

Histogram-Based Attribute Profiles for Classification of Very High Resolution Remote Sensing Images

Begüm Demir, *Member, IEEE*, and Lorenzo Bruzzone, *Fellow, IEEE*

Abstract—Morphological attribute profiles (APs) obtained by the sequential application of morphological attribute filters to images have been found very effective in remote sensing (RS) to characterize spatial properties of objects in a scene. However, a direct use of the APs can be insufficient to provide a complete characterization of spatial information when complex texture is present in the considered images. To overcome this problem, in this paper, we present the novel histogram-based morphological APs (HAPs). The HAPs model the marginal local distribution of attribute filter responses to better characterize the texture information, and they are obtained based on a three-step algorithm. In the first step, the standard APs are constructed by sequentially applying attribute filters to the considered image. In the second step, a local histogram is calculated for each sample of each image in the APs. Then, in the final step, the local histograms of the same pixel locations in the APs are stacked, resulting in a texture descriptor whose components represent local distributions of the filter responses for the related pattern. Finally, the very-high-dimensional HAPs are classified by a support vector machine (SVM) classifier with histogram intersection kernel. Experimental results obtained by considering two very high resolution panchromatic images show the effectiveness of the proposed HAPs, which sharply improve the accuracy of the SVM classifier with respect to standard AP-based methods.

Index Terms—Histogram features, image classification, morphological attribute filters, remote sensing (RS), very high resolution (VHR) images.

I. INTRODUCTION

DUE to the availability of a large number of very high resolution (VHR) images, one important research area in remote sensing (RS) is the generation of an accurate land-cover map by classifying VHR images. The VHR images provide detailed information on the shape and geometry of the objects present on the ground, simplifying the problem of mixed pixels existing in medium (or low) resolution RS images. However, the improvement in spatial resolution increases the intraclass variability of each land-cover class, while reducing the inter-class variability between different classes [1]. These critical issues decrease the statistical separability of the different land-cover classes in the spectral domain, thus limiting the performance of traditional pixel-based classification. To overcome the

mentioned problems, the spatial information of the scene should be integrated in the classification process.

In the RS literature, several techniques that characterize the spatial information for the classification of VHR images have been developed [2]–[14], [24]–[26], [31]. Among them, the morphological attribute profiles (APs) have become very popular in recent years [14]. This is because the APs provide a detailed multilevel characterization of a VHR image by the sequential application of morphological attribute filters that model the structural information according to different attributes. However, when complex texture information is present in RS images, the use of APs can be insufficient to provide a complete characterization of the spatial information. To capture the complex texture appearance (according to the spatial distribution and spatial dependence of samples), in this paper, we propose histogram-based APs (HAPs) that allow efficient and accurate modeling of texture information from APs. The proposed HAPs are generated based on a three-step algorithm. The first step is devoted to constructing the standard APs, whereas the second step aims at defining the local histograms of samples in each image in the APs (which model the local distributions of both the filter responses and the original image bands). The final step is focused on fusing local histograms obtained at the same pixel location within the APs to define a texture descriptor for the corresponding pixel. Then, HAPs are characterized by the texture descriptors of the whole profile. Differently from the AP-based methods presented in the RS literature, the HAPs consist of histograms of attribute filter responses and thus are more efficient for integrating both spatial and spectral information with respect to standard APs. After the HAPs are constructed, they are given as input to a support vector machine (SVM) classifier with a histogram intersection (HI) kernel, which is a very effective kernel for high-dimensional histogram-based feature representations. Experiments carried out on two data sets, including VHR panchromatic images, demonstrate the effectiveness of the proposed method.

The main novelty of the proposed HAPs consists in the design and development of a histogram-based strategy to capture efficiently the complex texture appearance from APs to produce accurate land-cover maps. Moreover, we introduce the use of a HI kernel in the framework of the RS image classification problems as a similarity measure of samples in the kernel space. Note that, in recent years, the HI kernel has gained an increasing interest for image retrieval problems in the computer-vision communities [16], [17], whereas its use in RS has been recently explored in the context of content-based RS image retrieval problems in [18].

Manuscript received April 21, 2015; revised September 13, 2015; accepted October 9, 2015. Date of publication December 17, 2015; date of current version March 9, 2016.

The authors are with the Department of Information Engineering and Computer Science, University of Trento, I-38123 Trento, Italy (e-mail: demir@disi.unitn.it; lorenzo.bruzzone@ing.unitn.it).

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/TGRS.2015.2496167

The remaining part of this paper is organized as follows. Section II reviews the methods that model spatial information in the context of VHR RS image classification. Section III introduces the proposed HAPs. Section IV describes the considered data sets and the design of experiments, whereas Section V illustrates the experimental results. Finally, Section VI draws the conclusion of this paper.

II. LITERATURE SURVEY

Several methods that characterize spatial information for the analysis of VHR RS images have been developed in RS. The spatial information can be characterized by the texture and geometrical characteristics of the objects within images. The Gray level cooccurrence matrix (GLCM) is one of the most widely used methods to compute texture measures [24]. Several texture features can be computed from the GLCM, e.g., angular second moment, contrast, correlation, entropy, variance, inverse difference moment, difference average, difference variance, difference entropy, sum average, sum variance, and sum entropy. Each feature models different properties of the statistical relation of pixels cooccurrence estimated within a given moving window and along predefined directions and interpixel distances [21]. Wavelet transform has been studied for texture analysis in VHR RS images in [25]. The texture features obtained from the GLCM is compared with those obtained by the scale-invariant feature transform (which is a translation, rotation, and scale-invariant texture feature extraction technique [20]) and by a bank of Gabor filters in [31]. A pointwise approach embedded into a graph model to achieve a graph-based local texture representation for VHR image classification has been recently presented in [7]. Markov random fields (MRFs) have proven effective in [2] to model the texture information associated with VHR images. The MRFs, Gibbs random fields, and local linear transformations have been compared in [30] to characterize texture features. The use of rotation-invariant descriptors based on local binary patterns and local phase quantization for texture characterization in the context of VHR RS images has been introduced in [26]. Segmentation-based methods have been widely used for driving the extraction of texture features in homogeneous regions in VHR images [4]–[6]. As an example, in [6], texture features have been extracted by exploiting an adaptive, multilevel, and hierarchical segmentation method.

Recently, the use of morphological filters to model the geometrical and texture properties of objects in a scene has attracted great attention in RS. As an example, morphological profiles (MPs) in the framework of the classification of VHR RS images are described in [8] and [9] for the characterization of spatial information. The MPs are built by applying opening and closing operations to the image features and are found effective for representing the multiscale variability of the structures in the image. Nevertheless, they are not appropriate to model all geometrical features. This limitation is overcome by the use of the morphological APs, which provide a detailed multilevel characterization of a VHR image. It is achieved by sequentially applying morphological attribute filters that model the spatial characteristics of the structures in a scene according to different

attributes [10]. The attribute filters cannot be directly applied to multispectral and hyperspectral images due to the lack of total order between vector values. To overcome this issue, different strategies have been presented in the literature. For example, in [11], the concept of APs is successfully extended to hyperspectral image classification problems by using the principle component analysis to reduce the dimensionality of the data. APs are initially generated only on the first few principle components and then stacked into a single profile [named extended APs (EAPs)]. Another strategy that relies on considering the attribute-constrained connectivity concept (where attributes are computed on the connected components of a hierarchy of partitions relying on a dissimilarity measure computed for every pair of adjacent pixels) is presented in [27] and [28]. APs can have a high dimensionality if a wide-range attribute filters and related parameter values are considered. This may result in the curse of dimensionality problem and thus in a poor generalization ability for classification systems. To address this issue, recent studies have been focused on developing feature selection algorithms applied to APs. To this end, in [12], a genetic-algorithm-based feature selection method is introduced, whereas sparse representations of the APs (which aim at exploiting the inherent low-dimensional structure of the APs) are presented in [13]. Sparse representations assume that the AP feature vectors lie in class-dependent low-dimensional subspaces (or submanifolds) even if they may have high dimensionality [13]. To approximately solve sparse problems in the context of sparse classification of APs, the orthogonal matching pursuit and variable splitting and augmented Lagrangian techniques have been used in [13]. To the best of our knowledge, current research on AP-based RS image classification focuses on reducing the dimensionality of APs and automatically selecting the attribute filtering parameters [14], whereas there is no work existing on further analysis of the texture information from APs to capture complete spatial characteristics of the structures in the scene.

