Parametric and Nonparametric Methods for SAR Patch Scene Categorization

Dušan Gleich, *Member, IEEE*, Jagmal Singh, and Peter Planinšič

Abstract—This paper presents synthetic aperture radar (SAR) image categorization based on feature descriptors within the discrete wavelet transform (DWT) domain using nonparametric and parametric features. The first and second moments, Kolmogorov Sinai entropy and coding gain, are used for the nonparametric features within an oriented dual tree complex wavelet transform (2D ODT℃WT). A Gauss–Markov random field (GMRF), triplet Markov random field (TMRF), and autobinomial model (ABM) are used for feature extraction using a parametric approach within an image domain. A single parameter of GMRF, TMRF, or ABM is used for characterizing an entire patch; therefore, higher model orders (MOs) are used. A database with 2000 images representing 20 different classes with 100 images per class is used for estimating classification efficiency. A supervised learning stage is implemented within a support vector machine (SVM) using 10% and 20% of the test images per class. The experimental results showed that the nonparametric features achieved better results when compared to the parametric features.

I. INTRODUCTION

A SYNTHETIC aperture radar (SAR) system is an imaging system which provides data day and night and in all weather conditions. The SAR data are two-dimensional (2-D) data, the properties of which, within the azimuth and range directions, depend on SAR system parameters [1]. The SAR data can be acquired in single polarization, can be polarized, and can be acquired at the same time using a pair of satellites. Repetitive pass interferometry is enabled using TerraSAR-X and Cosmo Skymed missions, and bistatic interferometry is possible with TanDEM-X. Radar data have been used for estimating the Earth's surface parameter [2], [3], soil moisture estimation [4], characterization of polarimetric features [5], properties of permanent scattering using interferometry techniques [6], land sliding applications, biomass estimation [7], crisis management in the case of flooding [8], earthquake detection and damage assessment after earthquakes [9], fire damage assessment [10], oil slick detection [11], etc.

The goal of this paper was to develop simple methods which would be able to characterize the type of terrain within a SAR data using small patches. In the literature, there are two approaches for scene description. The first group of methods

use theoretical scene models [12]–[14], the parameters of which are estimated from models representing feature parameters. Information extraction using a model-based approach was proposed in [13] and [14], where theoretical texture models were fitted over the real SAR scene. Parametric approaches using theoretical models were also considered in [12], where the Markovian model was used for classification and the statistics of a generalized Gamma distribution were estimated for texture component [15]. A nonstationary multinomial logistic latent class label model was used as a mixture density to obtain spatially smooth class segments of SAR images in [16].

The second group of methods belongs to the nonparametric approaches, which use objective parameters which are estimated from original or transformed data. The nonparametric feature-based classification methods are also well known within optical and remote sensing communities [17] and content-based image retrieval in multimedia [18]. In [19], feature descriptors (first- and second-order statistical moments, Spectral Centroid, Spectral Flux, Spectral Rolloff, Cepstral Coefficients) for SAR-image patches were computed within the Fourier domain. The feature descriptor, based on log cumulants (LC) of Fourier coefficients [20], has also been applied to the same database and those results have been used in this paper for comparison purposes. Wavelet-based subband modeling was used for change detection within a time series of SAR data [21]. SAR image scene classification requires computing the compact feature descriptors of scene contents. An image patch classification approach was proposed in [22] for detecting man made objects within satellite data and extended to SAR patch categorization [23] using a content-based approach. SAR patch categorization was also proposed in [24]. A motivation of this paper was to develop a nonparametric patch-based method for SAR image categorization and compare it using parametric methods. In this paper, a nonparametric feature descriptors represent first and second moments, Kolmogorov Sinai entropy and coding gain within an oriented dual tree complex wavelet transform (2D ODT℃WT). A model-based approach to information extraction using a triplet Markov random field (TMRF), Gauss–Markov random fields (GMRF) [14], and autobinomial model (ABM) [13] were extended to patch characterization. The methods were tested using a database with 2000 images representing 20 different classes with 100 images per class. A supervised learning stage was implemented using a support vector machine (SVM) [25]. The presented feature descriptors were compared with the local binary pattern (LBP) [26] and MR8 [27], which are very effective methods for texture feature extraction.

II. Nonparametric Features

The nonparametric features presented within this section use a dual tree complex wavelet transform. The SAR data are complex-valued data; therefore, the complex-valued wavelets can be constructed using complex-valued filter banks [28] within the framework of the Daubechies orthogonal filter banks. The complex-valued wavelet transform was later generalized to the complex wavelet transform ($\mathbb{C}$WT) [29]. $\mathbb{C}$WT is a complex-valued extension to the standard real-valued discrete wavelet transform (DWT). $\mathbb{C}$WT has an analytical wavelet function, by which the imaginary part is the Hilbert transform of the real part. By the implementation of $\mathbb{C}$WT with Finite Impulse Response FIR-filters, this condition is approximately fulfilled. In contradiction to classical DWT it does not suffer from aliasing and oscillations, purveys a high degree of shift-invariance magnitude and orientation within two and higher dimensions, which are the properties desired for pattern and texture classification and recognition. It exhibits redundancy of 2^d (where d is the dimension of the signal being transformed), which is significantly lower than nondecimated DWT.

This paper uses the dual-tree approach for complex wavelet transform (DT$\mathbb{C}$WT), which calculates the complex transform of a signal using two separate real DWT-decompositions (tree a and tree b). If the filters used in tree a are specifically designed to be different from those in tree b, then it is possible for one DWT to produce the real coefficients and other the imaginary. More exactly, to really achieve the desired analytical properties, the filters in the trees are interchanged at each decomposition level and during the first stage any perfect reconstruction filters can be used within each tree. In two dimensions (2D), six 2D ODT$\mathbb{C}$WT can be achieved by properly combining (sum and difference) two separable transform's subbands (two HL, two LH, and two HH) for each level of decomposition. The details regarding the implementing a 2D ODT$\mathbb{C}$WT can be found in [29].

The 2D $\mathbb{C}$WT provides multiresolution, sparse representation, and useful categorization of an image's structure. In the area of computer vision, by exploiting the concept of visual contexts, one can quickly focus on candidate regions where objects of interest may be found, and then additional features computed through the $\mathbb{C}$WT for those regions only. These additional features, while not necessary for global regions, are useful for the accurate detection and recognition of smaller objects. $\mathbb{C}$WT may be applied for detecting the activated voxels of cortex and temporal independent component analysis may be additionally utilized for extracting the underlying independent sources the number of which is determined by Bayesian information criterion [29].

Scene classes from SAR-image patches can be determined using different descriptors (feature vectors). Descriptors within the wavelet domain are very appropriate and a natural choice due to the multiresolution properties of wavelet transform and on the other hand, the deterministic or stochastic self-similarity properties of many natural and urban amplitude SAR-image's scenes. Different wavelet transforms can be used from continuous (CWT) to DWT [30] and discrete wavelet packets transform (DWPT) [31]. DWPT can be considered as the generalization of DWT, where the detailed wavelet coefficients are also further decomposed with low-pass and high-pass filters. It was discovered in [32] that the DWPT tree for amplitude SAR-images reduces to a DWT (Laplacian pyramid) using Shannon's entropy criteria for any decision on further decomposition. As will be shown later, TerraSAR-X images with their high-resolution properties enable the classification of scenes, which on bigger regions and globally look similar or almost equal. The classification efficiency is improved by using feature vectors with descriptors within the DWT domain.

Nonparametric feature extraction is performed using an image patch of $M \times M = 200 \times 200$ pixels. The descriptors were evaluated for each subband of the (ODT$\mathbb{C}$WT) with j decomposition levels. The patch with dimension $M \times M$ pixels was characterized by the feature vector, with descriptors for all subbands $S_{j,i}$. Subband $S_{j,1}$ is obtained by low-pass filtering in both the horizontal and vertical directions (LL), subband $S_{j,2}$ by low-pass filtering in the horizontal direction and high-pass filtering in the vertical direction (LH), subband $S_{j,3}$ by high-pass filtering in both direction (HH), and subband $S_{j,4}$ by high-pass filtering in the horizontal and low-pass filtering in the vertical direction (HL).

The feature vector within the wavelet transform consists of subbands variances $\delta_{j,i}^2$, mean values $\mu_{j,i}$, $L2$-norm $\|S_{j,i}\|_2$ and $L1$-norm $\|S_{j,i}\|_1$, entropy and coding gain. These parametric features were also presented in [42].

A. Mean and Variance of Subband

Thus, the variance of subband $S_{j,i}$ with dimension $M_j \times M_j$ is given by

$$\delta_{j,i}^2 = \frac{1}{M_j \cdot M_j} \sum_{m=0}^{M_j-1} \sum_{n=0}^{M_j-1} \left(S_{j,i}[m,n] - \mu_{j,i}\right)^2 \qquad (1)$$

where $S_{j,i}[m,n]$ represents a wavelet coefficient with indexes m, n, and μ_j is the mean of the subband given by

$$\mu_{j,i} = \frac{1}{M_j \cdot M_j} \sum_{m=0}^{M_j-1} \sum_{n=0}^{M_j-1} S_{j,i}[m,n]. \qquad (2)$$

B. First- and Second-Order Norm of Subband

The norm of order 2 is given by

$$\|S_{j,i}\|_2 = \frac{1}{M_j \cdot M_j} \sum_{m=0}^{M_j-1} \sum_{n=0}^{M_j-1} |S_{j,i}[m,n]|^2 \qquad (3)$$

and the norm of order 1 is given by

$$\|S_{j,i}\|_1 = \frac{1}{M_j \cdot M_j} \sum_{m=0}^{M_j-1} \sum_{n=0}^{M_j-1} |S_{j,i}[m,n]|^1. \qquad (4)$$

C. Entropy

From the information point of view, the entropy of a subband is the measure of information contained within that subband.

