Multilevel Local Pattern Histogram for SAR Image Classification

Dengxin Dai, Wen Yang, *Member, IEEE*, and Hong Sun, *Member, IEEE*

***Abstract*—In this letter, we propose a theoretically and computationally simple feature for synthetic aperture radar (SAR) image classification, the multilevel local pattern histogram (MLPH). The MLPH describes the size distributions of bright, dark, and homogenous patterns appearing in a moving window at various contrasts; these patterns are the elementary properties of SAR image texture. The MLPH is a very powerful descriptor of SAR images because it captures both local and global structural information. Additionally, it is robust to speckle noise. Experiments on a TerraSAR-X data set demonstrate that MLPH significantly outperforms four other widely used features in SAR image classification.**

***Index Terms*—Image classification, multilevel local pattern histogram (MLPH), synthetic aperture radar (SAR).**

I. Introduction

SYNTHETIC aperture radar (SAR) image classification has been actively researched in recent years for its broad applications in both civil and military domains [1]. The most challenge is the difficulty of extracting informative, distinguishing, and speckle-counteracting features. Although considerable efforts have been made, reliable and efficient features have remained out of reach. In this letter, we devise a suitable feature for high-resolution SAR image classification.

Statistical models such as the gamma, Weibull, and K distributions are widely used to describe scattering intensities in SAR images [2]–[6]. These mathematical models have done well in characterizing low-resolution SAR images. However, they cannot handle the abundant structural information in high-resolution SAR images, where classification requires the analysis of textures. Several texture descriptors have been adopted for SAR image classification [4], [7], [8]. The gray-level co-occurrence matrix (GLCM) [9] estimates the image properties from a distribution matrix of co-occurring gray values at a given offset in several defined directions. Gabor filters [10] capture textures by convolving the image data with Gabor wavelets tailored for a number of dilations and rotations. The Gaussian Markov random field (GMRF) [11] applies several Gaussian distributions with various orientations and predefined offsets to the pixels in a neighborhood. Although many breakthroughs have been made [4], [7], [8], [12], these tools were not originally designed for SAR image classification and cannot fully exploit the specific properties of SAR images.

To partially solve the aforementioned problems, we propose a theoretically and computationally simple feature for SAR image classification. The proposed feature is a concatenated histogram of local patterns extracted from the SAR image at multiple contrast levels, hereafter referred to as a multilevel local pattern histogram (LPH) (MLPH). Specifically, SAR images are first quantized into several channels at various particular contrasts by using a similar contrast technique in [13] and [14]. Then, the MLPH describes the size distributions of bright, dark, and homogenous patterns appearing in a moving window at these contrasts.

The remainder of this letter is organized as follows. The design objectives and principles behind the MLPH are described in Section II. Section III is devoted to comparative experimental results. We conclude this letter in Section IV.

II. MLPH

In this section, we first define the design objectives of the MLPH. Then, we will demonstrate the procedure of extracting an LPH at a specified contrast level. Finally, the MLPH is constructed.

A. Objectives

As we all know, the roughness (structure) of the Earth's surface is the dominant factor in SAR scattering. Fig. 1(a) shows a SAR image of a structure-dominated area (i.e., buildings). The local pattern is characterized by large coexisting bright and dark regions. For a texture-dominated area such as a woodland or a farmland, the local pattern is a scattering of brilliant "spots," often modulated by a macrotexture as shown in Fig. 1(b) and (c). For a very flat region such as water or a road surface, the local pattern is a homogeneous dark region [Fig. 1(d)]. The differences between these four patterns are mainly characterized by two factors: *size* (of the bright/dark regions, spots, etc.) and *level* (the contrast level of the local pattern). The MLPH is based on the idea that the *size* and *level* of local patterns are sufficient to identify the terrain category in SAR imagery.