III. PROPOSED METHOD

Let $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_P]$ be an RS image made up of P spectral channels, where $\mathbf{X}_p = \{x_1, x_2, \dots, x_B\}$ is the p th channel made up of B samples, and x_j is the j th pixel. The standard AP-based techniques proposed in the RS literature exploit the stacked filtered images (obtained by applying a sequence of attribute filters) together with the original image features. Depending on the parameters of the considered attribute filters, it is possible to model different structures in a scene with the AP-based methods. In order to increase the discrimination capability of features of APs in difficult classification tasks, here, we introduce HAPs. The goal of the HAPs is to integrate the spatial and spectral information by extracting APs and considering the texture information present on each single image in the APs to characterize accurately the texture elements and their distributions. To this end, the HAPs are constructed based on a three-steps procedure: 1) generation of the APs; 2) estimation of a local histogram for each sample within each image in APs (which represents local distributions of the filter responses); and 3) concatenation of local histograms obtained at the same pixel location within APs (which defines a texture descriptor for the

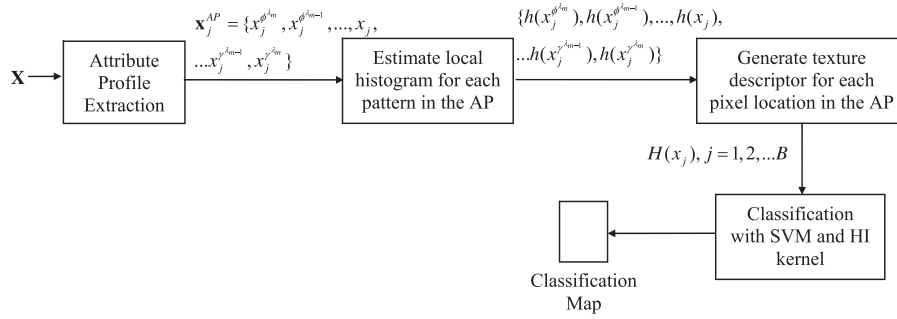


Fig. 1. Block diagram of the proposed HAP approach.

corresponding pixel in APs). Fig. 1 shows the block scheme of the proposed HAPs generation. Each step of the proposed method is explained in detail in the following.

A. Generation of Morphological APs

The first step aims at extracting the AP from an image \mathbf{X} . This is achieved by the sequential application of attribute filtering operation that is based on the evaluation of how a given attribute is computed for each connected component of an image band for a given threshold λ [14]. In other words, attribute filters are connected filters, and they operate on an image by only merging its connected components. A connected component is considered a group of isolevel pixels that are connected associated to a connectivity rule. The common connectivity rules are four and eight connected, for which a pixel is considered adjacent to four or eight of its neighbor pixels, respectively [14]. The connected components are the foreground and background regions existing in the image for binary images. Considering this information, the set of connected components for grayscale images can be achieved by representing the image as a stack of binary images generated by thresholding the image at all its gray-level values [14]. This is known as the threshold decomposition principle. For a connected component of the image, if the attribute value is higher than the threshold λ , the region is retained unchanged, whereas the attribute is merged with a surrounding connected component if its value is lower than the threshold λ [14]. The operation of merging the region to the adjacent region of a lower gray level is called attribute thinning, whereas it is called attribute thickening if the region is fused to the adjacent region of a greater gray level.

Without losing any generality, let us consider that the image \mathbf{X} contains only a single spectral channel, i.e., $P = 1$; thus, $\mathbf{X} = \mathbf{X}_1$. The filtered images $\mathbf{X}^{\phi^\lambda}$ and $\mathbf{X}^{\gamma^\lambda}$ are obtained by applying the attribute thickening ϕ^λ and attribute thinning γ^λ operations, respectively, to \mathbf{X} with the filter threshold λ as follows (see [10] for details):

$$\mathbf{X}^{\phi^\lambda} = \min \{s : x \in \Phi^\lambda [\text{Th}_s(\mathbf{X})]\} \quad (1)$$

$$\mathbf{X}^{\gamma^\lambda} = \max \{s : x \in \Gamma^\lambda [\text{Th}_s(\mathbf{X})]\} \quad (2)$$

where Φ^λ and Γ^λ are the binary attribute thickening and thinning operations, respectively, which preserve or remove the connected region based on the filter threshold λ . $x \in \mathbf{X}$

and $\text{Th}_s(\mathbf{X})$ is the binary image obtained by the threshold decomposition principle by thresholding \mathbf{X} at gray level s (that ranges on the gray levels of \mathbf{X}). According to (1) and (2), the binary attribute thickening and thinning operations are initially applied to each binary image, respectively. Then, the grayscale attribute thickening and thinning are achieved by the minimum and maximum gray level of the filtering results for each pixel $x \in \mathbf{X}$, respectively [10].

For a given attribute and an ordered sequence of m thresholds $\{\lambda_1, \lambda_2, \dots, \lambda_m\}$, the AP of \mathbf{X} (denoted $\text{AP}(\mathbf{X})$) is given by

$$\text{AP}(\mathbf{X}) = \{\mathbf{X}^{\phi^{\lambda_m}}, \dots, \mathbf{X}^{\phi^{\lambda_1}}, \mathbf{X}, \mathbf{X}^{\gamma^{\lambda_1}}, \dots, \mathbf{X}^{\gamma^{\lambda_m}}\} \quad (3)$$

where $\mathbf{X}^{\phi^{\lambda_i}} = \{x_1^{\phi^{\lambda_i}}, x_2^{\phi^{\lambda_i}}, \dots, x_B^{\phi^{\lambda_i}}\}$ and $\mathbf{X}^{\gamma^{\lambda_i}} = \{x_1^{\gamma^{\lambda_i}}, x_2^{\gamma^{\lambda_i}}, \dots, x_B^{\gamma^{\lambda_i}}\}$ are the filtered images obtained after applying the attribute thickening and thinning, respectively, by considering the i th threshold λ_i . The j th pattern $x_j \in \mathbf{X}$, $j = 1, 2, \dots, B$, can be represented by a feature vector \mathbf{x}_j^{AP} in the AP as $\mathbf{x}_j^{\text{AP}} = \{x_j^{\phi^{\lambda_m}}, x_j^{\phi^{\lambda_{m-1}}}, \dots, x_j^{\phi^{\lambda_1}}, x_j, x_j^{\gamma^{\lambda_1}}, \dots, x_j^{\gamma^{\lambda_m}}\}$, $j = 1, 2, \dots, B$. In other words, \mathbf{x}_j^{AP} is constructed by the stacked filtered responses associated to the related sample x_j . Accordingly, the AP of \mathbf{X} can be rewritten in terms of \mathbf{x}_j^{AP} , $j = 1, 2, \dots, B$, as follows:

$$\text{AP}(\mathbf{X}) = \{\mathbf{x}_1^{\text{AP}}, \mathbf{x}_2^{\text{AP}}, \dots, \mathbf{x}_B^{\text{AP}}\}. \quad (4)$$

It is worth noting that the set of connected components can be represented by a tree image representation, and such representation of an image can be efficiently used for the computation of morphological attribute filtering [14]. As an example, thinning and thickening operations can be obtained from a max-tree and min-tree image representations, respectively. For an illustrative example, we refer the reader to [14].

The APs can extract spatial information of the considered image and capture various types of structural and texture information depending on the selected attributes and filtering parameters. Different attributes can be computed for characterizing different properties. Commonly used attributes are: 1) area (the size of the regions); 2) moment of inertia (which models the elongation of the objects in the scene); and 3) standard deviation (which models region homogeneity). Characterizing an image with different types of attributes results in an efficient description of the spatial structures. It is worth noting that if more than one attribute is considered, the multiple APs are defined by stacking APs obtained for each single attribute by

using (1) and (2). In the case of considering more than one image band, i.e., $P > 1$ (e.g., with multispectral/hyperspectral images) one can use the EAPs that can be defined by: 1) applying a feature extraction method to the data to reduce the feature dimension; and then 2) constructing the APs on the reduced features and stacking them. For further information on the APs, the reader is referred to [10]–[14].

B. Estimation of Local Histograms

This step aims at extracting the texture information for each sample that describes the local spatial information in each image in the APs to provide complete spatial characteristics of the structures in the scene. To this end, we analyze the distributions of neighborhood pixels of each pattern $\{x_j^{\phi^{\lambda_m}}, x_j^{\phi^{\lambda_{m-1}}}, \dots, x_j, \dots, x_j^{\gamma^{\lambda_{m-1}}}, x_j^{\gamma^{\lambda_m}}\}_{j=1}^B$ in each image in the APs. Due to the effectiveness of histograms to represent textures efficiently, here, we estimate local histograms $\{h(x_j^{\phi^{\lambda_m}}), h(x_j^{\phi^{\lambda_{m-1}}}), \dots, h(x_j), \dots, h(x_j^{\gamma^{\lambda_{m-1}}}), h(x_j^{\gamma^{\lambda_m}})\}_{j=1}^B$ of each sample $\{x_j^{\phi^{\lambda_m}}, x_j^{\phi^{\lambda_{m-1}}}, \dots, x_j, \dots, x_j^{\gamma^{\lambda_{m-1}}}, x_j^{\gamma^{\lambda_m}}\}_{j=1}^B$, where $h(x_j^{\phi^{\lambda_m}})$ shows the local histogram of $x_j^{\phi^{\lambda_m}}$.