There are many types of entropies; however, the more common is Shannon's entropy given by

$$E_{j,i} = \sum_k p_k \log_2(1/p_k) \tag{5}$$

where p_k is the probability of the pixels with value S_k occurring within subband $S_{j,i}$.

D. A Kolmogorov Sinai (K-S) Entropy

A Kolmogorov Sinai (K-S) entropy [33] is an extended version of information entropy. The K-S is based on the modeling of sequential probability, entropy rate, and limiting conditions. Modeling of the probabilities of a sequence is estimated using a window space and the entropy is given by

$$H_{\Delta\beta} = -\sum_{s=1}^N p_s \log p_s \tag{6}$$

where $H_{\Delta\beta}$ is the entropy computed over a window space, p_s denotes the sequence of probabilities in the sth phase-space, and N is the number of possible phase-spaces routes. The sequence probabilities are the likelihoods that the state of the system will be taking various possible routes.

The average rate of K-S entropy is the average rate of the creation of new information, which is represented by the entropy per unit space $H_{\Delta\beta}/w$. The K-S entropy is computed within the limit, where the box size ϵ goes toward zero and the space w grows to infinity. The K-S entropy is given by

$$
\begin{aligned}
H_{K-S} &= \lim_{\epsilon \to 0} \lim_{w \to \infty} \left(\frac{H_{\Delta\beta}}{w} \right) \\
&\lim_{\epsilon \to 0} \lim_{w \to \infty} \frac{-\sum_{s=1}^N p_s \log p_s}{w} \\
&\lim_{\epsilon \to 0} \lim_{w \to \infty} (H_\beta - H_{\beta-1}).
\end{aligned} \tag{7}
$$

The K-S entropy is computed for each subband of the oriented wavelet transform, using window sizes of 16×16 pixels, and difference of entropies between two neighboring windows representing a subband feature.

E. Coding Gain

An another important descriptor is coding gain [34], which can be estimated using the flatness of the wavelet spectrum. A coding gain also represents a patch smoothness and is estimated using variances of subbands within the DWT domain using the expression

$$cg = \frac{\prod_{j=1}^J \prod_{i=1}^3 \alpha_{j.i} \cdot \sigma_{j,i}{}^2}{\prod_{j=1}^J \prod_{i=1}^3 (\sigma_{j,i}{}^2)^{\alpha_{j,i}}} \tag{8}$$

where $a_{j,i} < 1$ is the relative area of the subband with subscripts j, i. Note that the sum of all relative areas is equal to one.

Nonparametric features (1)–(5) and (7) were ordered into a descriptor $\boldsymbol{d}$ given by

$$d_j = [\delta_{j,i}^2, \mu_{j,i}, \|S_{j,i}\|_2, \|S_{j,i}\|_1, E_{j,i}, K_{K-S,j}] \tag{9}$$

where j represented the decomposition level and i belonged to the horizontal, vertical, or diagonal details of wavelet decomposition.

F. Outline of Nonparametric Feature Extraction

A nonparametric feature extraction is very simple. For each subband of the 2D ODTℂWT transformed image, a feature vector, given in (9), was computed and those parameters were categorized using an SVM.

III. PARAMETRIC FEATURES

In this paper, parametric or model-based methods model a scene using a theoretical probability density function. The parametric-based methods adapt their model parameters to the scene's content. The model parameters usually correspond to the physical parameters of the scene. The Markov random field models used in this paper, measure the interaction of neighboring pixels within an image. In [13], it was shown that the synthetically generated scene using estimated texture parameters corresponded well to the real SAR scene. The MRF-based methods describe scene content on a pixel-based level; therefore, each pixel is described by a set of parameters. In this paper, we used the MRF models, which will describe a content within a patch using a single parameter.

The prior pdf was theoretical modeled in this paper using a GMRF, triplet Gauss–Markov random field (TGMRF), and an ABM. All three models belong to the family of Markov random fields and they model a scene using model parameters, which represent synthetic scene parameters. Bayesian inferences of first and second orders were used for scene modeling [13], [14], [35]. The first order of Bayesian inference maximizes the product of a likelihood $p(y|x)$ and prior $p(x)$ and represents a maximimum *a posteriori* (MAP) estimate. The likelihood models speckle using a theoretical Nagakami distribution and a prior is determined with an MRF model. The likelihood is given by

$$p(y_s|x_s) = 2\left(\frac{y_s}{x_s}\right)^{2L-1} \frac{L^L}{x_s \, \Gamma(L)} \exp\left(-L\left(\frac{y_s}{x_s}\right)^2\right) \tag{10}$$

where y is the noisy pixel, x is an estimated speckle-free pixel, and L is the equivalent number of looks. A GMRF prior is a texture model, which represents an autoregressive process. All the models used in this paper are able to analytically describe scene contents. The GMRF pdf is given by

$$p(x_i|\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\left(x_i - \sum_{j \in N_s} \theta_j x_j\right)^2}{2\sigma^2}\right) \tag{11}$$

where $\boldsymbol{\theta}$ weights the interaction of neighboring pixels and represents texture parameters. The details can be found in [14]. In this paper, we extended the GMRF model to a conditional tripled GMRF, where any information regarding the pdf of the observed image within a window is added to the prior, which models the mutual influences between neighboring pixels. The TGMRF is given by

$$p(x_s|\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\sigma^2\sigma_\theta^2}}$$

$$\times \exp\left(-\frac{\left(x_s - \sum\limits_{j \in N_s} \theta_j x_j\right)^2}{2\sigma^2} - \frac{(x_s - \mu_\theta)^2}{2\sigma_\theta}\right)$$

$$\cdot g_s(x_s, \theta_s) \tag{12}$$

where g_s is a term reflecting interactions between observations given by

$$g_s(x_s, \theta_s) = \exp\left(-\frac{d_x^2}{\sigma} - \frac{d_\theta}{\sigma_\theta}\right) \tag{13}$$

$$d_x = \|x_{s+r} + x_{s-r} - (x_{s+r+1} + x_{s-r-1})\| \tag{14}$$

$$d_\theta = \|\theta_{s+r} - \theta_{s+r+1}\| \tag{15}$$

μ_θ and σ_θ represent mean value and standard deviation of estimated texture parameters. d_x and d_θ represent Euclidian distances of gray values and texture features of a site s. The parameter g_s is large when neighboring sites have similar features and intensities.

The TMRF [36] was used for classification, but in this paper, we modified a model in order to use it for despeckling and estimation of texture features. The model considers the observed pixel x_s and estimated texture parameters. The posterior is modeling an observed pixel using neighborhood of pixels, which are weighted with texture parameter θ. The texture parameter θ depends on distribution of pixels; therefore, the model considers conditional density of all pixels.

The third model used in this paper is an ABM [37], which represents binomial Markov random fields. This model can be used for the modeling of binary textures as well as gray scale textures [37]. In this paper, we used an ABM for modeling grayscale data. The ABM is given by

$$p(x_s|\theta) = \binom{G}{x_s} \rho_s^{x_s}(1 - \rho_s)^{G - x_s} \tag{16}$$

where interactions between sites are described with parameters ρ and η given by

$$\rho_s = \frac{1}{1 + \exp(-\eta_s)} \tag{17}$$

$$\eta_s = a_s + \sum_{r \in N_s} b_r (x_{s+r} + x_{s-r})/G \tag{18}$$

$$\theta = \left[a, \underbrace{b_{11}, b_{12}}_{1st}, \underbrace{b_{21}, b_{22}}_{2nd}, \underbrace{b_{31}, b_{32}}_{3rd}, \underbrace{b_{41}, b_{42}, b_{43}, b_{44}}_{4th}, \ldots\right]. \tag{19}$$

5	4	3	4	5
4	2	1	2	4
3	1	x_s	1	3
4	2	1	2	4
5	4	3	4	5

Fig. 1. Example of a neighborhood of model order 5.

The parameter θ describes the ABM parameters, where the sum of all b coefficients should be equal to the first parameter a. Fig. 1 shows a window of 5×5 pixels where a model order 5 is depicted. Coefficients b_{11} and b_{12} weight interactions of horizontal and vertical pixels marked as number 1, coefficients b_{21} and b_{22} weight diagonal interactions, etc.

In order to estimate the scene content using texture parameters, a second-order Bayesian inference is applied to the Bayesian rule, given by

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)}. \tag{20}$$

The first-order Bayesian inference is given by

$$\hat{x} = \arg\max p(y|x)p(x). \tag{21}$$

The GMRF, TGMRF, and ABM priors have a texture parameter θ, which represents texture information. The parameter θ is estimated using image content. A two-stage estimation of parameter θ was proposed in [14], by computing an MAP estimate and maximizing evidence in order to find the best texture parameter, which optimally fits the data.

The MAP estimates for three different priors GMRF, TGMRF, and ABM are given by (22)–(24)

$$x^4 + x^3 \sum_{j \in N_s} \theta j x_j + 2L\sigma^2 - 2L\sigma^2 y^2 = 0 \tag{22}$$

$$x^4\left(\sigma^2 + \sigma_\theta^2\right) + x^3\left(\sum_{j \in N_s} \theta_j x_j + \mu_\theta \sigma_\theta^2\right)$$

$$+ 2Lx^2\left(\sigma^2 + \sigma_\theta^2\right) - 2Ly^2\left(\sigma^2 + \sigma_\theta^2\right) = 0 \tag{23}$$

$$-\frac{2L}{x_s} + 2L\frac{y_s^2}{x_s^3} + \log\left(\frac{G - x_s}{x_s + 1}\right) + \log\left(\frac{\rho_s}{1 - \rho_s}\right) = 0. \tag{24}$$

The MAP estimate also represents a despeckled image. In order to find the best parameters of the model, the second-order Bayesian inference deals with model selection. The idea was to iteratively change the model's parameters and evaluate the evidence $p(y|\theta) = \int p(y|x)p(x|\theta)dx$. MAP estimates given in (22)–(24) were found using a Brent's algorithm [38]. The evidence depends on MAP estimate, which produces an estimate of speckle free image pixel x and texture parameter θ. The best

Algorithm 1. Parametric Feature Extraction.

