HISTOGRAM BASED ATTRIBUTE PROFILES FOR CLASSIFICATION OF VERY HIGH RESOLUTION REMOTE SENSING IMAGES

Begüm Demir and Lorenzo Bruzzone

Dept. of Information Engineering and Computer Science, University of Trento,
Via Sommarive, 14, I-38123 Trento, Italy

ABSTRACT

This paper presents a novel histogram based attribute profiles (HAPs) technique for classification of very high resolution remote sensing images. The HAPs characterize the marginal local distribution of attribute filter responses to model the texture information. This is achieved based on a two steps algorithm. In the first step the standard attribute profiles (AP) are built through sequential application of attribute filters to the considered image. In the second step a local histogram is initially computed for each sample of each image in the APs. Then the local histograms of the same pixel locations in the APs are concatenated. Accordingly, each sample is characterized by a texture descriptor whose components model local distributions of the filter responses. Finally the very high dimensional HAPs are classified by a Support Vector Machine classifier with histogram intersection kernel, which is very effective for high dimensional histogram-based feature representations. Experimental results confirm the effectiveness of the proposed HAPs with respect to standard APs.

Index Terms— Histogram features, morphological attribute filters, mathematical morphology, classification, very high resolution images, remote sensing

1. INTRODUCTION

Accurate classification of very high resolution (VHR) remote sensing (RS) images is very important for generating land-cover maps with high amount of geometrical details. The improvement in the image spatial resolution decreases the problem of mixed pixels present in moderate resolution RS images. Nevertheless, the use of VHR images decreases the statistical separability of different land-cover classes in the spectral domain due to three crucial issues: 1) increased intra-class variability of each land-cover class; 2) reduced inter-class variability between different classes; and 3) relatively low spectral resolution [1]. These issues can be addressed by taking into account in the classification process the spatial information of the scene that characterizes geometrical properties of the objects in an image.

Several methods for the classification of VHR RS images have been developed, differing from each other with respect to adopted strategies to extract the spatial information. Among them, morphological attribute profiles (APs) have gained an increasing interest in recent years [2]-[4]. APs provide a detailed multilevel characterization of a VHR image. This is achieved by the sequential application of morphological attribute filters that characterize the structural information associated to different attributes. The use of multiple attributes (that leads to multi APs) has also been proven effective on the RS images. In [3], the APs are applied to hyperspectral image classification problems. To this end, the Principle Component Analysis is used before applying the attribute filtering to the hyperspectral image features. Although the filtering is applied only to a small number of principal components, APs can still have high dimensionality due to the wide range of possible attributes and their parameters values. This may cause the curse of dimensionality problem (i.e., Hughes phenomenon), and thus resulting in a poor generalization capability. In order to address this problem, the most recent studies have been focused on reducing the dimensionality of APs and automatically selecting the attribute filtering parameters [3]. As an example, a genetic-algorithm based feature selection method is introduced in [4]. For a complete review of AP based methods, we refer the reader to [2]-[4].

Although APs provide high classification accuracies, a direct use of the APs (i.e., the use of only stacked filtered images) can be inadequate to provide a complete characterization of spatial information when complex texture information is present in RS images. To capture the complex texture information, we introduce Histogram based APs (HAPs) that allow an efficient and accurate modeling of texture information from APs. The HAPs consist of histograms of attribute filter responses and are capable of integrating spatial and spectral information more efficiently than the standard APs. After the HAPs are constructed, they are given as input to a Support Vector Machine (SVM) classifier with Histogram Intersection (HI) kernel. The HI kernel is a very effective kernel for high dimensional histogram based feature representations. Due to the classification of HAPs with SVMs using HI kernel, unlike the other AP based methods, the HAPs are not affected by the Hughes phenomenon and result in a high generalization capability of SVMs.

2. HISTOGRAM BASED ATTRIBUTE PROFILES

Let $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, ..., \mathbf{X}_P]$ be a RS image made up of P spectral channels, where $\mathbf{X}_p = \{x_1, x_2, ..., x_B\}$ is the p-th channel consisting of B samples and x_j is the j-the pixel. The standard AP-based techniques proposed in the literature exploit the stacked filtered images (obtained by applying a sequence of attribute filters) with the original image features. In order to characterize the texture elements as well as their local distributions accurately, in this paper we introduce the HAPs. The HAPs are obtained based on a two steps algorithm. In the first step the standard APs are built. Then, in the second step, the HAPs are generated by: i) estimating a local histogram at each sample location within each filtered image in APs, and ii) stacking local histograms to obtain in a texture descriptor for the corresponding pattern. Therefore, the HAPs consist of marginal local distributions of responses from a series of morphological attribute filters by concatenating local histograms of all images in APs. Each step is explained in detail in the following.