In order to estimate the local histogram $h(x_j^{\phi^{\lambda_m}})$ of a sample $x_j^{\phi^{\lambda_m}}$, initially, the range of possible values are defined by the minimum and maximum pixel values of the related image $\mathbf{X}^{\phi^{\lambda_m}}$ (independently from the other images), and the range is divided into b histogram intervals (i.e., b histogram bins). Then, the histogram $h(x_j^{\phi^{\lambda_m}})$ is computed by counting how many of the neighborhood patterns within a $w \times w$ -sized spatial window centered at the corresponding pixel $x_j^{\phi^{\lambda_m}}$ location belongs to each interval. Accordingly, the local histogram describes a feature vector consisting of a marginal local distribution of a filter response for each sample and thus represents efficiently the texture of the related sample. Note that if a sufficient number b of histogram bins is defined, the local histograms can represent the underlying local sample distribution with high precision. Thus, the local histogram associated with each sample in each image in the APs is capable of effectively capturing the texture content of the related sample.

C. Generation of the Texture Descriptors via Concatenating Local Histograms

This step aims at integrating all the local histograms $\{h(x_j^{\phi^{\lambda_m}}), h(x_j^{\phi^{\lambda_{m-1}}}), \dots, h(x_j), \dots, h(x_j^{\gamma^{\lambda_{m-1}}}), h(x_j^{\gamma^{\lambda_m}})\}_{j=1}^B$ of the j th pixel locations in the APs to define a texture descriptor $H(x_j)$ for each pixel $x_j, j = 1, 2, \dots, B$. This is achieved by stacking all local histograms of the j th pixel locations in APs as follows:

$$H(x_j) = \left\{ h(x_j^{\phi^{\lambda_m}}), h(x_j^{\phi^{\lambda_{m-1}}}), \dots, h(x_j), \dots, h(x_j^{\gamma^{\lambda_{m-1}}}), h(x_j^{\gamma^{\lambda_m}}) \right\}. \quad (5)$$

In this way, the texture descriptor $H(x_j)$ integrates responses of attribute filters to form a texture feature for the sample x_j and defines a region appearance based on spectral and spatial

properties. In other words, the local histograms are concatenated into one histogram in $H(x_j)$ that captures the texture information present in all images in the APs. The $H(x_j)$ can be also rewritten as

$$H(x_j) = \left\{ \zeta_j^1, \zeta_j^2, \dots, \zeta_j^{|H(x_j)|} \right\} \quad (6)$$

where ζ_j^k represents the histogram feature value (i.e., frequency) given at the k th component, and the dimension $|H(x_j)|$ of $H(x_j)$ is defined as $|H(x_j)| = P \times b \times (2m + 1)$. It is worth noting that, in (6), the first b components are the elements of the local histogram $h(x_j^{\phi^{\lambda_m}})$ of $x_j^{\phi^{\lambda_m}}$, whereas the following b components are the elements of the local histogram $h(x_j^{\phi^{\lambda_{m-1}}})$ of $x_j^{\phi^{\lambda_{m-1}}}$ and so on. In order to better explain this procedure, Fig. 2 illustrates a qualitative example based on a single image band (i.e., $P = 1$). Note that, for simplicity, the example shows an AP obtained by including one attribute with only two different threshold values, i.e., $m = 2$.

After estimating the texture descriptors $H(x_j), j = 1, 2, \dots, B$ for each pixel, we construct the HAP using all the texture descriptors obtained for all the patterns in a scene. Accordingly, the HAP provides an efficient description of texture information and produces promising texture discriminative features. After the HAP of an image is constructed, it is given as input to a classifier to generate the corresponding land-cover map. It is worth noting that the HAPs can be very high dimensional when: 1) a wide range of filters (i.e., a dense sampling of the values of the parameters of the filters) is selected to generate APs; 2) a high histogram bin number is chosen; and 3) the image consists of several spectral channels. Depending on the considered classifier [1], this issue may lead to the Hughes phenomenon, i.e., the classification accuracy decreases by increasing the number of features given as input to the classifier over a given threshold. To overcome this problem, we propose to apply the SVM classifier to the HAPs by considering the HI kernel that has been extensively used for histogram comparisons in the computer vision community [16], [17].

To measure the similarities between the texture descriptors $H(x_j)$ and $H(x_k)$ for samples x_j and x_k , respectively, the HI kernel is defined as

$$K(H(x_j), H(x_k)) = \sum_{l=1}^{P \times b \times (2m+1)} \min(\zeta_j^l, \zeta_k^l) \quad (7)$$

where $\zeta_j^l \in H(x_j), j = 1, 2, \dots, B$ and $\zeta_k^l \in H(x_k), j = 1, 2, \dots, B$ denote histogram features related to l th histogram bin. Note that the HI kernel is a positive definite parameter-free kernel for histogram-based features. SVMs using the HI kernel can generalize well on challenging classification problems (where the feature space dimension may vary from thousands to tens of thousands [15]) through efficiently modeling the sparsity of data. It is worth recalling that the HAP consists of histogram-based features that model marginal sample distributions and thus represents the discrete densities that are sparse (i.e., several histogram bins having zero entries). Accordingly, even in the case of high-dimensional HAPs, due to the high generalization capability of SVMs with the HI kernel, the

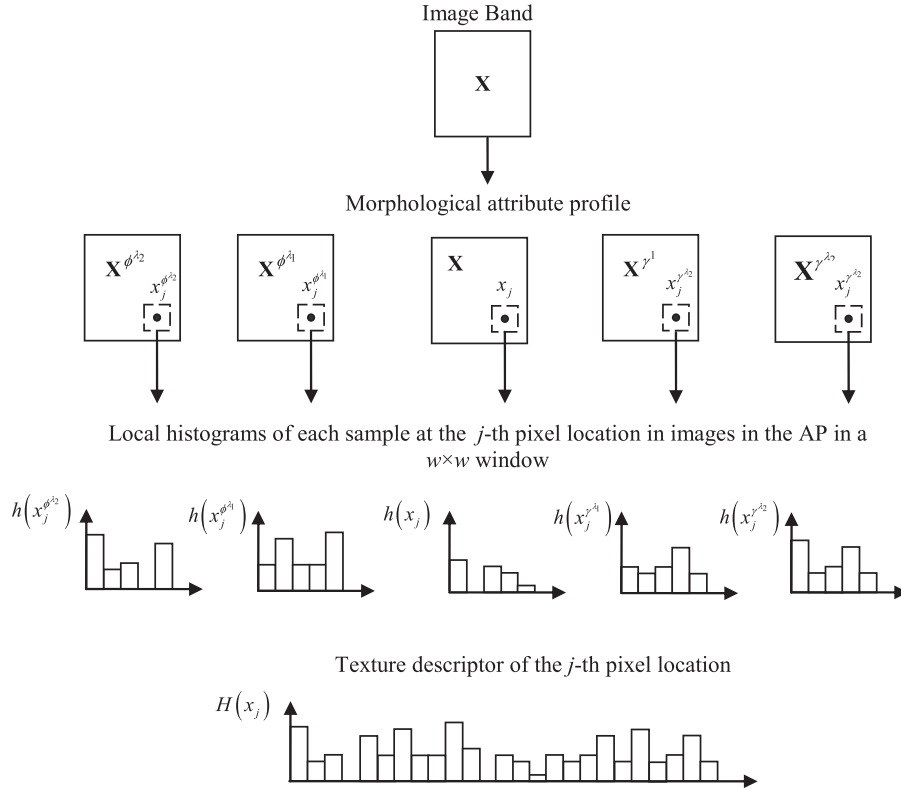


Fig. 2. Qualitative example that illustrates how to generate a texture descriptor for the j th pixel location in the AP.

performance of the proposed HAP is not degraded by the Hughes phenomenon.

The overall computational complexity of the HAPs is associated to the computational time required for: 1) the generation of the APs; and ii) the computation of the texture descriptors from the APs. As stated in [10], the computational cost to generate an AP (based on the max/min-trees) is associated to $O(BG + 4mB)$, where B is the number of pixels in the image, G is the number of gray levels in the image, and m is the number of filter thresholds. However, the computation can be significantly speeded up based on a concurrent computation of attribute filters on shared memory parallel machines [29]. The computational complexity of the texture descriptor generation is much smaller than that required for the generation AP and is $O(B^2m(w^2 + b))$. To reduce the computational cost, methods that can estimate the local histograms in $O(B^2mb)$ [22], [23] can be used.