Although many texture descriptors have been investigated for SAR image classification [4], [7], [8], a relatively few of these

Manuscript received September 30, 2009; revised June 22, 2010; accepted July 5, 2010. Date of publication September 7, 2010; date of current version February 25, 2011. This work was supported in part by the National Natural Science Foundation of China under Grants 40801183, 60890074, and 60872131 and in part by the National High-Technology Research and Development Program of China under Grants 2007AA12Z180 and 155.

The authors are with the School of Electronic Information and the State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China (e-mail: dxdai@mail.whu.edu.cn; yangwen@whu.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LGRS.2010.2058997

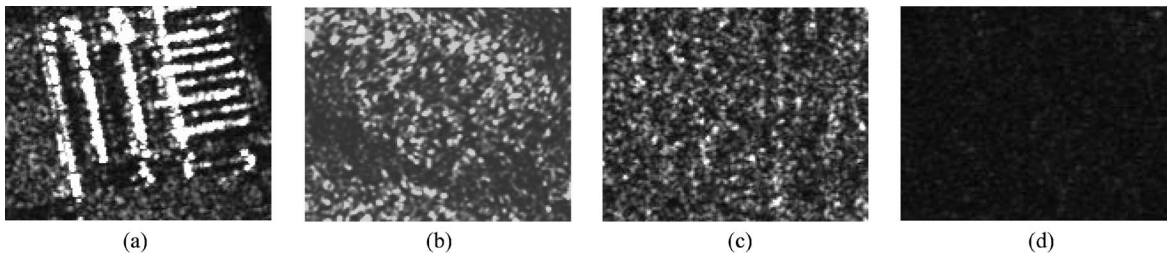

Fig. 1. Examples of images belonging to the three types of areas (see text for more details). (a) Building. (b) Woodland. (c) Farmland. (d) Water.

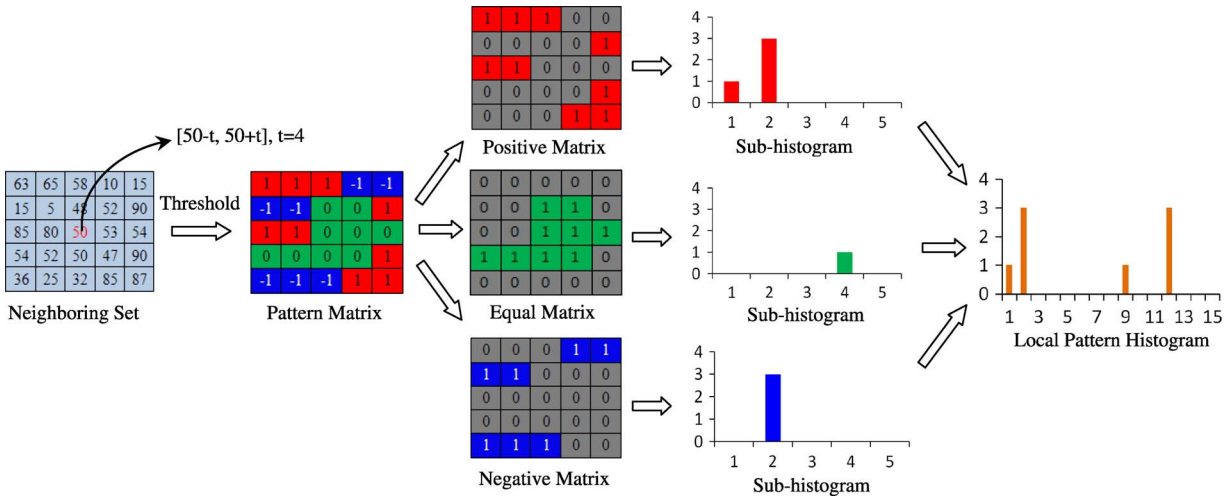

Fig. 2. Extracting an LPH from the neighborhood of a pixel. From left to right, the panels show the original intensity values, the quantization step, the matrix-splitting procedure, the subhistograms derived from the split matrices, and the LPH. The LPH is formed by concatenating the three subhistograms, and the subhistograms are computed based on (6) with $B = 2$.

have proven particularly suited for the task. To the best of our knowledge, this is the first paper to specifically investigate the discrimination power of the *size* and the *level* for SAR image classification.