Compute ENL
Set windows $N \times N = 200 \times 200$ and $M \times M = 200 \times 200$
Set model order mo and number of iterations
Set initial values for $\theta = 0.5/model\_par[mo]$
repeat
  repeat
    Compute MAP estimate for GMRF (22) or TGMRF (23) or ABM (24)
    Compute θ. Estimate evidence (25) using selected model ((26) for GMRF, (27) for TGMRF or (28) and (29) for ABM).
    Use Amoeba algorithm for texture estimation and iterate until evidence grows
    Compute MAP estimate using estimated θ
    within window $N \times N$ write estimated speckle free pixels $\hat{x}_s$ and estimated texture parameters $\hat{\theta}$
    Next window
  until end of an image
until end of iterations

model is chosen by selecting parameter θ, which gives the maximal evidence.

The outline of the algorithm is depicted in Algorithm 1. The user chooses size of window, the model, model order (MO), and sets initial texture parameters. The algorithm computes MAP estimates and θ parameters within selected windows. The algorithm shows that the parameter θ is iteratively changed while maximizing evidence. The parameter θ at the highest evidence is then used for MAP estimation. It means that the prior's parameters are changed until the best model is found and then the MAP estimate is computed using estimated parameter θ for the observed pixels within a window of $N \times N$.

A. Evidence Maximization

The computation of the evidence integral was not straightforward but was estimated using a Hessian approach [14] and is given by

$$
\begin{aligned}
\log p(y|\theta) &\approx \frac{1}{2}(M \log 2\pi - \log |\mathbf{H}|) \\
&+ \sum_{i=1}^{N} \log p(y|\hat{x}_i) + \log p(\hat{x}_i|\theta) \\
&\approx \sum_{i=1}^{N} \frac{1}{2}(\log 2\pi - \log h_{ii}) + \log p(y_i|\hat{x}_i) \\
&+ \log p(\hat{x}_i|\theta)
\end{aligned}
\tag{25}
$$

where h_{ii} are the components of matrix $\mathbf{H}$ and M represents the number of pixels inside the window. The matrix $\mathbf{H}$ is computed using the second derivative of the product of likelihood and prior. The GMRF, TGMRF, and ABM have different prior models; therefore, their computations of the evidence are numerically different. In (25), $\hat{x}_i$ represents an

estimated speckle-free pixel. The Hessian approximation h_{ii} for the GMRF is given by

$$
h_{iiGMRF} = \frac{6Ly_s^2}{x_s^2} - \frac{2L}{x_s^2} + \frac{1}{\sigma^2}\left(1 + \sum_{j \in \zeta} \theta_j^2\right).
\tag{26}
$$

The evidence in the case of TGMRF is given by

$$
h_{iiTGMRF} = \frac{6Ly_s^2}{x_s^2} - \frac{2L}{x_s^2} + \frac{1}{\sigma^2}\left(1 + \sum_{j \in \zeta} \theta_j^2\right) + \frac{\mu}{\sigma_x^2}.
\tag{27}
$$

The evidence for auto-binomial prior is given by

$$
\begin{aligned}
\log p(\mathbf{y}|\boldsymbol{\theta}) = \sum_{i=1}^{M} &\left[\frac{1}{2}\left(M \log 2\pi - \log h_{ii}\right) \right.\\
&+ \log\left(2\left(\frac{y_i}{x_i}\right)^{2L-1} \frac{L^L}{x_i \Gamma(L)} \exp\left(-L\left(\frac{y_i}{x_i}\right)^2\right)\right) \\
&\left.+ \log\left(\binom{G}{x_i} \rho_i^{x_i}(1-\rho_i)^{G-x_i}\right)\right]
\end{aligned}
\tag{28}
$$

where h_{ii} is the approximated component of Hessian matrix

$$
\begin{aligned}
h_{iiABM} \approx \sum_{i=1}^{M} &\left(\frac{6Ly_i^2}{x_i^{4MAP}} - \frac{2L}{x_i^{2MAP}} + \frac{1}{G - x_i^{MAP}} \right.\\
&\left. + \frac{1}{x_i^{MAP} + 1} \right).
\end{aligned}
\tag{29}
$$

The algorithm has two steps. Within the first step, an MAP estimate is estimated using the initial texture parameters. Within the second step, an iterative procedure is applied in order to achieve maximal evidence. The maximal evidence is achieved by changing the texture parameters. The amoeba algorithm [38] is applied when searching for a global maximum of the evidence. The evidence is computed within a window of $M \times M$ pixels and is the most computational demanding part of the algorithm. The adaptation of the existing algorithms is automatic MO selection and using a window size which corresponds to the site of the image patch.

IV. Experimental Results

A. Database

A database of 2000 patches was generated within an image size of 200×200 pixels using SLC VHR Spotlight TerraSAR-X images. Each patch covered approximately 200 m^2 of ground. This size could be generally used for defining a particular category, as it is large enough to contain the contextual information needed for defining an object's structure in the case of HR, yet is also suitable for the texture computation of homogeneous areas. The presented experimental database was composed of 20 well-defined object/texture categories with 100 patches in each category. Care was taken to generate a database with diverse textures. Examples of one patch from each category are shown in Fig. 2. These
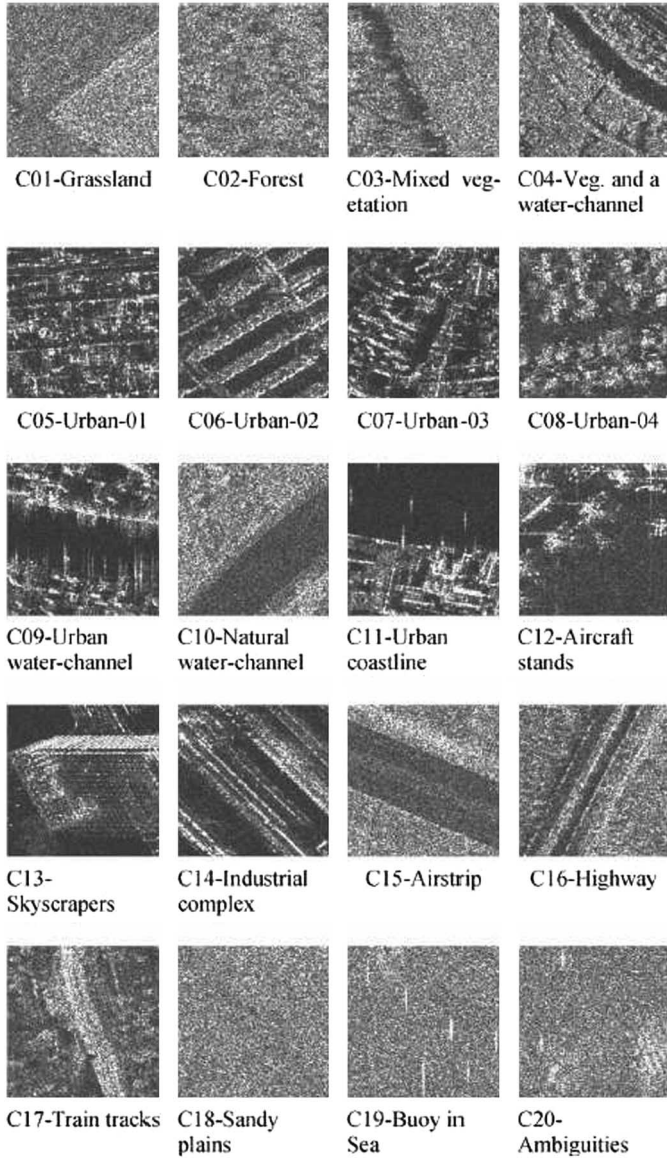
Fig. 2. Examples of the SAR image patches (of size 200×200 pixels) in different object/texture categories. The image patches have been obtained from SLC VHR Spotlight TerraSAR-X images from different areas in order to include the diversity of objects/textures when analyzing feature descriptors.

categories included land-cover topologies (C01-Grassland, C02- Forest, C03-Mixed vegetation, C04-Vegetation and a water channel, and C18-Sandy plains), urban areas with a particular texture (C05-Urban-01, C06-Urban-02, C07-Urban-03, and C08-Urban-04), urban and natural water channels (C09-Urban water-channel and C10-Natural water-channel), regions with strong scatterers along both surfaces exhibiting specular reflection (C11-Urban coastline and C12-Aircraft stands), regions with very strong scatterers with multiple bounces (C13-Skyscrapers and C14-Industrial complex), and other special categories like C15-Airstrip, C16-Highway, and C17-Train tracks. Two categories from the sea-water region (C19-Buoy in Sea and C20-Ambiguities) were also considered. This database was generated manually; thus, it could be used as a ground truth for the assessment of various algorithms.

B. Assessment

This section exploits the efficiencies of nonparametric features within wavelet transform, as presented in Section IV-B1, and parametric features within the image domain, as presented in Section IV-B2.

1) Nonparametric Features: A feature vector given in (9) was used for image patches categorization. The coding gain was estimated for a SAR patch and was added as a single parameter at the end of a feature vector. Three levels of dyadic decomposition were used ($j = 3$); therefore, d consisted of $3 \cdot ((6 \cdot 12)) + 1 = 217$ features within ODT$\mathbb{C}$WT transform.