IV. DATA-SET DESCRIPTION AND DESIGN OF EXPERIMENTS

Experimental analyses are conducted on a panchromatic image of 10000×10000 pixels that was acquired by the Quickbird scanner on the city of Trento, Italy, on July 2006 with geometric resolution of 0.6 m. In order to show the effectiveness of the proposed method, two test sites for which the reference samples are available are selected from the whole panchromatic image. The first test site (denoted as Trento-1 data set) is an image made up by 400×400 pixels [see Fig. 3(a)], whereas the second test site (denoted as Trento-2 data set) consists of 900×900 pixels [see Fig. 3(b)]. Both images represent two

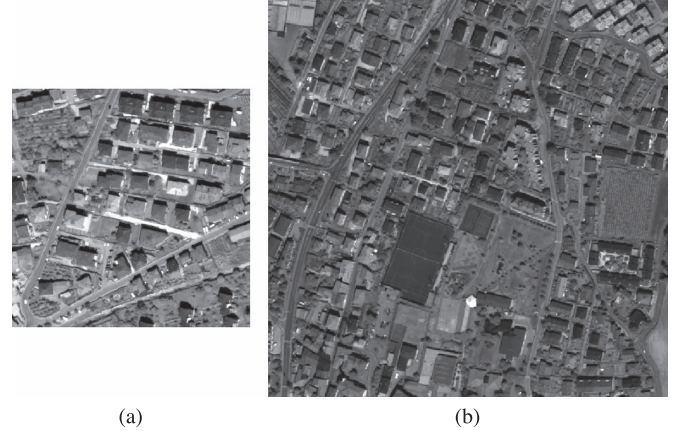


Fig. 3. Portions of a panchromatic image used for quantitative accuracy assessments. (a) Trento-1 data set. (b) Trento-2 data set.

complex urban areas due to the presence of: 1) different size and shape residential buildings and large industrial buildings; and 2) shadows mainly present close to the buildings. For both images, the available ground reference samples are representative of the four land-cover classes (i.e., road, building, shadow, and vegetation). The available labeled data for the Trento-1 data set (7005 samples) were divided to derive a training set of 885 samples and a test set of 6120 samples, whereas those for the Trento-2 data set (14735 samples) were divided to derive a training set of 1200 samples and a test set of 13 535 samples. Table I shows the land-cover classes and the related number of samples used in the experiments for both data sets.

In the experiments, three different attributes were considered for constructing the AP: 1) the area; 2) the standard deviation;

TABLE I
NUMBER OF SAMPLES OF EACH CLASS IN THE TRAINING SET
AND THE TEST SET FOR BOTH CONSIDERED DATA SETS

Land-Cover Class	Data Set			
	Trento-1		Trento-2	
	Training	Test	Training	Test
Roads	199	907	300	3068
Buildings	209	3330	300	4184
Shadow	255	853	300	1715
Vegetation	222	1030	300	4568
Total	885	6120	1200	13535

and 3) the moment of inertia. As suggested in [10] for the same data sets, the AP with the area attribute is obtained with λ values of 49, 169, 361, 625, 961, 1369, 1849, and 2401, whereas that with the standard deviation is obtained with λ values of 10, 20, 30, 40, 50, 60, 70, and 80. The AP with the moment of inertia attribute is constructed by selecting values of λ as 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, and 0.9 as in [10]. In the experiments, we have also considered APs constructed by the use of multiple attributes such as: 1) the moment of inertia and the standard deviation; and 2) the moment of inertia, the standard deviation, and the area. In the experiments, to construct APs, a four-connected connectivity rule is considered. After the HAPs are constructed, they are given as input to the SVM classifier (denoted as HAP-SVM). The proposed HAP-SVM is achieved by using the HI kernel that is parameter free. In the model selection, the regularization parameter of SVM is tested between [10–2000] with a step size increment of 20, the number b of histogram bins is tested in the range of [5–15] with a unit step size increment, and the local window size w is tested within the range of [5–11] with a step size of 2. Then, the best values are obtained by using a fivefold cross-validation approach for both data sets.

We compare the results of the proposed HAP-SVM with those obtained by: 1) the SVM classification applied to original image features (denoted as SVM); 2) the SVM classification applied to the standard APs (denoted as AP-SVM) [10]; and 3) the SVM classification applied to the textural features (i.e., mean, variance, homogeneity, contrast, and entropy) extracted from the GLCM computed on a window size of 9×9 pixels (denoted as GLCM-SVM). The SVM, the AP-SVM, and the GLCM-SVM are accomplished by using the radial basis function (RBF) kernel that is defined as [19]

$$K(H(x_j), H(x_k)) = \exp\left(-\gamma \|H(x_j), H(x_k)\|^2\right), \quad \gamma > 0. \quad (8)$$

This choice is done because the RBF is much effective than the HI kernel if the features are not histograms (note that the RBF kernel is used also in the related paper of the AP [10]). In the experiments, the value of the γ parameter of the RBF kernel (which is tested between 0.1 and 4 with a step size increment of 0.1) and that of the regularization parameter of SVM (which is tested between 10 and 2000 with a step size of 20) were also obtained using a 5-fold cross-validation approach for both data sets.

All the experiments were implemented via MATLAB on a standard personal computer with 2.93-GHz Intel Core i7 CPU and 8-GB RAM.

TABLE II
OVERALL ACCURACY OBTAINED WITH HAP-SVM BY VARYING THE
NUMBER OF BINS AND THE WINDOW SIZE WHEN STANDARD DEVIATION
ATTRIBUTE IS USED. THE BEST RESULTS ARE GIVEN IN BOLD
(TRENTO-1 DATA SET)

Bin number (b)	Window size (w)			
	5×5	7×7	9×9	11×11
5	70.2	80.5	77.1	72.61
6	69.4	73.6	71.7	69.6
7	66.7	66.8	70.8	71.2
8	73.2	72.7	72.7	72.7
9	71.6	70.9	72.6	73.9

TABLE III
OVERALL ACCURACY OBTAINED WITH HAP-SVM BY VARYING THE
NUMBER OF BINS AND THE WINDOW SIZE WHEN STANDARD DEVIATION
AND MOMENT OF INERTIA ATTRIBUTES ARE JOINTLY USED. THE BEST
RESULTS ARE GIVEN IN BOLD (TRENTO-1 DATA SET)

Bin number (b)	Window size (w)			
	5×5	7×7	9×9	11×11
5	88.9	88.6	88.2	88.3
6	89.8	90.1	89.4	88.9
7	90.1	89.4	90.0	89.6
8	91.4	92.4	92.4	91.7
9	91.3	92.6	91.8	90.7

V. EXPERIMENTAL RESULTS

A. Results on the Trento-1 Data Set

Here, we present the results of the proposed HAP for the classification of the Trento-1 data set. The parameters required to estimate the HAPs related to each attribute, i.e., the number b of histogram bin and the window sizes w are important to have a good approximation of the underlying distribution of samples for defining accurate local histograms. In the experiments, as aforementioned, the best parameters are selected according to the 5-fold cross-validation procedure. In particular, from our experiments, we have observed that, under proper definition of the range of parameters, the accuracies of the HAP-SVM are always higher than those obtained by the AP-SVM. For example, Tables II and III show the overall accuracies obtained by varying the number b of bins and window sizes w when the HAPs are obtained with: 1) the standard deviation; and 2) the joint use of inertia and standard deviation attributes, respectively. Note that the AP-SVM provides an accuracy of 56.4% when the standard deviation attribute is used, whereas it yields an accuracy of 74.2% when the moment of inertia and standard deviation attributes are jointly used. From tables, one can see that the accuracies obtained by the HAP-SVM are in general higher than those obtained with the standard AP-SVM for both cases at any value of the corresponding parameters. In greater details, the highest accuracies of the HAP are obtained with: 1) $(b, w) = (5, 7)$ in the case of standard deviation attribute; and 2) $(b, w) = (9, 7)$ in the case of joint use of the moment of inertia and the standard deviation attributes. It is worth noting that the same behavior is also observed in the results obtained for the other attributes (not reported due to space constraints). Specifically, the highest accuracies of the HAP are achieved with: 1) $(b, w) = (5, 11)$ in the case of area attribute; 2) $(b, w) = (10, 9)$ in the case of the moment