B. LPH

In this section, we describe how to construct the LPH from an image. As an intuitive extension of LPH, the MLPH is proposed in the next section.

The LPH is based on the local behavior of pixel intensities; each pixel is compared to neighboring pixels within a moving window. There are three steps to the analysis: image quantization, matrix splitting, and histogram computation.

In the quantization stage, all pixel intensities within the window are compared to the intensity of the central pixel. Let g_c denote the intensity of the central pixel. Then, all intensities in the range $g_c \pm t$ are replaced with 0, intensities above this range are quantized to $+1$, and all others are replaced with -1. Mathematically, this quantization procedure can be formulated as an indicator function

$$s(i) = \begin{cases} +1, & g_i > g_c + t \\ 0, & |g_i - g_c| \le t; \\ -1, & g_i < g_c - t \end{cases} \quad i \in [1, \ldots, h^2] \tag{1}$$

where g_i is the intensity of pixel i in the local window and t is a predefined threshold. The moving window is a square of size h. We refer to the quantized image in a local window as the "pattern matrix." An example of the quantization ($h = 5$, $t = 4$) is shown in the left panel of Fig. 2.

In the second step, the pattern matrix is split into three matrices: a "positive matrix" (PM), an "equal matrix" (EM), and a "negative matrix" (NM). They are defined by the following indicator functions:

$$PM(i) = \begin{cases} 1, & s(i) > 0 \\ 0, & \text{others} \end{cases}, \quad i \in [1, \ldots, h^2] \tag{2}$$

$$EM(i) = \begin{cases} 1, & s(i) = 0 \\ 0, & \text{others} \end{cases}, \quad i \in [1, \ldots, h^2] \tag{3}$$

$$NM(i) = \begin{cases} 1, & s(i) < 0 \\ 0, & \text{others} \end{cases}, \quad i \in [1, \ldots, h^2] \tag{4}$$

where i refers to a specific element of any matrix and $s(i)$ is the ith element of the pattern matrix. The splitting procedure is shown in Fig. 2.

The three matrices play different roles in image representation. The positive matrix captures brilliant patterns, i.e., points or regions which are significantly brighter than the central pixel. The negative matrix describes dark patterns, and the equal matrix captures a homogenous region that includes the central pixel. The three matrices may seem redundant, but as the following histogram computation will show, this is not really true.

After splitting, a subhistogram is calculated for each matrix. The LPH is then formed by concatenating the three

subhistograms. In each matrix, elements with the value one are considered a foreground, while those with the value zero are considered a background. The local patterns counted in the LPH are continuous fragments of the foregrounds. Each fragment is assigned to a bin according to its number of pixels, thereby capturing the *size* of features in the foregrounds. For a given submatrix, let $bins(k)$ denote the value of the kth bin and N the number of local patterns in the matrix. The subhistogram can then be built as follows:

$$bins(k) = \sum_{n=1}^{N} \delta\left[num(n) = k\right] \tag{5}$$

where $num(n)$ denotes the number of pixels in the nth local pattern. The function $\delta[\cdot]$ is one if its argument is true and zero otherwise.

Theoretically, each subhistogram has h^2 bins. However, to obtain a more compact representation, we combine them into a much smaller set of K (a hand-tuned parameter) bins. Several strategies are possible, but it is best to require that higher order bins represent a wider range than lower order bins. There are two reasons for this approach.

1) The larger the $num(n)$, the rarer the corresponding pattern is.
2) The ability of humans to discriminate local patterns decreases as $num(n)$ increases. For instance, local patterns with 1 pixel and 3 pixels are obviously different, while patterns with 20 pixels and 22 pixels may be difficult to distinguish.