The proposed nonparametric features were compared to the gray-level cooccurrence matrix (GLCM) features. The second-order statistics of the amplitude SAR image patches were computed using GLCMs, extracting 12 textural features and used for a comparative assessment. This implementation allowed four options for the orientation (0., 45., 90., and 135.). Each GLCM feature descriptor was formed with 12 parameters based on the normalized matrix: mean, variance, entropy, angular second moment, energy, correlation, maximum probability, contrast, homogeneity, dissimilarity, cluster shade, and cluster prominence. This provided a GLCM feature descriptor of length 48. A spectral descriptor based on a fractional Fourier transform and LC was applied to this database [20], where 27 features were used.

The presented parametric and nonparametric methods were compared with the LBP [26] and MR8 [27] methods for texture description. LBP is a method, which observes a neighbor pixels and codes a neighborhood regarding intensity of pixels. Neighborhoods $LBP^{riu2}_{8,1}$, $LBP^{riu2}_{8,1+16,2}$, $LBP^{riu2}_{8,1+16,2+24,3}$, and $LBP^{riu2}_{8,1+16,2+24,3+24,4}$ were considered, where subscript denotes a type of a neighborhood according [26], where (8,1), (16,2), and (24,3) denote the use of radii $R = 1, 2, 3$ with neighboring samples $N = 8, 16, 24$. $riu2$ denotes rotation- invariant pattern. The MR8 method consists of eight filter responses derived from the original responses of 38 filters [27].

Table I shows the efficiency of the proposed methods using nonparametric feature extraction. The efficiency of the class recognition is measured in percentage of the recognized images within a specific class. 10% and 20% randomly chosen image patches represented 200 and 400 image patches, which were used only at the learning stage. Those image patches were not used for recognition; therefore, 1800 and 1600 image patches were used within the recognition procedure, respectively. Supervised learning using an SVM [39] with a multiclass classifier [25] was used for all experiments in this paper. A multiclass SVM was implemented in MATLAB using a polynomial kernel, sequential minimal optimization (SMO) method for separating hyperplane, and allowing maximum 1500 iterations for SMO. A package can be found in [40]. Table I reports the results for those nonparametric features obtained using ODT$\mathbb{C}$WT. The average recognition rates for the wavelet transform-based methods were generally better when compared to the GLCM method [41] and LC [20] of the fractional Fourier transform methods. The DWT-based patch categorization achieved 83.8% of recognition rate for the used database, when 10% of the data were used at the learning stage and 93.4% when 20% of

TABLE I
ACCURACY OF CLASS RECOGNITION IN % FOR NONPARAMETRIC FEATURES USING 10% AND 20% OF IMAGES FOR LEARNING STAGE

Class	Nonparametric features					
	ODT$\mathbb{C}$WT		GLCM	LC [20]	LBP [26]	MR8 [27]
	10%	20%	20%	20%	20%	20%
C1	85.2 ± 0.13	**94.2 ± 0.1**	39.8 ± 0.1	92.5 ± 0.21	92.8 ± 0.11	93.8 ± 0.13
C2	86.9 ± 0.13	**90.8 ± 0.12**	60.6 ± 0.5	90.7 ± 0.23	89.5 ± 0.12	90.4 ± 0.11
C3	84.0 ± 0.12	88.6 ± 0.11	34.1 ± 0.7	81.9 ± 0.14	88.5 ± 0.13	**92.4 ± 0.14**
C4	92.3 ± 0.14	**96.1 ± 0.13**	49.2 ± 0.012	95.4 ± 0.12	93.3 ± 0.14	95.3 ± 0.16
C5	90.1 ± 0.13	**98.0 ± 0.02**	58.9 ± 0.6	95.4 ± 0.14	94.2 ± 0.14	93.9 ± 0.18
C6	78.5 ± 0.16	86.4 ± 0.12	44.5 ± 0.8	**95.7 ± 0.16**	82.6 ± 0.16	87.3 ± 0.13
C7	87.4 ± 0.12	**95.3 ± 0.09**	37.3 ± 0.6	80.8 ± 0.14	92.4 ± 0.15	83.4 ± 0.14
C8	90.1 ± 0.11	**95.3 ± 0.12**	53.8 ± 0.6	89.7 ± 0.15	94.1 ± 0.17	88.4 ± 0.17
C9	87.4 ± 0.12	**95.6 ± 0.13**	50.8 ± 0.6	87.5 ± 0.12	93.2 ± 0.14	87.9 ± 0.11
C10	93.8 ± 0.14	**100 ± 0.12**	65.4 ± 0.4	87.0 ± 0.11	95.8 ± 0.15	97.3 ± 0.18
C11	85.3 ± 0.18	**95.1 ± 0.13**	71.1 ± 0.3	94.5 ± 0.16	93.2 ± 0.16	92.8 ± 0.15
C12	86.7 ± 0.15	**95.5 ± 0.11**	70.7 ± 0.2	78.0 ± 0.13	92.7 ± 0.18	93.7 ± 0.13
C13	87.7 ± 0.14	**94.9 ± 0.13**	36.6 ± 0.3	79.8 ± 0.11	93.3 ± 0.12	93.4 ± 0.18
C14	84.4 ± 0.23	93.0 ± 0.09	30.4 ± 0.6	67.5 ± 0.15	92.1 ± 0.11	**93.8 ± 0.16**
C15	82.8 ± 0.23	**87.9 ± 0.11**	43.4 ± 0.6	63.2 ± 0.11	87.2 ± 0.14	85.2 ± 0.18
C16	86.8 ± 0.23	94.2 ± 0.11	40.9 ± 0.5	93.2 ± 0.12	**96.5 ± 0.13**	93.1 ± 0.16
C17	93.7 ± 0.131	95.7 ± 0.06	47.3 ± 0.5	76.4 ± 0.13	**95.8 ± 0.15**	94.1 ± 0.12
C18	87.3 ± 0.13	90.8 ± 0.11	50.83 ± 0.4	89.0 ± 0.15	89.6 ± 0.12	**91.1 ± 0.13**
C19	92.5 ± 0.12	95.3 ± 0.05	89.3 ± 0.4	94.3 ± 0.18	94.5 ± 0.13	**95.6 ± 0.11**
C20	75.9 ± 0.12	85.5 ± 0.11	47.96 ± 0.3	96.9 ± 0.13	93.5 ± 0.21	**95.1 ± 0.1**
Total	83.8 ± 0.16	**93.4 ± 0.13**	51.4 ± 0.3	86.2 ± 0.14	92.2 ± 0.14	92.1 ± 0.14

This table represents the accuracy of features within ODT$\mathbb{C}$WT, GLCM, and LC.

data were used at the learning stage. All classes had very similar recognition rates (approx. between 80% and 90%). Classes C6-Urban-02, C15-Airstrip, and C20-Ambiguities had lower recognition rates. The GLCM features gave very low recognition rate, which varied between 30% and 90% and the average recognition rate was 51.4%. The LC of fractional Fourier transform gave accuracies between 60% and 97% and on average 86.2%. The LC badly recognized categories C14-Industrial Complex, C15-Airstrip, and C17-Traintracks. The LBP method gave the best results for classes C-16-Highway and C-17-Train tracks. The average performances of the LBP method gave 1.2% lower results compared with the ODT$\mathbb{C}$WT method and all classes were estimated with accuracies between 88% and 96%. The standard deviation of all classes for the nonparametric methods was between 0.11 and 0.21. The MR8 method gave the best results results for classes C3-Mixed vegetation, C14-Industrial Complex, C18-Sandy plains, C-19 Buogy in Sea, and C20-Ambiguities. The overall accuracy of the MR8 method was 92.1%. All the presented methods gave very similar results regarding accuracies, when 20% of the test data were used. The best results were obtained with the ODT$\mathbb{C}$WT, followed by the MR8, LBP, LC, and GLCM methods.

2) Parametric Features: The GMRF, TGMRF, and ABM algorithms require determination of a window size for evidence computation, and the number of iterations needed for the amoeba algorithm, which is an iterative algorithm used for estimating of texture parameters. The window size of the proposed algorithms was set experimentally to 31×31 pixels. The average convergence of the amoeba algorithm was 453 iterations. Areas with few strong scatters usually caused problems for converging of the amoeba algorithm. Texture parameter estimation is very important during the MAP estimation process since it determines the scene features. Fig. 3(a) shows the original SAR image. The MAP estimated SAR images obtained using the GMRF model are shown in Fig. 3(b) and (c). In Fig. 3(b),

the texture parameters, as shown in Fig. 3(e)–(f), converged when using the amoeba algorithm. An MAP estimated image, as shown in Fig. 3(c), had texture parameters obtained using an evidence maximization algorithm, but they were obtained with an insufficient number of iterations of amoeba algorithm. The parameters did not converge, as can also be noticed in the texture parameters, shown in Fig. 3(h)–(j). The MAP estimated image, as shown in Fig. 3(c) also has distortions, which represent wrongly estimated texture parameters. Fig. 3(d) shows the MAP estimated image using an ABM and the corresponding texture parameters are shown in Fig. 3(k)–(m). All the texture parameters were obtained using a model order 4, a window size of 31×31 pixels for evidence estimation, and texture parameters representing horizontal, vertical, and diagonal interactions between the second MO neighbors.

The other very important parameter is the MO of an MRF, which was set experimentally. Fig. 4(a) and (b) shows the MAP estimated images using the GMRF model orders of 4 and 8, respectively. The MO of MRF defines their abilities for estimating complex textures. The difference within MAP estimated images using model orders 4 and 8 is within the textured area, but the differences were minor, as could be noticed within the ratio images, as shown in Fig. 4(c) and (d). In this paper, the model-based approach for patch categorization was performed by manually selecting MRF MOs. The parameters of the patch using a window size, which corresponded to the image size were used for SAR patch categorization. The MO was set at 4, 6, and 10 and supervised learning was used by using a multiclass SVM.