2.1. Step 1: Generation of morphological attribute profiles

The first step aims at extracting the AP of an image $\mathbf{X}$. This is achieved by applying a sequence of attribute thinning ϕ^λ and attribute thickening γ^λ operations with different filter thresholds λ to each image band $\mathbf{X}_p \in \mathbf{X}, p = 1, 2..., P$, independently from each other (note that a reduced subset of features extracted from the P spectral images can be used). Without losing in generality, let us assume that the image $\mathbf{X}$ contains only one spectral channel, i.e., P=1 and accordingly $\mathbf{X} = \mathbf{X}_1$.The attribute profile of $\mathbf{X}$ for a given attribute and an ordered sequence of m thresholds $\{\lambda_1, \lambda_2, ..., \lambda_m\}$ is attained as:

$$AP(\mathbf{X}) = \{\mathbf{X}^{\phi^{\lambda_m}}, ..., \mathbf{X}^{\phi^{\lambda_1}}, \mathbf{X}\ , \mathbf{X}^{\gamma^{\lambda_1}}, ..., \mathbf{X}^{\gamma^{\lambda_m}}\} \tag{1}$$

where $\mathbf{X}^{\phi^i} = \{x_1^{\phi^{\lambda_i}}, x_2^{\phi^{\lambda_i}}, ..., x_B^{\phi^{\lambda_i}}\}$ and $\mathbf{X}^{\gamma^i} = \{x_1^{\gamma^{\lambda_i}}, x_2^{\gamma^{\lambda_i}}, ..., x_B^{\gamma^{\lambda_i}}\}$ are the filtered images obtained after applying the attribute thinning and the thickening, respectively, by considering the i-th threshold λ_i. The j-th pattern $x_j \in \mathbf{X}, j = 1, 2..., B$ is characterized by a feature vector $\mathbf{x}_j^{AP}$ in the AP defined as:

$$\mathbf{x}_j^{AP} = \{x_j^{\phi^{\lambda_m}}, x_j^{\phi^{\lambda_{m-1}}}, ..., x_j, ...x_j^{\gamma^{\lambda_{m-1}}}, x_j^{\gamma^{\lambda_m}}\} \tag{2}$$

In other words, $\mathbf{x}_j^{AP}$ is defined by the stacked filtered responses associated to the related sample x_j.

Depending on the attributes selected, the APs capture different types of structural information. Possible attributes are: i) area (which models the size of the regions); ii) moment of inertia (which models the elongation of the objects in the scene); and iii) standard deviation (which models region homogeneity). If more than one image band (i.e., multispectral and hyperspectral images) is considered, extended APs are defined by concatenating different APs computed on each image band. For further information on the APs, the reader is referred to [2]-[4].

2.2. Step 2: Generation of the texture descriptors and characterization of the HAPs

The second step aims at extracting texture information from each filtered image in the APs to provide a complete spatial characterization of the considered image. To this end, local histograms $\{h(x_j^{\phi^{\lambda_m}}), h(x_j^{\phi^{\lambda_{m-1}}}), ..., h(x_j), ...h(x_j^{\gamma^{\lambda_{m-1}}}), h(x_j^{\gamma^{\lambda_m}})\}_{j=1}^B$ of each sample $\{x_{j,1}^{\phi^m}, x_{j,1}^{\phi^{m-1}}, ..., x_{j,1}, ...x_{j,1}^{\gamma^m}, x_{j,1}^{\gamma^{m-1}}\}_{j=1}^B$, where $h(x_j^{\phi^{\lambda_m}})$ is the local histogram of $x_j^{\phi^{\lambda_m}}$, are initially computed. Local histograms are built to represent the texture in a neighborhood system of considered samples. In order to estimate the local histogram $h(x_j^{\phi^{\lambda_m}})$ of a sample $x_j^{\phi^{\lambda_m}}$, 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). Then, the range is divided into b histogram intervals. The histogram $h(x_j^{\phi^{\lambda_m}})$ is calculated by counting how many of the neighborhood samples within a $w \times w$ pixels spatial window centered at the corresponding pixel $x_j^{\phi^{\lambda_m}}$ location belong to each interval. Thus, the local histogram is a feature vector that consists of marginal local distributions of filter responses at each pixel location. Then, a texture descriptor $H(x_j)$ for each sample $x_j, j = 1, 2, ...B$ is defined by integrating all the local histograms $\{h(x_j^{\phi^{\lambda_m}}), h(x_j^{\phi^{\lambda_{m-1}}}), ..., h(x_j), ...h(x_j^{\gamma^{\lambda_{m-1}}}), h(x_j^{\gamma^{\lambda_m}})\}_{j=1}^B$. This is achieved by stacking the local histograms of the j-th pixel location in APs as:

$$H(x_j) = \{h(x_j^{\phi^{\lambda_m}}), h(x_j^{\phi^{\lambda_{m-1}}}), ..., h(x_j), ...h(x_j^{\gamma^{\lambda_{m-1}}}), h(x_j^{\gamma^{\lambda_m}})\} \tag{3}$$

As seen in (3), texture descriptors integrate responses of different filters to form a texture feature vector that describes spectral and spatial properties. Fig.1 shows a qualitative illustration on how generating a texture descriptor for the j-th pixel location in the AP is shown in.