Here, one histogram-combining strategy is proposed

$$\begin{aligned} &vol(k) = B \times vol(k-1), \qquad k \in [2, \ldots, K] \\ &\text{s.t.} \sum_{l=1}^{K-1} vol(l) < h^2 \leq \sum_{l=1}^{K} vol(l) \end{aligned} \tag{6}$$

where $vol(k)$ denotes the "volume" [range in $num(n)$] of the kth bin in the simplified subhistogram and B is a parameter to control the growth rate of $vol(k)$. Obviously, with the increase of B from 1 to $+\infty$, the strategy controls the resource distribution for $vol(k)$ effectively. Therefore, it is a very general strategy to decrease the number of dimensions. Furthermore, it makes LPH invariant to scaling (normalized by the window size h), an expected property in SAR image classification. LPH is obtained by concatenating the three simplified subhistograms.

C. MLPH

Until now, we have only discussed the case of a single level (the fixed threshold t). However, as shown in Fig. 1, different terrain categories will exhibit textures in SAR imagery at different *levels* (contrast levels). It is therefore important to analyze the image with multiple thresholds in order to obtain an informative and robust descriptor. The multilevel quantization technique is superior to the single-level technique because it can capture the detailed variations of intensities around the central pixel and because it is resistant to speckle noise. The MLPH is formed by concatenating LPHs extracted at multiple *levels*.

Several strategies may be used to select appropriate values of t. However, it is more effective to put resources into smaller t because pixel intensities within a local window are usually similar to each other. Thus, the spacing between consecutive thresholds should increase with t. In the following, we suggest a strategy to account for the fact that uncertainty (local variability) increases with the intensity threshold.

In this study, one type of t-defining strategy is proposed

$$\begin{aligned} &t_m = T \times t_{m-1}, \qquad m \in [2, \ldots, M] \\ &\text{s.t. } t_M < C \leq t_{M+1} \end{aligned} \tag{7}$$

where M, the total number of levels, is again a hand-tuned parameter, T is a parameter to control the growth rate of t, and C, a constant, is the maximum contrast value in the image. This is also a general strategy to select appropriate values of t by tuning T in the range $[1, +\infty]$. Thus, the total dimension of the MLPH is $M \times 3 \times K$.

III. Experiments

A. Data Set and Experiment Setting

In this section, we analyze the performance of the MLPH in terrain classification on TerraSAR-X imagery. Our data set is based on two entire TerraSAR-X images of Foshan in Guangdong province, China, and of Wuhan in Hubei province, China, both of which are 48 189 × 25 255 pixels. The spatial resolution is about 3 m × 3 m. The true terrain classes for each region are labeled manually according to optical remote sensing imagery (SPOT) and related geographic information. Each pixel is assigned to one of the four semantic classes or *void*. The four classes are buildings, woodland, farmland, and water. Void pixels either do not belong to one of the four classes or lie near boundaries between classes, and were labeled as such to simplify the task of manual segmentation. About 11% of the pixels are unclassified ("void") in the experimental imagery. We ignore these pixels and their neighbors (within an $h \times h$ window) both in training and testing. We randomly select 20% pixels of the Wuhan image as training data and adopt the entire Foshan image as test data. All reported results represent averages over ten train–test partitions.

Two sets of experiments are performed in this study. The first set compares the performances of the MLPH under various parameter configurations. The second set compares the MLPH to three other widely used texture descriptors and to a gray histogram.

The support vector machine (SVM) methodology is adopted for classification. The implementation used in this research is LIBSVM[1] with a linear kernel [15].

