuracies of the developed Copula-DSEM-MRF approach: "by class", "average" (i.e., the arithmetic mean of the class accuracies) and "overall" (i.e., the percentage of correctly classified test samples, irrespective of their classes). The classification results are compared visually (Fig. 1, Fig. 2) and quantitatively (Table 3) with the nonparametric K-nearest neighbors [13] classification method combined with an MRF, with $K^* = 35$ was estimated by cross-validation [13]. On both images the accuracy demonstrated by our method is higher, especially for the "urban" class. On RS1 image, we also provide a comparison with Copula-DSEM-MRF classification based on 2 channels HH/VV to demonstrate the accuracy gain from model generalization to $D \geqslant 3$.

CONCLUSIONS

In this paper a generalized method for supervised multichannel SAR image classification is developed, based on a previously proposed approach [6]. It generalizes the concept of finite mixture combined with Markov random field classification approach to D-channel SAR images, $D \geqslant 3$, via copulas. The resulting classification technique is flexible and performs well in experiments with high resolution Quad-Pol RADARSAT-2 images, thanks to a modeling flexibility provided by copulas and a dictionary-based copula selection approach.

The directions of further research include the study of a more specific copula dictionary, e.g., by including non-symmetric copulas, and the specialization of the model to urban area classification, e.g., by incorporating geometrical features into the MRF model.

ACKNOWLEDGMENTS

REFERENCES

1. C. Oliver, and S. Quegan, *Understanding Synthetic Aperture Radar Images*, Artech House, Norwood, 1998.
2. J. Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society B* **48**, 259–302 (1986).
3. G. Moser, S. B. Serpico, and J. Zerubia, "Dictionary-based Stochastic Expectation Maximization for SAR amplitude probability density function estimation," *IEEE Trans. Geosci. Remote Sens.* **44**, 188–199 (2006).
4. 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. Rem. Sens. Lett.* (2010), (in press). DOI:10.1109/LGRS.2010.2053517.
5. R. B. Nelsen, *An Introduction to Copulas*, Springer, New-York, 2006, 2nd edn.
6. G. Moser, V. Krylov, S. B. Serpico, and J. Zerubia, "High resolution SAR-image classification by Markov random fields and finite mixtures," in *Proc. of SPIE*, San Jose, USA, 2010, 753308.
7. S. Geman, and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Trans. Patt. Anal. Mach. Intell.* **6**, 721–741 (1984).
8. Z. Kato, J. Zerubia, and M. Berthod, "Satellite image classification using a modified Metropolis dynamics," in *Proc. of ICASSP*, San Francisco, USA, 1992, pp. 573–576.
9. I. Sneddon, *The use of integral transforms*, McGraw-Hill, New-York, 1972.
10. G. Celeux, D. Chauveau, and J. Diebolt, On stochastic versions of the EM algorithm, Research Report 2514, INRIA, France (1995).
11. I. Kojadinovic, and J. Yan, "Comparison of three semiparametric methods for estimating dependence parameters in copula models," *Insurance: Mathematics and Economics* **47**, 52–63 (2010).
12. E. Lehmann, and J. Romano, *Testing statistical hypotheses*, Springer, New York, 2005, 3rd edn.
13. C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New-York, 2006.