Table II summarizes the recognition rates of the parametric features using GMRF, TGMRF, and ABM. Model orders 4, 8, and 10 were used for all the presented parametric models. The window size of texture parameter estimation was set to 200×200 in order to characterize an image with a single texture parameter. The overall recognition rate of parametric
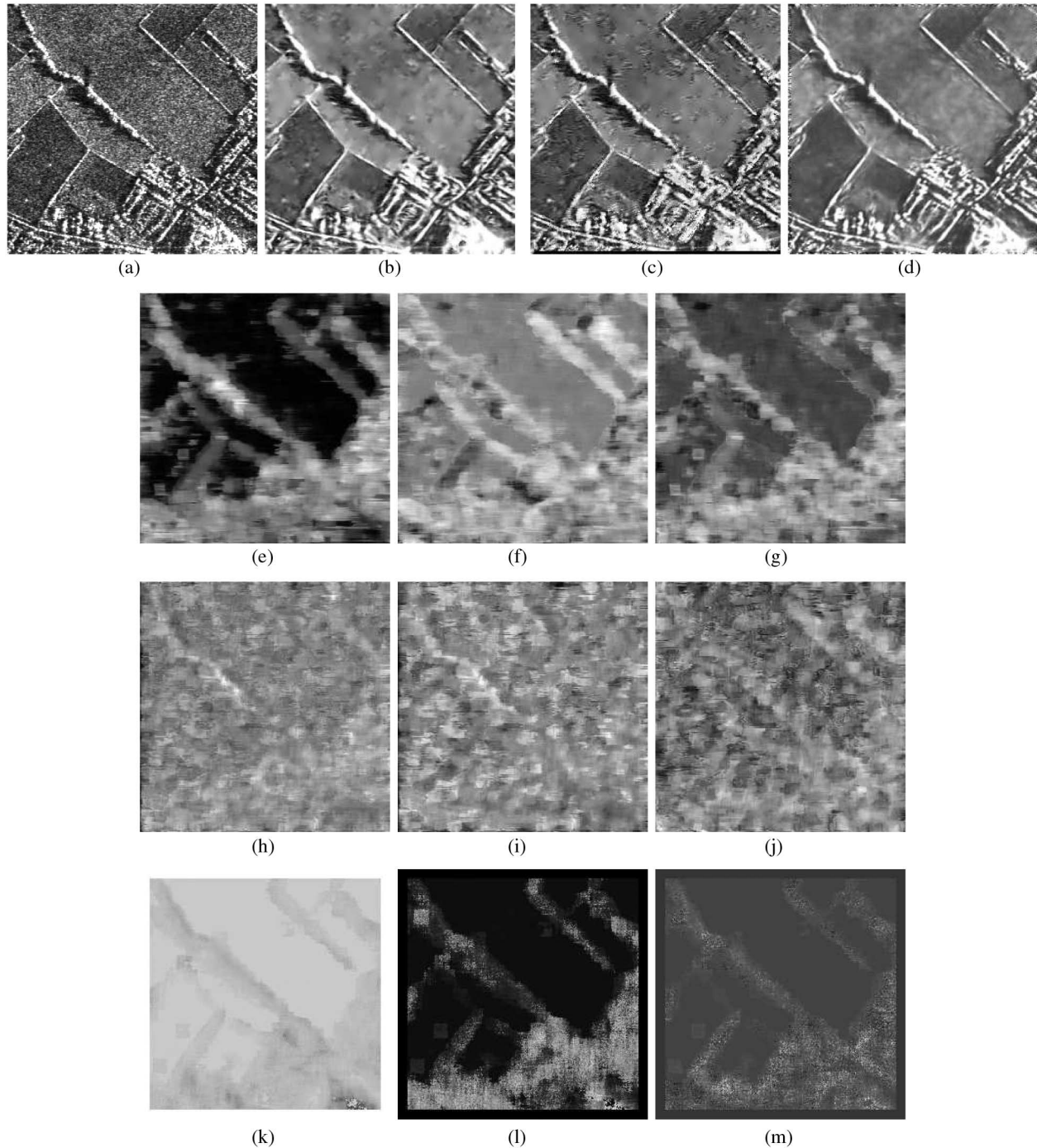
Fig. 3. (a) Original SAR image. (b) Despeckled with MAP-GMRF with converged texture parameters. (c) Despeckled using MAP-GMRF where parameters did not converge. (d) Despeckled using MAP-ABM. Converged texture parameters of MAP-GMRF (e) horizontal, (f) vertical, and (g) diagonal texture parameters. (h)–(j) Texture parameters of the MAP-GMRF that did not converge. (k)–(m) Converged texture parameters of MAP-ABM.

features was smaller compared to the parametric approach. The ABM was able to recognize patch scenes and had a rate of recognition similar to the nonparametric features. The performance of the GMRF-based patch characterization increased by increasing the MO. The homogeneous areas could be well recognized when using model orders 2–4 but the complex texture structures were more difficult to categorize; therefore, MOs greater than 6 should be used for patch categorization. The MRF models are scale and rotationally variant, thus causing the order of parameters to be changed. The GMRF model order 4

gave accuracies of 70.7% and 77.8% when 10% and 20% of images within a database were used for supervised learning. Most of the classes were recognized with accuracies between 60% and 70%. The recognition rate increased by increasing the number of training samples. In Table II, the recognition rates above average are marked with bold. The classes C02-Forest, C04-Vegetation and a water channel, C6-Urban-02, C7-Urban-03, C11-Urban and Coastline, and C13-Skyscrapers had recognition below the average rate for all GMRF MOs. The best accuracy of the GMRF model was achieved using model order
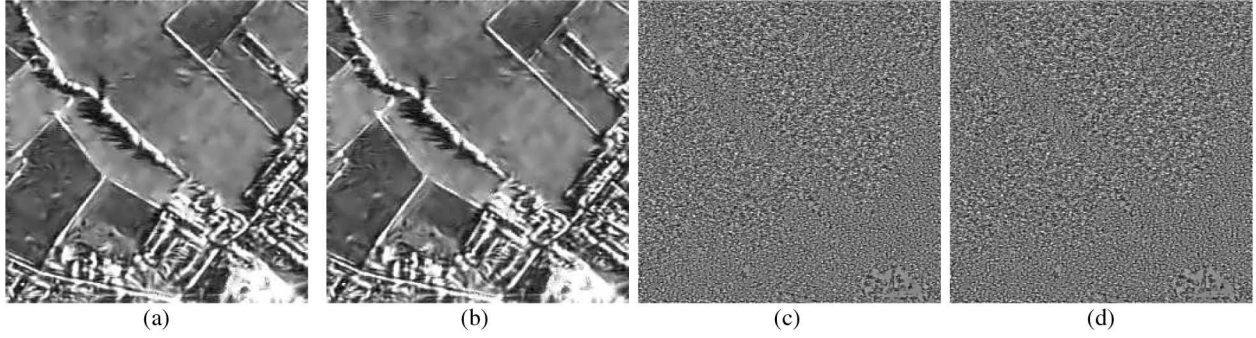
Fig. 4. (a) Despeckled with MAP-GMRF using model order 4. (b) Despeckled with MAP-GMRF and model order 8. (c) Ratio image for MAP-GMRF using model order 4. (d) Ratio image for MAP-GMRF using model order 8.