B. Results on the TerraSAR-X Data Sets

In this section, we demonstrate the MLPH in TerraSAR-X image classification and analyze its performance for various choices of its input parameters. MLPH is calculated from a

[1]The software is available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

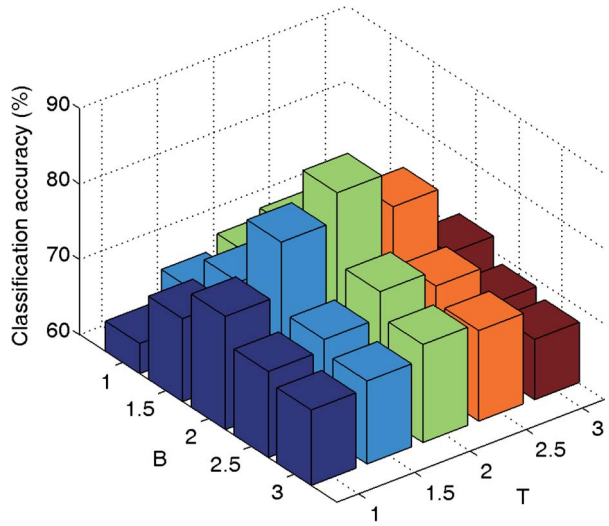

Fig. 3. Exemplar illustration ($h = 5$) of classification accuracy as a function of B and T. Other values of h obtain similar results.

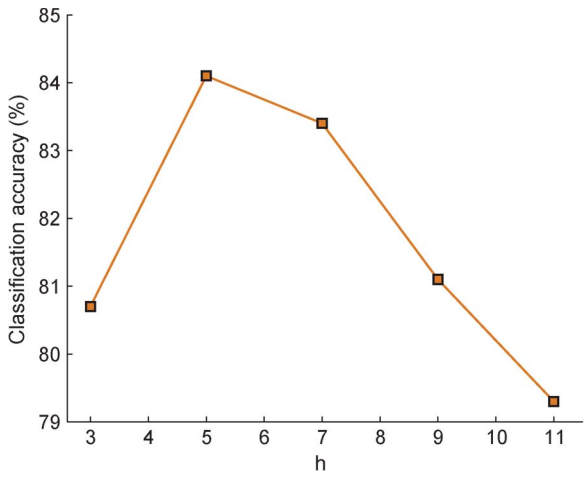

Fig. 4. Classification accuracy as a function of h.

TABLE I
CLASSIFICATION ACCURACIES ON THE TERRASAR-X DATA SET BY THE SVM WITH DIFFERENT FEATURES (IN PERCENT)

	Hist	GLCM	Gabor	GMRF	MLPH
Buildings	85.01	87.72	86.47	87.41	**88.08**
Farmland	68.11	79.71	64.62	65.56	**82.77**
Woodland	73.84	61.86	57.01	59.13	**76.52**
Water	90.69	88.74	79.09	87.15	**90.16**
Average	78.73	79.66	71.97	74.52	**84.08**

quantized 8-bit SAR image, namely, $C = 255$. M is set to five since a larger one cannot boost the final performance. K is set to five to strike a balance between discrimination power and computational burden. Therefore, we mainly test the effects of the other three parameters: h, B, and T.

Fig. 3 shows the classification accuracies with different K and T. Here, for ease of illustration, we fixed h to five. Other values of h achieve similar results in our experiments. In histogram combination, $B = 2$ achieves the best results. In level determination, $T = 2$ achieves the best results. What we can see from Fig. 3 is that too large values of B and T degrade the performance of MLPH, and so do too small values. It is worth emphasizing that the "optimal" B and T are almost certainly task dependent. Therefore, the various parameter configurations should be tested again when classifying other types of images. New binning and level strategies could also be designed for individual tasks.

Fig. 4 shows the classification accuracies with different h. From Fig. 4, we can conclude that h affects the performance of MLPH substantially. Therefore, a suitable selection of h is important for using MLPH. It is worth noticing that the optimal h is dependent on the image resolution and the classification task. In this work, with the resolution of 3 m $\times$ 3 m and the four defined categories (buildings, woodland, farmland, and water), $h = 5$ achieves the best performance in TerraSAR-X image classification.