TABLE IV
OVERALL CLASSIFICATION ACCURACY OBTAINED BY THE STANDARD SVM, THE GLCM-SVM, THE AP-SVM
AND THE PROPOSED HAP-SVM. THE BEST RESULTS ARE GIVEN IN BOLD (TRENTO-1 DATA SET)

Features/Attributes	SVM	GLCM-SVM	AP-SVM	HAP-SVM
Only Panchromatic	55.5		-	-
GLCM Texture Features		61.8	-	-
Area	-		60.3	82.3
Moment of Inertia	-		78.4	92.0
Standard Deviation	-		56.4	80.5
Moment of Inertia+Standard Deviation	-		74.2	92.6
Moment of Inertia+ Area	-		83.7	89.2
Standard Deviation+ Area	-		64.6	83.8
Area+Moment of Inertia+Standard Deviation	-		86.1	90.7

TABLE V
CLASS BY CLASS SVM ACCURACY OBTAINED WHEN i) THE ORIGINAL FEATURES, ii) THE MOMENT OF INERTIA AND iii) THE JOINT USE OF MOMENT OF INERTIA AND STANDARD DEVIATION ATTRIBUTES ARE USED TO BUILD THE APs. THE BEST RESULTS ARE GIVEN IN BOLD (TRENTO-1 DATA SET)

Land-cover classes	Only Panchromatic	Moment of Inertia		Moment of Inertia+ Standard Deviation	
	SVM	AP-SVM	HAP-SVM	AP-SVM	HAP-SVM
Roads	81.4	82.0	92.9	85.8	97.8
Buildings	35.4	72.6	89.4	61.1	88.0
Shadow	97.2	95.2	97.8	98.8	99.5
Vegetation	63.4	80.2	94.9	85.5	97.4
Overall Accuracy	55.5	78.4	92.0	74.2	92.6

of inertia attribute; 3) $(b, w) = (8, 9)$ in the case of jointly using all the attributes; 4) $(b, w) = (6, 11)$ in the case of jointly using the area and the moment of inertia attributes; and 5) $(b, w) = (9, 11)$ in the case of jointly using the area and the standard deviation attributes.

Table IV shows the classification results obtained by the standard SVM, the GLCM-SVM, the AP-SVM, and the proposed HAP-SVM associated with the different attributes. In Table IV, we reported the highest accuracies of the HAP-SVM obtained with the best values of the parameters b and w identified through the 5-fold cross-validation model selection. From Table IV, one can observe that, as expected, the accuracies obtained by the standard SVM are significantly improved by using the GLCM texture features and APs in the context of both the AP-SVM and the HAP-SVM. In addition, the accuracies achieved by the proposed HAP-SVM are sharply higher than those yielded by the GLCM-SVM and the AP-SVM independently from the selected attributes. As an example, the proposed HAP-SVM obtains an accuracy of 92% when the moment of inertia attribute is used, whereas the AP-SVM provided an accuracy of 78.4% with the same attribute and the GLCM-SVM resulted in an accuracy of 61.8%. By using the standard deviation attribute, the accuracy of HAP-SVM is 25% higher than that of standard SVM and 24% higher than that of AP-SVM. Note that, since the GLCM-SVM proved less accurate than the AP-SVM at most of the cases, in the following, the detailed analysis is focused on the comparison of the HAP-SVM with the AP-SVM for this data set. Accordingly, Table V shows the class-by-class accuracies obtained by the standard SVM, AP-SVM, and HAP-SVM when we used: 1) the moment of inertia; and 2) joint use of the moment of

inertia and standard deviation attributes. From Table V, one can see that, for all the classes, the proposed HAP-SVM provides the highest accuracies for both cases. For example, using the moment of inertia attribute, the proposed HAP-SVM achieves an accuracy on the “buildings class” of 89.4%, whereas the standard SVM and the AP-SVM obtain accuracies of 35.4% and 72.6%, respectively. For a qualitative analysis, Figs. 4 and 5 show the related classification maps. From the figures, one can observe that the thematic maps obtained by the SVM and the AP-SVM are less accurate than those obtained by the proposed HAP-SVM. This shows that the proposed HAP-SVM allows the available spatial/texture information to be incorporated into the classification process much more accurately than the AP-SVM. It is worth noting that the HAP-SVM (due to the use of local histograms) results in smoother classification maps than the AP-SVM. The level of smoothing depends on the window size w . Overall, experimental results obtained on the Trento-1 data set show that the HAP-SVM can considerably improve the classification accuracy of VHR images compared with direct SVM and AP-SVM.

B. Results on the Trento-2 Data Set

Here, we present the results of the proposed HAP for the classification of the Trento-2 data set. As in the case of the first data set, the values of the parameters required to construct the HAPs associated with each attribute (i.e., the number b of histogram bin and window size w) are chosen by the fivefold cross-validation procedure. Moreover, for the Trento-2 data set, from the experiments, we have observed that under a reasonable selection of the range of the HAP parameters, the accuracies

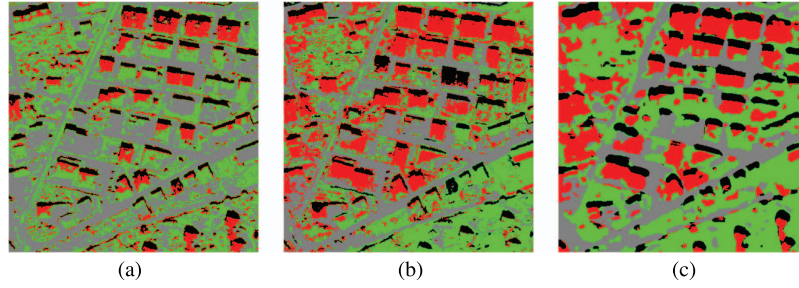


Fig. 4. Classification maps of the Trento-1 data set obtained by (a) the SVM when only the panchromatic image is used, (b) by the AP-SVM, and (c) by HAP-SVM when the moment of inertia attribute is used (gray: roads; red: buildings; black: shadow; green: vegetation).

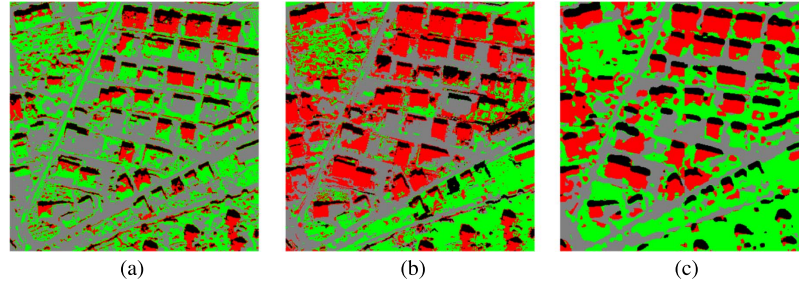


Fig. 5. Classification maps of Trento-1 data set obtained by (a) the SVM when only the panchromatic image is used, (b) by the AP-SVM, and (c) by HAP-SVM when the moment of inertia and standard deviation attributes are jointly used (gray: roads; red: buildings; black: shadow; green: vegetation).

TABLE VI
OVERALL ACCURACY OBTAINED WITH THE HAP-SVM BY VARYING THE NUMBER OF BINS AND THE WINDOW SIZE WHEN THE MOMENT OF INERTIA ATTRIBUTE IS USED. THE BEST RESULT IS GIVEN IN BOLD. (TRENTO-2 DATA SET)

Bin number (b)	Window size (w)			
	5×5	7×7	9×9	11×11
4	68.8	69.0	68.3	67.4
5	66.3	67.4	68.1	66.2
6	66.5	68.7	71.3	68.9
7	64.4	66.2	69.1	68.0
8	68.8	66.2	69.3	69.7

TABLE VII
OVERALL ACCURACY OBTAINED WITH HAP-SVM BY VARYING THE NUMBER OF BINS AND THE WINDOW SIZE WHEN THE STANDARD DEVIATION ATTRIBUTE IS USED. THE BEST RESULT IS GIVEN IN BOLD. (TRENTO-2 DATA SET)

Bin number (b)	Window size (w)			
	5×5	7×7	9×9	11×11
4	62.6	62.5	63.3	63.0
5	61.2	63.8	64.9	65.4
6	68.2	71.8	71.7	70.3
7	64.8	65.9	69.3	67.9
8	65.0	66.2	67.5	68.5

of the HAP-SVM are always higher than those yielded by the AP-SVM. For example, Tables VI and VII show the overall accuracies of the HAP-SVM with respect to different values of b and w when the moment of inertia and the standard deviation attributes are used, respectively. Note that the AP-SVM provides an accuracy of 62.4% when the moment of inertia attribute is used, whereas it yields an accuracy of 47.1% when the standard deviation attribute is considered. From the table, one can observe that the HAP-SVM, with any combination of b and w , performs better than AP-SVM. Specifically, the highest accuracies of the HAP are obtained with: 1) $(b, w) = (6, 9)$ in the case of the moment of inertia attribute; and 2) $(b, w) = (6, 7)$ in the case of standard deviation attribute. It is worth noting that the same behavior is also observed in the results for the other attributes (not reported for space constraints). The highest HAP accuracies are obtained with: 1) $(b, w) = (7, 9)$ in the case of area attribute; 2) $(b, w) = (6, 11)$ in the case of joint use of the moment of inertia and standard deviation attributes; 3) $(b, w) = (6, 7)$ in the case of joint use of all the attributes, 4) $(b, w) = (6, 5)$ in the case of jointly using the area and the moment of inertia attributes; and 5) $(b, w) = (7, 9)$ in the case of jointly using the area and the standard deviation attributes.