TABLE II
ACCURACY OF CLASS RECOGNITION IN % FOR PARAMETRIC METHODS USING 10% AND 20% OF IMAGES AT LEARNING STAGE

Class	Parametric features																	
	GMRF						TGMRF						ABM					
	MO4		MO6		MO10		MO4		MO6		MO10		MO4		MO6		MO10	
	10%	20%	10%	20%	10%	20%	10%	20%	10%	20%	10%	20%	10%	20%	10%	20%	10%	20%
C1	66.5±0.22	77.4±0.21	68.6±0.23	76.6±0.29	71.3±0.21	87.2±0.32	64.4±0.25	74.8±0.24	66.8±0.23	73.4±0.21	69.3±0.23	82.6±0.24	71.3±0.21	76.3±0.23	77.3±0.13	80.9±0.16	81.6±0.14	84.4±0.15
C2	75.9±0.3	85.1±0.21	79.6±0.23	82.6±0.22	90.5±0.23	91.7±0.23	76.6±0.25	83.1±0.27	75.4±0.21	81.4±0.23	81.7±0.21	89.9±0.2	77.3±0.24	87.6±0.24	76.9±0.12	78.8±0.14	80.1±0.35	86.8±0.15
C3	70.2±0.31	77.3±0.22	70.1±0.13	72.2±0.22	79.3±0.22	75.2±0.33	68.5±0.31	76.2±0.24	72.1±0.23	76.5±0.21	74.7±0.24	82.3±0.22	76.9±0.18	81.8±0.19	79.9±0.38	85.1±0.18	87.4±0.19	96.4±0.13
C4	65.3±0.3	73.3±0.12	66.6±0.13	72.3±0.14	69.4±0.15	79.9±0.13	66.3±0.42	72.6±0.31	68.5±0.41	75.0±0.25	77.8±0.24	83.8±0.14	82.3±0.19	87.9±0.13	86.3±0.13	92.6±0.12	93.3±0.14	94.4±0.14
C5	76.5±0.19	80.6±0.24	80.8±0.24	84.4±0.28	85.4±0.15	87.9±0.32	74.3±0.23	84.1±0.23	75.5±0.3	77.8±0.45	86.5±0.18	93.8±0.12	83.8±0.13	86.9±0.21	89.6±0.22	97.9±0.09	95.8±0.23	96.6±0.12
C6	60.3±0.23	68.8±0.29	70.5±0.28	75.7±0.24	76.1±0.22	80.5±0.21	60.5±0.24	69.5±0.23	72.6±0.24	76.5±0.24	79.3±0.25	86.9±0.15	72.8±0.19	78.1±0.17	76.3±0.22	83.3±0.12	79.3±0.21	93.5±0.11
C7	66.3±0.31	71.2±0.22	68.1±0.25	76.1±0.24	75.3±0.25	79.2±0.26	66.5±0.25	77.0±0.23	62.3±0.26	66.9±0.29	64.4±0.32	72.1±0.25	78.6±0.21	88.5±0.25	78.6±0.12	80.6±0.22	77.7±0.25	90.3±0.12
C8	69.1±0.25	78.9±0.14	76.4±0.15	77.8±0.22	85.2±0.12	87.5±0.11	70.2±0.23	79.7±0.13	75.6±0.14	78.0±0.14	81.3±0.13	90.4±0.10	88.3±0.11	96.3±0.10	90.3±0.11	98.3±0.09	91.7±0.12	97.8±0.12
C9	76.3±0.16	81.7±0.15	78.8±0.25	83.1±0.22	87.8±0.21	91.3±0.23	72.6±0.22	76.9±0.23	70.5±0.25	74.9±0.26	78.6±0.26	85.3±0.26	84.2±0.27	89.5±0.12	84.6±0.13	89.2±0.12	90.5±0.13	99.7±0.11
C10	81.2±0.13	88.7±0.12	82.1±0.13	90.2±0.13	93.0±0.12	99.6±0.09	81.1±0.16	91.4±0.11	83.8±0.12	91.1±0.13	88.6±0.13	90.4±0.11	73.6±0.22	79.2±0.21	82.3±0.11	91.3±0.13	90.9±0.13	89.8±0.14
C11	59.3±0.33	63.8±0.34	68.7±0.29	72.8±0.2	73.3±0.18	95.3±0.11	66.3±0.29	75.9±0.26	63.4±0.42	70.3±0.25	70.9±0.23	95.2±0.10	86.4±0.2	90.1±0.2	87.6±0.16	93.8±0.13	91.7±0.11	95.3±0.09
C12	67.9±0.28	72.7±0.27	76.4±0.25	80.6±0.19	80.9±0.20	98.0±0.11	69.8±0.24	72.9±0.21	70.1±0.20	73.0±0.19	77.2±0.19	98.1±0.09	87.6±0.11	96.5±0.10	88.3±0.13	95.7±0.11	98.0±0.11	
C13	60.1±0.31	63.5±0.30	69.9±0.28	72.0±0.25	76.4±0.24	79.2±0.21	65.6±0.31	73.6±0.28	64.8±0.33	69.6±0.31	79.1±0.32	70.5±0.2	82.1±0.15	88.9±0.14	85.6±0.15	91.6±0.11	85.5±0.14	93.9±0.10
C14	73.3±0.23	83.5±0.24	74.2±0.22	82.7±0.18	78.4±0.21	88.3±0.14	78.4±0.19	88.8±0.14	78.5±0.21	81.0±0.12	81.4±0.18	84.2±0.17	77.3±0.16	84.2±0.16	83.5±0.15	85.4±0.16	84.5±0.14	95.6±0.11
C15	70.3±0.25	77.2±0.24	77.3±0.24	82.6±0.18	80.7±0.18	88.7±0.16	73.5±0.22	81.6±0.18	70.5±0.23	72.4±0.21	74.7±0.21	84.2±0.17	83.9±0.16	91.7±0.12	86.5±0.13	87.8±0.15	93.4±0.12	95.8±0.13
C16	76.4±0.23	82.4±0.17	79.5±0.22	88.1±0.15	82.8±0.16	92.6±0.12	77.5±0.16	82.1±0.15	77.5±0.21	84.9±0.16	78.9±0.19	92.6±0.13	86.7±0.18	92.6±0.13	84.2±0.19	86.4±0.19	91.9±0.12	92.6±0.11
C17	66.8±0.33	75.5±0.25	73.6±0.22	80.6±0.18	82.1±0.17	94.4±0.13	70.6±0.24	80.2±0.24	72.6±0.29	77.8±0.28	79.6±0.25	94.1±0.16	75.6±0.16	82.2±0.14	83.4±0.15	92.1±0.13	88.3±0.14	94.4±0.11
C18	80.1±0.18	88.0±0.15	86.5±0.18	91.4±0.13	93.8±0.14	97.6±0.15	75.5±0.25	84.4±0.17	69.8±0.25	71.2±0.23	74.4±0.24	81.6±0.18	84.6±0.18	91.4±0.14	86.7±0.16	90.1±0.14	94.2±0.13	92.5±0.13
C19	72.1±0.33	79.6±0.21	81.4±0.22	84.8±0.23	87.7±0.23	91.7±0.17	71.5±0.29	79.0±0.28	50.2±0.42	54.9±0.44	58.1±0.42	57.8±0.32	85.7±0.42	95.7±0.12	82.3±0.14	88.6±0.15	88.8±0.12	98.1±0.10
C20	79.2±0.24	86.9±0.28	82.9±0.15	89.5±0.16	89.3±0.17	93.3±0.15	75.4±0.29	85.5±0.28	69.5±0.12	76.9±0.26	77.0±0.25	79.9±0.25	88.6±0.15	96.1±0.13	84.5±0.17	92.4±0.12	88.4±0.16	89.1±0.19
Total	70.7±0.3	77.8±0.2	75.6±0.2	80.8±0.2	81.95±0.2	88.9±0.2	71.3±0.2	79.5±0.2	70.5±0.2	75.2±0.2	76.4±0.2	84.8±0.2	81.3±0.2	88.1±0.2	83.8±0.2	89.2±0.2	90.4±0.2	93.55±0.2

Accuracy is presented for the following models: GMRF, TGMRF, and ABM for model orders 4, 6, and 10.

10 and the classes C04-Vegetation and a water channel, C13-Skyscrapers, and C14-Industrial complex had low recognition levels, when using model order 10.

The TGMRF model is a very similar model to the GMRF, which included the gray levels of the image and texture information within the model. The local gray levels of the patches changed for the SAR scenes, which contain strong scatterers or rivers, lakes, roads, or airports. The recognition rates of the TGMRF were lower when compared to the GMRF model for model orders 6 and 10. Classes below the average recognition rate were C01-Grassland, C04-Vegetation and a water channel, C11-Urban and Coastline, C13-Skyscrapers, and C19-Buoy in Sea, which were almost the same categories as in the case of GMRF.

The ABM model achieved the highest recognition rate, when compared to the GMRF and TGMRF models. It achieved 81%, 83%, and 90% of the recognition rate when 10% of the data were used for learning, and 88%, 89%, and 93% when 20% of the data were used at the learning stage, for model orders 4, 8, and 10, respectively. The classes below average recognition rate were C01-Grassland, C02-Forest, C06-Urban-02,

C07-Urban-03, C10-Natural water-channel, and C14-Industrial Complex. It could be noticed that ABM had lower deviation in the recognition rate than the GMRF and TGMRF models.

The nonparametric methods and parametric methods used an SAR patch with a dimension of 200 × 200 pixels. The size of the patch was chosen experimentally and was cut from different single look complex (SLC) SAR scenes, which have one look. The rotation invariant features for SAR data cannot be observed, since the features of the rotated data are changing, because the SAR data represent backscattering from targets. Signatures are also changing with the incidence angle of the data. The image patches were chosen using different data sets, a different incidence angle and different orientations of the data. We observed any efficiencies of the parametric and nonparametric methods, when the patch size was changing from 100 × 100 pixels to 400 × 400 pixels. Table III shows the mean accuracies and standard deviations for all 20 classes, when the patch size was changing. Table III shows that the best accuracy for the nonparametric wavelet method was achieved for a patch size of 300 × 300 pixels and for the parametric method the best accuracy was achieved for 200 × 200 pixels.

TABLE III

DEPENDENCY OF CLASS RECOGNITION ACCURACY VERSUS PATCH SIZE FOR PARAMETRIC AND NONPARAMETRIC METHODS USING 20% OF TEST IMAGES AT LEARNING STAGE