C. Comparison With Other Methods

In this section, we compare the MLPH to four widely used features in SAR image classification: GLCM, Gabor filters, GMRF, and the gray histogram. Each feature is implemented with its optimal parameters, which are determined experimentally. GLCM is implemented in a moving window of 5 $\times$ 5 pixels on a quantized 4-bit SAR image and calculates four statistics: contrast, entropy, correlation, and homogeneity. Four directions (0°, 45°, 90°, and 135°) and two distances (one or two pixels) are used. MLPH, GMRF, Gabor filters, and the gray histogram are used on a quantized 8-b SAR image. Gabor filters are implemented on six scales and eight orientations, and their energies are used as classification features. GMRF is calculated in a moving window of 25 $\times$ 25 pixels. MLPH is used with the optimal settings obtained in the previous section.

Table I lists the classification accuracies obtained by SVM for the five descriptors. The accuracies are the percentage of image pixels belonging to each class that were assigned to the correct class. From Table I, we conclude that MLPH significantly outperforms the other four widely used descriptors in SAR image classification.

The superiority of MLPH can be attributed to four advantages. First, MLPH directly exploits the *size* and the *level* of the local patterns. Second, the MLPH is shape invariant, rotation invariant, and scale invariant. These properties are important because textures representing the same terrain type may vary with respect to the shapes and contrasts of bright local patterns and the direction of any large-scale trends. Third, MLPH is robust to speckle noise because it employs multiple thresholds. Fourth, in addition to capturing local information, the MLPH contains information on the global structure of the image. This is evident in the strong correlations obtained between MLPHs calculated from neighboring positions, for example, $MLPH(x, y)$ and $MLPH(x + 1, y)$. These two histograms must be correlated to each other since they both compare the intensities of pixel (x, y) and pixel $(x + 1, y)$. Many other direct and indirect relationships exist in the two windows, of course. Both types of constraints can propagate information regarding the local structure from one window to the next; therefore, the relationship between local windows can persist over large distances. The MLPH therefore implicitly captures

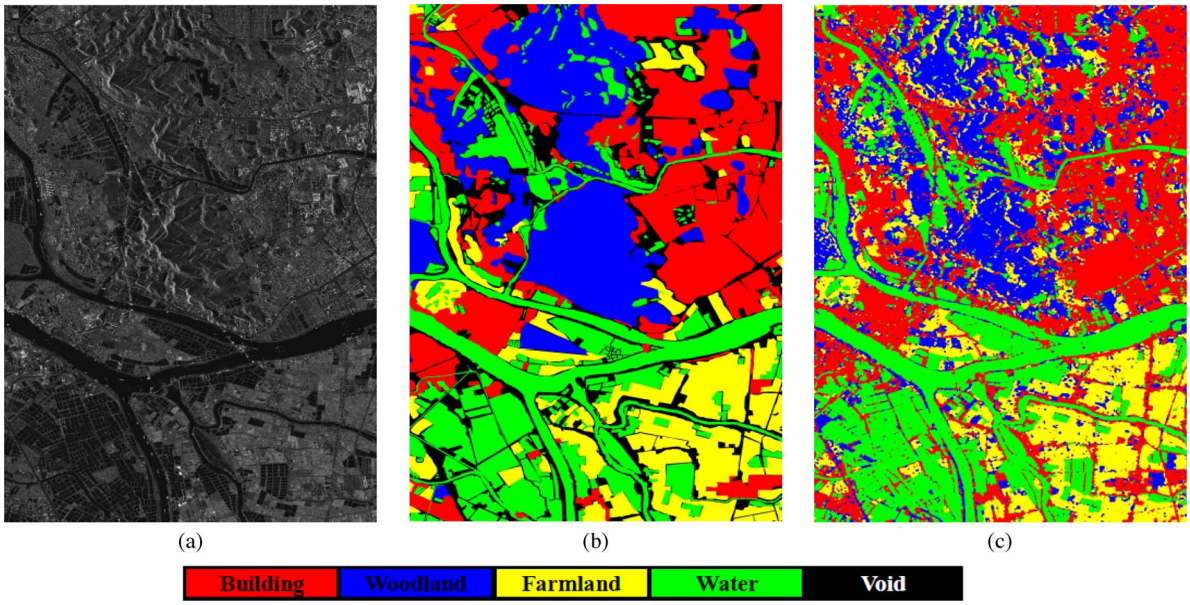

Fig. 5. TerraSAR-X image labeling results. (a) 88 006 400-pixel region of the original image. (b) and (c) show its ground truth and classification obtained by SVM using the MLPH, respectively.

large-scale image cues, which are particularly valuable for recognizing building areas and woodland.