Table VIII shows the classification accuracies obtained by the SVM, the GLCM-SVM, the AP-SVM, and the proposed HAP-SVM for the Trento-2 data set. In Table VIII, we reported the highest average accuracy of the HAP-SVM obtained with the best values of the parameters b and w . By analyzing the figure, one can observe that the proposed HAP-SVM, again, leads to the highest accuracies for all the attributes and significantly outperforms the standard SVM, GLCM-SVM, and AP-SVM methods. As an example, the HAP-SVM provides an accuracy of 71.8% when the standard deviation attribute is used, whereas the accuracy obtained by the AP-SVM with the same attribute is 48.1% and that obtained by the GLCM-SVM is 48.9%. In addition, the HAP-SVM achieves an accuracy of 77.8% when all the attributes are used, whereas the AP-SVM could never reach such accuracy with any of the attribute. The AP-SVM provided higher accuracies than the GLCM-SVM in most of the cases. Accordingly, the results of the GLCM-SVM are no more reported for this data set.

Table IX shows the class-by-class accuracies obtained by the standard SVM, the AP-SVM, and the HAP-SVM for the moment of inertia and the standard deviation attributes. The related classification maps are given in Figs. 6 and 7, respectively. As

TABLE VIII
OVERALL CLASSIFICATION ACCURACY OBTAINED BY THE STANDARD SVM, THE GLCM-SVM, THE AP-SVM,
AND THE PROPOSED HAP-SVM. THE BEST RESULTS ARE GIVEN IN BOLD (TRENTO-2 DATA SET)

Features/Attributes	SVM	GLCM-SVM	AP-SVM	HAP-SVM
Only Panchromatic	41.1	-	-	-
GLCM Texture Features		48.9	-	-
Area	-	-	48.2	60.5
Moment of Inertia	-	-	62.4	71.3
Standard Deviation	-	-	47.1	71.8
Moment of Inertia+Standard Deviation	-	-	64.3	77.7
Moment of Inertia+ Area	-	-	48.3	71.1
Standard Deviation+ Area	-	-	49.8	65.6
Area+Moment of Inertia+Standard Deviation	-	-	64.8	77.8

TABLE IX
CLASS-BY-CLASS SVM ACCURACY OBTAINED WHEN: 1) THE ORIGINAL FEATURES, 2) THE INERTIA, AND 3) THE STANDARD DEVIATION ATTRIBUTES
ARE USED TO BUILD THE APs, THE BEST RESULTS ARE GIVEN IN BOLD (TRENTO-2 DATA SET)

Land-cover classes	Only Panchromatic	Moment of Inertia		Standard Deviation	
	SVM	AP-SVM	HAP-SVM	AP-SVM	HAP-SVM
Roads	67.4	69.0	71.1	31.2	46.3
Buildings	43.5	47.3	61.9	24.7	73.3
Shadow	82.2	81.9	88.9	92.0	95.9
Vegetation	29.0	70.4	73.6	61.3	78.4
Overall Accuracy	41.1	62.4	71.3	47.1	71.8

one can see from the tables and figures, the class accuracies obtained by the proposed HAP-SVM are significantly higher than those yielded by the SVM and the AP-SVM for all the land-cover classes. As an example, when the standard deviation attribute is used, the “vegetation class” accuracy achieved using HAP-SVM is 78.4%, whereas those obtained using the standard SVM and the AP-SVM are 29% and 61.30%, respectively. It is worth noting that the results obtained on the Trento-2 data set are lower than those achieved on the Trento-1 data set since the Trento-2 data set represents more complex urban areas.

From all the results, we can observe that, when the HAPs are considered, the classification accuracy is significantly improved with respect to the standard APs. This demonstrates the importance of information conveyed by marginal local distributions for optimizing the classification performance and the effectiveness of SVM with the HI kernel in exploiting the high-dimensional HAPs.

C. Analysis of the Computational Complexity

Here, we evaluate the computational time required: 1) to construct the APs and the HAPs; and 2) to classify them. In order to show the computational complexities for a large image, in this part of the experiments, we considered the full image having size of 10000×10000 pixels rather than the considered two smaller test sites. Due to the unavailability of the ground reference samples that are representative of all land-cover classes present in the entire image, here, we only analyze the computational complexity without the classification accuracies (note that the accuracies obtained on the available reference samples are already provided earlier for the two test sites). Tables X and XI show the computational time taken by the construction of the AP and the HAP and by their

classification with SVMs. The results given in Table X are obtained: 1) when all the attributes considered in this paper (i.e., the area, the moment of inertia, and the standard deviation) are used to build the APs (thus, the feature length in the AP is 49); and 2) when the HAP is constructed by selecting $(b, w) = (8, 9)$ (thus, the feature length in the HAP is 392). The results reported in Table XI are achieved: 1) when the moment of inertia and standard deviation attributes are used to generate the APs (thus the feature length in the AP is 33); and 2) when the HAP is constructed by selecting $(b, w) = (9, 7)$ (thus, the feature length in the HAP is 297). By analyzing the tables, one can see that the time taken by the HAPs is slightly higher than that required by the APs due to the estimation of the texture descriptors. It is worth noting that the amount of computational time differences between the APs and the HAPs depends on the number of pixels in the image and the total number of filtered images in the APs (see Section III). Thus, it can be significantly reduced by using a cluster-based parallel computing approach to estimate the samples’ texture descriptors. Moreover, from the tables, one can also see that the computational time taken by the HAPs classification by SVMs is higher than that of the APs due to the increased number of features. Nevertheless, also in this case, one can use parallel processing in order to significantly decrease computational time required for the HAP classification.

VI. DISCUSSION AND CONCLUSION

In this paper, we have proposed HAPs, which provide a more complete and detailed characterization of the spatial information than the standard APs for the classification of VHR RS images. Unlike the standard APs, the proposed HAPs model the marginal local distribution of attribute filter responses by

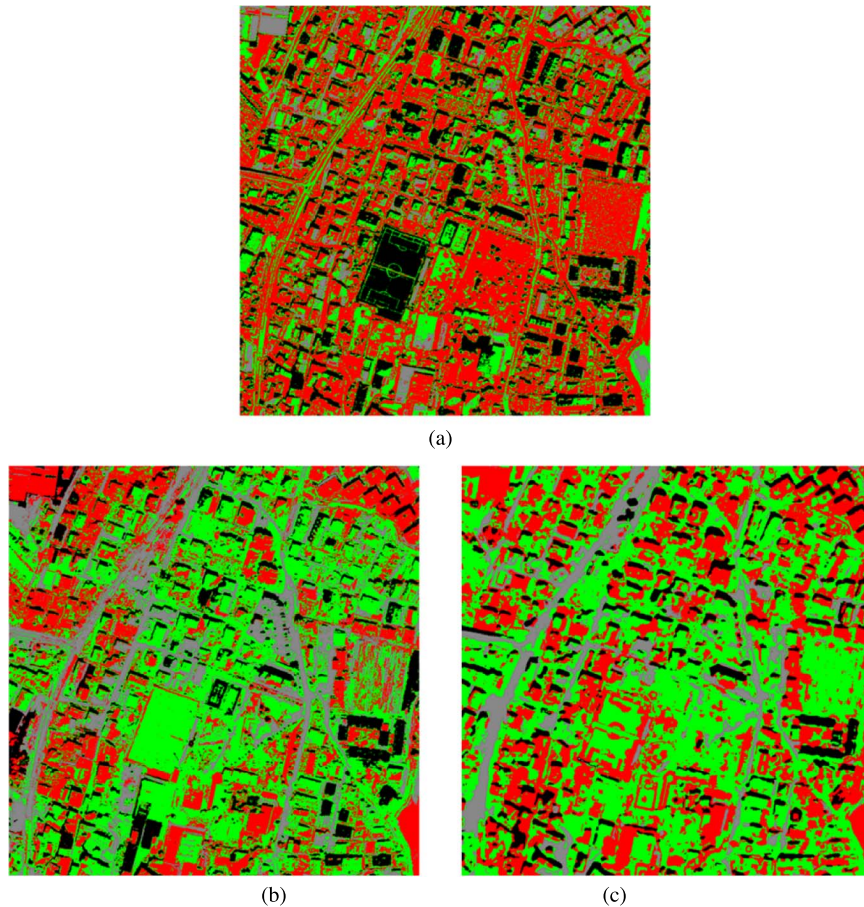


Fig. 6. Classification maps of the Trento-2 data set obtained by (a) the SVM when only the panchromatic image is used, (b) by the AP-SVM, and by (c) HAP-SVM when the inertia attribute is used (gray: roads; red: buildings; black: shadow; green: vegetation).