Class	100×100			200×200			300×300			400×400		
	ABM	ODTℂWT	MR8	ABM	ODTℂWT	MR8	ABM	ODTℂWT	MR8	ABM	ODTℂWT	MR8
C1	87.9 ± 0.19	94.1 ± 0.18	92.9 ± 0.14	84.4 ± 0.15	94.2 ± 0.11	93.8 ± 0.13	78.4 ± 0.19	**96.6 ± 0.18**	96.5 ± 0.16	71.2 ± 0.23	95.6 ± 0.25	94.7 ± 0.17
C2	91.3 ± 0.19	90.1 ± 0.19	89.9 ± 0.11	86.8 ± 0.15	90.8 ± 0.12	90.4 ± 0.11	79.1 ± 0.19	92.6 ± 0.19	90.9 ± 0.13	72.6 ± 0.23	**93.3 ± 0.26**	**91.3 ± 0.16**
C3	99.4 ± 0.17	88.4 ± 0.18	91.3 ± 0.16	96.4 ± 0.13	88.6 ± 0.11	92.4 ± 0.14	90.1 ± 0.17	91.2 ± 0.18	93.5 ± 0.17	84.2 ± 0.21	90.8 ± 0.25	94.1 ± 0.16
C4	95.6 ± 0.18	94.4 ± 0.2	92.0 ± 0.19	94.4 ± 0.14	96.1 ± 0.13	95.3 ± 0.16	89.9 ± 0.18	97.9 ± 0.21	96.4 ± 0.19	84.2 ± 0.22	**96.8 ± 0.27**	95.4 ± 0.18
C5	99.3 ± 0.16	96.4 ± 0.18	91.8 ± 0.18	96.6 ± 0.12	98 ± 0.11	93.9 ± 0.18	89.3 ± 0.16	98.3 ± 0.18	95.7 ± 0.17	83.1 ± 0.2	95.4 ± 0.25	94.6 ± 0.16
C6	**98 ± 0.15**	85.8 ± 0.19	84.5 ± 0.15	93.5 ± 0.11	86.4 ± 0.12	87.3 ± 0.13	85.5 ± 0.15	**88.8 ± 0.19**	90.9 ± 0.13	81.1 ± 0.19	87.1 ± 0.26	87.7 ± 0.14
C7	93.7 ± 0.16	94.6 ± 0.16	81.1 ± 0.14	90.3 ± 0.12	95.3 ± 0.09	83.4 ± 0.14	84.4 ± 0.16	97.8 ± 0.16	**83.4 ± 0.14**	78.7 ± 0.2	**98.5 ± 0.23**	82.3 ± 0.16
C8	**98.8 ± 0.16**	93.4 ± 0.19	84.9 ± 5 0.16	97.8 ± 0.12	93.8 ± 0.12	88.4 ± 0.17	91.3 ± 0.16	95.3 ± 0.19	89.9 ± 0.22	85.1 ± 0.2	96.1 ± 0.26	88 ± 0.22
C9	**99.4 ± 0.15**	93.1 ± 0.20	86.5 ± 0.14	99.7 ± 0.11	95.6 ± 0.13	87.9 ± 0.11	92.3 ± 0.15	**98.0 ± 0.2**	88.1 ± 0.13	84.7 ± 0.19	96.2 ± 0.27	**87.6 ± 0.15**
C10	93.8 ± 0.18	99.4 ± 0.19	93.1 ± 0.17	89.8 ± 0.14	100 ± 0.12	97.3 ± 0.18	83.4 ± 0.18	**99.8 ± 0.19**	**99.7 ± 0.19**	76.3 ± 0.22	99 ± 0.26	99.1 ± 0.2
C11	**99.1 ± 0.13**	95.1 ± 0.20	92.2 ± 0.14	95.3 ± 0.09	95.1 ± 0.13	92.8 ± 0.15	90.1 ± 0.13	**97.1 ± 0.2**	93.9 ± 0.15	84.9 ± 0.17	96.7 ± 0.27	90.7 ± 0.14
C12	95.2 ± 0.12	94.5 ± 0.18	89.5 ± 0.14	98.8 ± 0.08	95.5 ± 0.11	93.7 ± 0.13	94.2 ± 0.12	98.8 ± 0.18	**94.6 ± 0.14**	89.2 ± 0.16	**99.6 ± 0.25**	91.4 ± 0.14
C13	**97 ± 0.15**	93.2 ± 0.20	89.1 ± 0.21	93.9 ± 0.11	94.9 ± 0.13	93.4 ± 0.18	86.8 ± 0.15	**98.4 ± 0.2**	94.7 ± 0.19	78.9 ± 0.19	98.2 ± 0.27	91.8 ± 0.22
C14	**99.8 ± 0.15**	92.2 ± 0.16	91.1 ± 0.14	95.6 ± 0.11	93 ± 0.09	92.8 ± 0.16	90.4 ± 0.15	95.2 ± 0.16	94.3 ± 0.19	84.0 ± 0.19	95.3 ± 0.23	**94.7 ± 0.18**
C15	**99.2 ± 0.17**	88.1 ± 0.18	84.3 ± 0.19	95.8 ± 0.13	87.9 ± 0.11	85.2 ± 0.18	88.7 ± 0.17	89.6 ± 0.18	**85.7 ± 0.21**	81.3 ± 0.21	90.2 ± 0.25	83.3 ± 0.24
C16	95.4 ± 0.15	92.6 ± 0.18	91.3 ± 0.16	92.6 ± 0.11	94.2 ± 0.11	93.1 ± 0.16	86.3 ± 0.15	98.1 ± 0.18	**95.9 ± 0.15**	80.6 ± 0.19	**99.1 ± 0.25**	94.5 ± 0.14
C17	98.6 ± 0.15	93.0 ± 0.13	90.2 ± 0.12	94.4 ± 0.11	95.7 ± 0.06	94.1 ± 0.12	87.2 ± 0.15	**96.8 ± 0.13**	96.1 ± 0.11	79.9 ± 0.19	95 ± 0.2	**94.5 ± 0.11**
C18	97.3 ± 0.17	90.4 ± 0.18	89.2 ± 0.14	92.5 ± 0.13	90.8 ± 0.11	91.1 ± 0.09	85.8 ± 0.17	**93.5 ± 0.18**	**91.1 ± 0.12**	81.2 ± 0.21	91.6 ± 0.25	89 ± 0.11
C19	98.1 ± 0.14	92.1 ± 0.12	93.8 ± 0.13	98.1 ± 0.11	95.3 ± 0.05	95.6 ± 0.11	92.9 ± 0.14	**97.8 ± 0.12**	94.9 ± 0.1	87.1 ± 0.18	95.2 ± 0.19	**95.4 ± 0.1**
C20	**92.2 ± 0.23**	84.3 ± 0.18	91.3 ± 0.12	89.1 ± 0.18	85.5 ± 0.11	95.1 ± 0.1	84.3 ± 0.23	**86.9 ± 0.18**	94.9 ± 0.11	79.1 ± 0.27	85.5 ± 0.25	92.4 ± 0.11
Total	**96.5 ± 0.16**	92.3 ± 0.17	89.4 ± 0.14	93.7 ± 0.12	93.4 ± 0.10	91.8 ± 0.14	87.5 ± 0.16	95.4 ± 0.17	93.1 ± 0.15	81.34 ± 0.2	94.76 ± 0.24	91.6 ± 0.1595

TABLE IV

DEPENDENCY OF CLASS RECOGNITION ACCURACY VERSUS PARAMETERS OF THE SVM FOR PARAMETRIC AND NONPARAMETRIC METHODS USING 20% OF TEST IMAGES AT LEARNING STAGE

Kernel	ABM			ODTℂWT			MR8		
	SMO	QP	LS	SMO	QP	LS	SMO	QP	LS
Linear	94.1 ± 0.14	94.9 ± 0.14	89.4 ± 0.21	94.2 ± 0.15	95.8 ± 0.14	92.4 ± 0.18	92.3 ± 0.13	95.5 ± 0.15	88.6 ± 0.14
Quadartic	97.8 ± 0.15	97.8 ± 0.15	93.5 ± 0.19	96.8 ± 0.14	96.4 ± 0.14	93.4 ± 0.18	97.6 ± 0.13	97.3 ± 0.13	92.4 ± 0.14
Polynomial	96.5 ± 0.16	97.2 ± 0.15	91.7 ± 0.19	95.4 ± 0.17	96.1 ± 0.16	94.4 ± 0.18	93.1 ± 0.15	94.1 ± 0.13	89.5 ± 0.16

For ABM, ODTℂWT, MR8 methods and parameters of SVM were changed and accuracy was observed. Kernel parameters and methods for separating hyperplanes were observed.

The nonparametric methods used a subsampled dual tree wavelet transform; therefore, the scene content was separated better within the wavelet subbands, when patch size had 300×300 pixels. The gain was 2.5% on average, for nonparametric methods. By increasing the patch size, the parametric model was unable to estimate complex scenes; therefore, the ABM performed better with smaller patches, but the smaller patches insufficiently described the content of the SAR patch.

Table IV shows the dependence of the parametric ABM method, and the nonparametric ODTℂWT and MR8 on the parameters of the SVM classifier. In all experiments, polynomial kernel of order 3 and SMO were used. Table IV shows that better efficiency of presented methods can be achieved by using a quadratic kernel. The SMO hyperplane separation gave slightly better results when the same kernel was compared with the QP and least squares hyperplane separation.

C. Computational Complexity

The algorithms were executed using a GeForce GTX 480 graphical card, which has 16 graphical processing units (GPU) and Cuda platform, which enables parallel execution of algorithms. The nonparametric methods were much faster compared to the parametric information extraction methods. The execution times (s) of the parametric and nonparametric methods are summarized in Table V. The nonparametric method consisted of ODTℂWT, feature estimation within each subbands, and the learning phase where the SVM was used, as shown in Table V. The ODTℂWT, parameter estimation, and classification represented the recognition phase.

The learning stage of the parametric methods consisted of two phases, MAP estimation and evidence maximization using the amoeba algorithm, and the learning of texture parameters using SVM. The recognition phase consisted of MAP estima-

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

GLEICH *et al.*: PARAMETRIC AND NONPARAMETRIC METHODS FOR SAR PATCH SCENE CATEGORIZATION 11

TABLE V
EXECUTION TIMES OF PARAMETRIC AND NONPARAMETRIC METHODS (S)

Iter. no.	Nonparametric methods			GMRF			TGMRF			ABM		
	Operation	Time (s)		Operation	Time (s)		Operation	Time (s)		Operation	Time (s)	
		10%	20%		10%	20%		10%	20%		10%	20%
	Learning stage			Learning stage			Learning stage			Learning stage		
1.	ODTℂWT	10.5	21.1	MAP	915.2	1823.1	MAP	923.8	1878.4	MAP	997.2	1987.1
2.	Par. est.	265.1	597.6	Amoeba	1775.2	3576.1	Amoeba	1875.2	3634.1	Amoeba	2002.9	3935.6
3.	Learn	105.6	218.1	Learn	132.1	254.0	Learn	133.2	255.8	Learn	133.7	253.5
4.	Recognition			Recognition			Recognition			Recognition		
5.	ODTℂWT	94.5	84.8	MAP	8239.3	7331.3	MAP	8354.5	7386.0	MAP	8544.2	7573.4
6.	Par. est.	2120.4	1875.6	Amoeba	13972.2	12419.6	Amoeba	14030.2	12619.6	Amoeba	14381.3	12612.3
7.	Class.	10.4	10.3	Class.	10.4	10.8	Class.	10.9	10.2	Class.	10.5	10.8
	Total L&R	2606.5	2807.5		25044.4	25414.9		25227.8	25784.1		26069.8	26372.7

tion and evidence maximization using the rest of the data and classification of the textural parameters. The complexities of MAP-GMRF and MAP-TGMRF were very similar, meanwhile the MAP-ABM was more complex due to the solution of MAP estimates which were numerically solved using a Brent's algorithm [38].