Fig. 5 shows a portion of the TerraSAR-X image alongside the true terrain values and SVM classifications obtained with MLPH. It is worth emphasizing that, here, for ease of illustration, the training pixels in the portion are also classified and illustrated. From Fig. 5, we learn that satisfactory results can be obtained using only MLPH features. However, some isolated pixels are incorrectly classified even within homogenous regions, a consequence of processing each pixel independently. This deficiency can be alleviated by introducing some contextual information, such as a smoothing prior of Markov random fields.

IV. Conclusion

This letter has proposed a theoretically and computationally simple feature, referred to as the MLPH, for SAR image classification. The feature is based on the physical observations that the *size* and the *level* properties are effective for recognizing a terrain category. Our feature provides very promising performances on the TerraSAR-X data set and significantly outperforms other commonly used features. Future work will extend MLPH to polarimetric SAR image classification.

References

[1] C. J. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images*. New York: Artech House, 1998.

[2] P. Bernad, G. Denise, and R. Réfrégier, "Hierarchical feature-based classification approach for fast and user-interactive SAR image interpretation," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 1, pp. 117–121, Jan. 2009.

[3] Y. Dong, A. K. Milne, and B. C. Forster, "Segmentation and classification of vegetated areas using polarimetric SAR image data," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 2, pp. 321–329, Feb. 2001.

[4] C. Tison, J. M. Nicolas, F. Tupin, and H. Maître, "A new statistical model for Markovian classification of urban areas in high-resolution SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 10, pp. 2046–2057, Oct. 2004.

[5] H. Laur, "Analyse D'images Radar en Télédétection: Discriminateurs Radiométriques et Texturaux," Ph.D. dissertation, Paul Sabatier Univ., Toulouse, France, 1989.

[6] M. Petrou, F. Giorgini, and P. Smits, "Modelling the histograms of various classes in SAR images," *Pattern Recognit. Lett.*, vol. 23, no. 9, pp. 1103–1107, Jul. 2002.

[7] D. A. Clausi, "Comparison and fusion of co-occurrence, Gabor, and MRF texture features for classification of SAR sea ice imagery," *Atmos. Oceans*, vol. 39, no. 4, pp. 183–194, 2001.

[8] P. Maillard, D. A. Clausi, and H. W. Deng, "Operational map-guided classification of SAR sea ice imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 12, pp. 2940–2951, Dec. 2005.

[9] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-3, no. 6, pp. 610–621, Nov. 1973.

[10] A. C. Bovik, M. Clapk, and W. S. Geisler, "Multichannel texture analysis using localized spatial filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 1, pp. 55–73, Jan. 1990.

[11] R. Chellappa, "Two-dimensional discrete Gaussian Markov random field models for image processing," in *Progress in Pattern Recognition*, L. N. Kanal and A. Rosenfeld, Eds. New York: Elsevier, 1985, pp. 79–112.

[12] S. E. Franklin, *Remote Sensing for Sustainable Forest Management*. Boca Raton, FL: CRC Press, 2001.

[13] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," in *Proc. IEEE Int. Workshop Anal. Model. Faces Gestures*, Oct. 2007, pp. 168–182.

[14] T. Ojala, M. Pietikäinen, and T. Mäenpää, "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.

[15] C. C. Chang and C. J. Lin, LIBSVM: A Library for Support Vector Machines, 2001.