histogram-based texture descriptors extracted by a three-step algorithm. In the first step, the standard APs (which are sets of stacked attribute filtered images together with the original image) are obtained. In the second step, the local histograms of each pixel (which model the texture information of each sample in each image) in the APs are calculated. In the final step, the local histograms are concatenated to form a final texture descriptor for each pattern in the AP. It is worth noting that the texture descriptors of the HAPs are: 1) easy to compute; 2) sparse (i.e., several histogram bins have zero entries); and 3) highly discriminative when complex texture appearance is present. To classify the HAPs, we considered an SVM classifier with an HI kernel that has a high generalization capability even in the case of high-dimensional sparse histograms. The main novelties of the proposed HAPs are related to the proposed strategy to model the marginal local distribution of attribute filter responses through histogram-based texture descriptors. Moreover, we introduce the use of the HI kernel for VHR image classification problems (particularly in the context of the SVM classification) in RS.

Experimental results obtained on two VHR panchromatic images show that the proposed HAPs are very promising for characterizing the spatial/texture information. In the experimental analysis, we compared the proposed method with a standard AP-based method presented in the RS literature. By this comparison, we observed that the proposed HAP allows

one to significantly outperform the standard APs. The proposed method improves the classification accuracy without being affected by the Hughes phenomenon, due to the high generalization ability of SVMs with the HI kernel. We emphasize that the robustness to feature dimensionality is a very important advantage since, in RS, the image features can be tremendously high-dimensional. It is worth noting that the proposed HAP is intrinsically classifier independent. However, in this paper, we propose to classify it with SVMs by using the HI kernel due to the aforementioned advantages. One possible drawback of the proposed approach is that, based on the local window histograms, it produces classification maps slightly smoother than the standard approach. However, as in the computation of other texture measures, the smoothing effect can be controlled by reducing the window size. Moreover, the slight smoothing is compensated by both the sharp increase in accuracy and the overall improvement in the quality of the classification maps.

As a final remark, we would like to point out that the computational time required for both the generation of the proposed HAPs and their classification is slightly higher than those required by the APs. However, the order of magnitude does not change; thus, the operational implications are not significantly different from those of the APs. Note that the computational time can be reduced by adopting a parallel processing algorithm for the implementation.

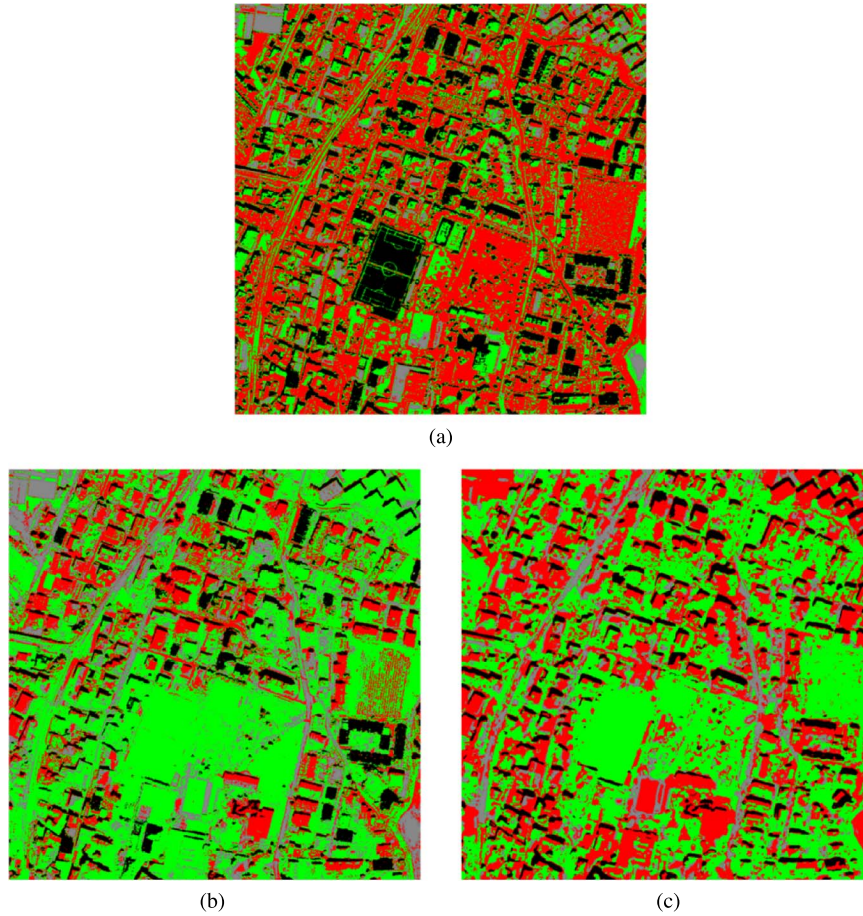


Fig. 7. Classification maps of Trento-2 data set obtained by (a) the SVM when only the panchromatic image is used, (b) by the AP-SVM, and by (c) the HAP-SVM when the standard deviation attribute is used (gray: roads; red: buildings; black: shadow; green: vegetation).

TABLE X

COMPUTATIONAL TIME (IN MINUTES) REQUIRED FOR THE GENERATION OF THE AP, THE HAP AND THEIR CLASSIFICATION WHEN ALL ATTRIBUTES CONSIDERED IN THIS PAPER ARE USED TO BUILD THE APS (ENTIRE TRENTO IMAGE)

Methods	Computational Time (minutes)		
	AP Generation	HAP Generation	Classification
AP-SVM	27	-	19
HAP-SVM	-	48	46

TABLE XI

COMPUTATIONAL TIME (IN MINUTES) REQUIRED FOR THE GENERATION OF THE AP, THE HAP AND THEIR CLASSIFICATION WHEN THE MOMENT OF INERTIA AND THE STANDARD DEVIATION ATTRIBUTES ARE JOINTLY USED TO BUILD THE APS (ENTIRE TRENTO IMAGE)

Methods	Computational Time (minutes)		
	AP Generation	HAP Generation	Classification
AP-SVM	18	-	13
HAP-SVM	-	32	29

As a future development of this paper, we plan to apply histogram-based dimension reduction methods to HAPs to remove the redundancy present in their components. Moreover, we also plan to exploit adaptive size windows (which depend on the outcomes of attribute filters) instead of fixed sized square windows to define the local histograms.

REFERENCES

- [1] L. Bruzzone and B. Demir, "A review of modern approaches to classification of remote sensing data," in *Land Use and Land Cover Mapping Europe*, Practices and Trends, ser. EARSeL Book Series, I. Manakos and M. Braun Eds. Dordrecht, The Netherlands: Springer-Verlag, ch. 9, 2014, pp. 127–143.
- [2] G. Moser, S. B. Serpico, and J. A. Benediktsson, "Land-cover mapping by Markov modeling of spatial-contextual information in very high resolution remote sensing images," *Proc. IEEE*, vol. 101, no. 3, pp. 631–651, Mar. 2013.
- [3] G. Moser and S. B. Serpico, "Combining support vector machines and Markov random fields in an integrated framework for contextual image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2734–2752, May 2013.
- [4] Y. Tarabalka, J. Chanussot, and J. Benediktsson, "Segmentation and classification of hyperspectral images using watershed transformation," *Pattern Recognit.*, vol. 43, no. 7, pp. 2367–2379, Jul. 2010.
- [5] K. Bernard, Y. Tarabalka, J. Angulo, J. Chanussot, and J. A. Benediktsson, "Spectral-spatial classification of hyperspectral data based on a stochastic minimum spanning forest approach," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 2008–2021, Apr. 2012.
- [6] L. Bruzzone and L. Carlini, "A multilevel context-based system for classification of very high spatial resolution images," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 9, pp. 2587–2600, Sep. 2006.
- [7] M.-T. Pham, G. Mercier, and J. Michel, "Pointwise graph-based local texture characterization for very high resolution multispectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 5, pp. 1962–1973, May 2015.
- [8] M. Pesaresi and J. Benediktsson, "A new approach for the morphological segmentation of high-resolution satellite imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 2, pp. 309–320, Feb. 2001.
- [9] M. Dalla Mura, J. A. Benediktsson, J. Chanussot, and L. Bruzzone, "The evolution of the morphological profile: From panchromatic to