Table V shows that the parametric methods were approx. 10 times more computationally demanding; therefore, the parametric methods are not very useful for the practical implementations.

V. CONCLUSION

This paper presented a comparison between the parametric and nonparametric approaches for SAR image categorization using ODTℂWT and model-based methods within an image domain. The nonparametric features within wavelet transformation better characterize features of the image patches in comparison with the parametric features, which used texture models. Among the presented GMRF, TGMRF, and ABM models, which were used within parametric-based methods, the ABM gave the highest accuracy of patch recognition and had a smaller deviation in patch categorization. From among them, the nonparametric features obtained within the ODTℂWT gave the best results. The presented parametric methods require MO selection, which is still an open issue and of disadvantage computationally regarding very demanding algorithm for second-order Bayesian inference, which represent texture parameter estimation.

REFERENCES

[1] I. G. Cumming and F. H.-C. Wong, *Digital Processing of Synthetic Aperture Radar Data: Algorithms and Implementation*. Norwood, MA, USA: Artech House, 1995.

[2] S. Wdowinski, S.-H. Hong, and S.-W. Kim, "Evaluation of TerraSAR-X observations for wetland InSAR application," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS'08)*, 2008, vol. 4, pp. IV-1233–IV-1236.

[3] M. Eineder, N. Nico, R. Bamler, N. Yague-Martinez, and H. Breit, "Spaceborne spotlight SAR interferometry with TerraSAR-X," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1524–1535, May 2009.

[4] M. Aubert *et al.*, "Toward an operational bare soil moisture mapping using TerraSAR-X data acquired over agricultural areas," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 900–916, Apr. 2013.

[5] A. Marino, "A notch filter for ship detection with polarimetric SAR data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 3, pp. 1219–1232, Jun. 2013.

[6] D. Perissin and T. Wang, "Time-series InSAR applications over urban areas in China," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 4, no. 1, pp. 92–100, Mar. 2011.

[7] T. Häme, Y. Rauste, O. Antropov, H. Ahola, and J. Kilpi, "Improved mapping of tropical forests with optical and SAR imagery, Part II: Above ground biomass estimation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 1, pp. 92–101, Feb. 2013.

[8] L. Pulvirenti, N. Pierdicca, M. Chini, and L. Guerriero, "Monitoring flood evolution in vegetated areas using COSMO-SkyMed data: The Tuscany 2009 case study," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.* vol. 6, no. 4, pp. 1807–1816, Aug. 2013.

[9] R. Dekker, "High-resolution radar damage assessment after the earthquake in Haiti on 12 January 2010," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 4, no. 4, pp. 960–970, Dec. 2011.

[10] K. Vadrevu *et al.*, "Hotspot analysis of vegetation fires and intensity in the Indian region," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 1, pp. 224–238, Feb. 2013.

[11] E. Valencia, A. Camps, N. Rodriguez-Alvarez, H. Park, and I. Ramos-Perez, "Using GNSS-R imaging of the ocean surface for oil slick detection," *IEEE J. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 1, pp. 217–223, Feb. 2013.

[12] C. Tison, J.-M. Nicolas, F. Tupin, and H. Matre, "New statistical model for Markovian classification of urban areas in high-resolution SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, pp. 2046–2057, Oct. 2004.

[13] M. Hebar, D. Gleich, and Z. Cucej, "Autobinomial model for SAR image despeckling and information extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2818–2835, Aug. 2009.

[14] M. Walessa and M. Datcu, "Model-based despeckling and information extraction from SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, pp. 2258–2269, Sep. 2000.

[15] G. Gao, X. Qin, and S. Zhou, "Modeling SAR images based on a generalized gamma distribution for texture component," *Prog. Electromagn. Res.*, vol. 137, pp. 669–685, 2013.

[16] K. Kayabol and J. Zerubia, "Unsupervised amplitude and texture classification of SAR images with multinomial latent model," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 561–572, Feb. 2013.

[17] C.-R. Shyu *et al.*, "Geoiris: Geospatial information retrieval and indexing system—Content mining, semantics modeling, and complex queries," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 4, pp. 839–852, Apr. 2007.

[18] T. Li and M. Ogihara, "Toward intelligent music information retrieval," *IEEE Trans. Multimedia*, vol. 8, no. 3, pp. 564–574, Jun. 2006.

[19] A. Popescu, I. Gavat, and M. Datcu, "Contextual descriptors for scene classes in very high resolution SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 1, pp. 80–84, Jan. 2012.

[20] J. Singh and M. Datcu, "SAR image categorization with log cumulants of the fractional Fourier transform coefficients," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 12, pp. 5273–5282, Dec. 2013.

[21] S. Cui and M. Datcu, "Statistical wavelet subband modeling for multitemporal SAR change detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 4, pp. 1095–1109, Aug. 2012.

[22] M. Molinier, J. Laaksonen, and T. Häme, "Detecting man-made structures and changes in satellite imagery with a content-based information retrieval system built on self-organizing maps," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 4, pp. 861–874, Apr. 2007.

[23] M. Molinier *et al.*, "Improving content-based target, and change detection in ALOS PALSAR images with efficient feature selection," in *Proc. ESA-EUSC2008 Symp. (WPP-278)*, Frascati, Italy, 2008.

[24] F. Bovolo and L. Bruzzone, "A split-based approach to unsupervised change detection in large-size multitemporal images: Application to tsunami-damage assessment," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, pp. 1658–1670, Jun. 2007.

[25] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines*. Cambridge, U.K.: Cambridge Univ. Press, 2000.

[26] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.

[27] M. Crosier and L. Griffin, "Using basic image features for texture classification," *Int. J. Comput. Vis.*, vol. 88, pp. 447–460, 2010.

[28] J.-M. Lina and L. Gagnon, "Image enhancement with symmetric daubechies wavelets," in *Proc. SPIE, Wavelet Appl. Signal Image Process. III*, 1995, vol. 2569, pp. 196–207.

[29] I. Selesnick, R. Baraniuk, and N. Kingsbury, "The dual-tree complex wavelet transform," *IEEE Signal Process. Mag.*, vol. 22, no. 6, pp. 123–151, Nov. 2005.

[30] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 1992.

[31] R. Coifman and M. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 713–718, Mar. 1992.

[32] B. Banjanin, B. Gergič, P. Planinšič, and Z. Čučej, "Entropy-threshold method for best basis selection," *Image Vis. Comput.*, vol. 19, no. 7, pp. 447–484, 2001.

[33] A. Kolmogorov, "New metric invariant of transitive dynamical systems and endomorphisms of lebesgue spaces," *Dokl. Russian Acad. Sci.*, vol. 119, pp. 861–864, May 1958.

[34] P. Planinšič, B. Gergič, D. Gleich, and Z. Čučej, "Fuzzy control of subband coded image quality using standard and fuzzy quality measure," *Opt. Eng.*, vol. 40, no. 8, pp. 1529–1544, 2001.

[35] D. Gleich, "Markov random field models for non-quadratic regularization of complex SAR images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 3, pp. 952–961, Jun. 2012.

[36] D. Bendoudjema and W. Pieczynski, "Unsupervised segmentation of non-stationary images using triplet Markov random fields," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 8, pp. 1367–1378, Aug. 2007.

[37] G. R. Cross and A. K. Janin, "Markov random field texture models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 5, no. 1, pp. 25–39, Jan. 1983.

[38] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recepies: The Art of Scientific Computing*, 3rd ed. Cambridge, U.K.: Cambridge Univ. Press, 2007, ch. 2.

[39] M. Hearst, S. Dumais, E. Osman, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intell. Syst. Appl.*, vol. 13, no. 4, pp. 18–28, Jul./Aug. 1998.

[40] A. Mishra. (2014). *Mathworks Homepage* [Online]. Available: http://www.mathworks.com/matlabcentral/fileexchange/33170-multi-class-support-vector-machine/

[41] F. Albgretzen. (2008). *Statistical Texture Measures Computed from Gray Level Co-Occurrence Matrices* [Online]. Available: http://www.uio.no/studier/emner/matnat/ifi/INF4300/h08/undervisningsmateriale/glcm.pdf

[42] P. Planinšič, J. Singh, and D. Gleich, "SAR image categorization using parametric and nonparametric approaches within a dual tree CWT," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 10, pp. 1757–1761, Oct. 2014.

Dušan Gleich (M'04) received the B.Sc, M.Sc, and Ph.D degrees from the University of Maribor, Maribor, Slovenia, in 1997, 2000, and 2002, respectively, all in electrical engineering.

He works as a Research Scientist with the Laboratory for Signal Processing and Remote Control, Faculty of Electrical Engineering and Computer Science, Maribor, Slovenia. He was a Visiting Scientist with the German Aerospace Center (DLR), Oberpffafenhofen, Germany, from 2004 to 2005, 2009, and 2012. From 2010, he was an Associate Professor with the University of Maribor and a Head of Laboratory for Signal Processing and Remote control. His research interests include image processing, data compression, and extraction information from SAR images.

Jagmal Singh photograph and biography not available at the time of publication.

Peter Planinšič photograph and biography not available at the time of publication.