- hyperspectral images,” in *Optical Remote Sensing*, Advances in Signal Processing and Exploitation Techniques, S. Prasad, L. Mann Bruce, and J. Chanussot Eds., Berlin, Germany: Springer Verlag, ch. 8, 2011, pp. 123–146.
- [10] M. Dalla Mura, J. A. Benediktsson, B. Waske, and L. Bruzzone, “Morphological attribute profiles for the analysis of very high resolution images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 10, pp. 3747–3762, Oct. 2010.
 - [11] M. Dalla Mura, J. A. Benediktsson, B. Waske, and L. Bruzzone, “Extended profiles with morphological attribute filters for the analysis of hyperspectral data,” *Int. J. Remote Sens.*, vol. 31, no. 22, pp. 5975–5991, Jul. 2010.
 - [12] M. Pedernana, P. Reddy Marpu, M. Dalla Mura, J. A. Benediktsson, and L. Bruzzone, “A novel technique for optimal feature selection in attribute profiles based on genetic algorithms,” *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 6, pp. 3514–3528, Jun. 2013.
 - [13] B. Song *et al.*, “Remotely sensed image classification using sparse representations of morphological attribute profiles,” *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 5122–5136, Aug. 2014.
 - [14] P. Ghamisi, M. Dalla Mura, and J. A. Benediktsson, “A survey on spectral-spatial classification techniques based on attribute profiles,” *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2335–2353, May 2015.
 - [15] O. Chapelle, S. P. Haffner, and V. Vapnik, “Support vector machines for histogram-based image classification,” *IEEE Trans. Neural Netw.*, vol. 10, no. 5, pp. 1055–1064, May 1999.
 - [16] S. Maji, A. C. Berg, and J. Malik, “Classification using intersection kernel support vector machines is efficient,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Anchorage, AK, USA, 2008, pp. 1–8.
 - [17] S. Maji, A. C. Berg, and J. Malik, “Efficient classification for additive kernel SVMs,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 66–77, Jan. 2013.
 - [18] B. Demir and L. Bruzzone, “A novel active learning method in relevance feedback for content based remote sensing image retrieval,” *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2323–2334, May 2015.
 - [19] G. Camps-Valls and L. Bruzzone, “Kernel-based methods for hyperspectral image classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 6, pp. 1351–1362, Jun. 2005.
 - [20] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.
 - [21] R. M. Haralick, Statistical and structural approaches to texture, *Proc. IEEE*, vol. 67, no. 5, pp. 786–804, May 1979.
 - [22] Y. Wei and L. Tao, “Efficient histogram-based sliding window,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, San Francisco, CA, USA, 2010, pp. 3003–3010.
 - [23] M. Sizintsev, K. G. Derpanis, and A. Hogue, “Histogram-based search: A comparative study,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Anchorage, AK, USA, 2008, pp. 1–8.
 - [24] E. Vansteenkiste, S. Gautama, and W. Philips, “Analysing multispectral textures in very high resolution satellite images,” in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Anchorage, AK, USA, 2004, vol. 5, pp. 3062–3065.
 - [25] S. K. Meher, B. U. Shankar, and A. Ghosh, “Wavelet-feature-based classifiers for multispectral remote-sensing images,” *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 6, pp. 1881–1886, May 2007.
 - [26] M. Musci, R. Q. Feitosa, G. A. O. P. Costa, and M. L. Fernandes Velloso, “Assessment of binary coding techniques for texture characterization in remote sensing imagery,” *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 6, pp. 1607–1611, Nov. 2013.
 - [27] P. Soille, “Constrained connectivity for hierarchical image partitioning and simplification,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 7, pp. 1132–1145, Jul. 2008.
 - [28] D. Ehrlich, T. Kemper, X. Blaes, and P. Soille, “Extracting building stock information from optical satellite imagery for mapping earthquake exposure and its vulnerability,” *Nat. Hazards*, vol. 68, no. 1, pp. 79–95, Aug. 2013.
 - [29] M. H. F. Wilkinson, H. Gao, W. H. Hesselink, J. E. Jonker, and A. Meijster, “Concurrent computation of attribute filters on shared memory parallel machines,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 10, pp. 1800–1813, Oct. 2008.
 - [30] A. Farag, R. Mohamed, and A. El-Baz, “A unified framework for map estimation in remote sensing image segmentation,” *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 7, pp. 1617–1634, Jul. 2005.
 - [31] N. Chehata, A. L. Bris, and P. Lagacherie, “Comparison of VHR panchromatic texture features for tillage mapping,” in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Melbourne, Vic., Australia, 2013, pp. 3128–3131.



Begüm Demir (S’06–M’11) received the B.S., M.Sc., and Ph.D. degrees from Kocaeli University, Izmit, Turkey, in 2005, 2007, and 2010, respectively, all in electronic and telecommunication engineering.

She is currently an Assistant Professor with the Department of Information Engineering and Computer Science, University of Trento, Trento, Italy. Her main research interests include image processing and machine learning with applications to remote sensing image analysis. In particular, she conducts

research on remote sensing single-date and time-series image classification, biophysical parameter estimation, content-based remote sensing image retrieval, and analysis of multitemporal images.

Dr. Demir has served as the Organizer for the Special Sessions on Remote Sensing Image Analysis in the IEEE Conference on Signal Processing and Communications Applications from 2010 to 2013. He serves as a Scientific Committee member for the SPIE International Conference on Signal and Image Processing for Remote Sensing and for the International Workshop on the Analysis of Multitemporal Remote Sensing Images. She also serves as a Co-Chair for the Image and Signal Processing for Remote Sensing Workshop organized within the IEEE Conference on Signal Processing and Communications Applications in Turkey since 2014. She is a referee for several journals (such as IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, IEEE TRANSACTIONS ON IMAGE PROCESSING, *Pattern Recognition*, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, *International Journal of Remote Sensing*), and several international conferences.



Lorenzo Bruzzone (S’95–M’98–SM’03–F’10) received the Laurea (M.S.) degree (*summa cum laude*) in electronic engineering and the Ph.D. degree in telecommunications from the University of Genoa, Genoa, Italy, in 1993 and 1998, respectively.

He is currently a Full Professor of telecommunications with the University of Trento, Trento, Italy, where he teaches remote sensing, radar, pattern recognition, and electrical communications. He is the Founder and the Director of the Remote Sensing Laboratory, Department of Information Engineering and Computer Science, University of Trento. He promotes and supervises research on these topics within the frameworks of many national and international projects. Among the others, he is the Principal Investigator of the *Radar for Icy Moon Exploration (RIME)* instrument in the framework of the *Jupiter Icy moons Explorer (JUICE)* mission of the European Space Agency. He is the author or coauthor of 161 papers in referred international journals (111 in IEEE journals), more than 220 papers in conference proceedings, and 17 book chapters. His papers are highly cited, as proven from the total number of citations (more than 12,700) and the value of the h-index (58) (source: Google Scholar). He is the Editor or Co-Editor of 16 books/conference proceedings. His current research interests include remote sensing, radar and synthetic aperture radar, signal processing, and pattern recognition.

Dr. Bruzzone was invited as a keynote speaker in more than 30 international conferences and workshops. He has served as a member of the Administrative Committee of the IEEE Geoscience and Remote Sensing Society since 2009. He has served as an appointed *Distinguished Speaker* of the IEEE Geoscience and Remote Sensing Society since 2012. He has served as the Chair of the SPIE Conference on Image and Signal Processing for Remote Sensing since 2003. He is the Cofounder of the IEEE International Workshop on the Analysis of Multi-Temporal Remote-Sensing Images (MultiTemp) series and is currently a member of the Permanent Steering Committee of this series of workshops. He served as a Guest Coeditor for different special issues of international journals. He has served as the founder Editor-in-Chief of the IEEE GEOSCIENCE AND REMOTE SENSING MAGAZINE since 2013. He currently serves as an Associate Editor for the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING and the *Canadian Journal of Remote Sensing*. He ranked first place in the Student Prize Paper Competition of the 1998 IEEE International Geoscience and Remote Sensing Symposium (Seattle, July 1998). Since that time, he was a recipient of many international and national honors and awards.