Note that, for simplicity, the example shows an AP obtained by including one attribute with only one threshold value, i.e., m=1.

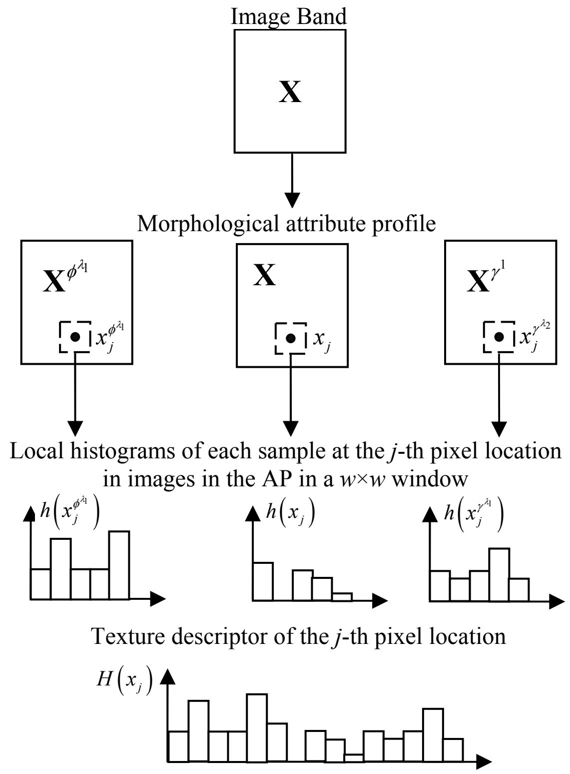

Fig. 1. Generation of a texture descriptor for the j-th pixel location in the AP

Finally, HAPs are built by all the texture descriptors obtained for all the samples in a scene. To classify the HAPs, we use an SVM classifier with histogram intersection (HI) kernel, which has been widely used for histogram comparisons in the computer vision community. Note that the HI kernel is a positive definite parameter-free kernel for histogram based features [6], and SVMs with HI kernel can generalize well on difficult classification problems where the features are high dimensional histograms (i.e., discrete densities) [5]. Robustness to feature dimensionality is crucial for the success of the proposed HAPs. This is due to the fact that the HAPs can have very high dimensionality if: 1) a dense sampling of the values of the filters' parameters is selected to generate APs; 2) the selected histogram bin number is very high; and 3) the image includes several spectral channels. This issue can lead to the Hughes phenomenon, i.e., the curse of dimensionality problem. However, due to the high generalization capability of SVMs with the HI kernel, this issue does not affect the performance of the proposed HAP. Note that SVM with HI kernel has been used with very high dimensional histogram features in computer vision problems (i.e., the dimension may vary from thousands to tens of thousands).

3. EXPERIMENTAL RESULTS

Experimental analyses are conducted on a panchromatic image having geometric resolution of 0.6 m and acquired by the Quickbird sensor on the city of Trento, Italy, on July 2006. The available ground reference samples are representative of the four main land-cover classes (i.e., road, building, shadow and vegetation). They were divided to derive a training set of 885 samples and a test set of 6120 samples. Table I shows the land-cover classes and the related number of samples used in the experiments.

Table 1: Number of samples of each class in the training and the test sets

Land-Cover Class	Training Set	Test Set
Roads	199	907
Buildings	209	3330
Shadow	255	853
Vegetation	222	1030
Total	885	6120

Note that although in the experiments we focus on the panchromatic channel only, the proposed HAP can be also applied to the other multispectral channels. Three different attributes were considered for the construction of the AP: i) the area; ii) moment of inertia; and iii) the standard deviation. In the experiments the AP with the area attribute is obtained with λ values of 49, 169, 361, 625, 961, 1369, 1849, 2401, while that with the standard deviation is obtained with λ values of 10, 20, 30, 40, 50, 60, 70, 80. The AP with the moment of inertia attribute is constructed by selecting values of λ equal to 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 as in [2]. After the HAPs are constructed, they are given as input to the SVM classifier (denoted as HAP-SVM). The proposed HAP-SVM is implemented by using the HI kernel, which is parameter free. In the model selection, the regularization parameter of the 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 an unit step size increment and the local window size is tested within the range of [5-11] with a step size of 2. Then, the best values are obtained using a 5-fold cross validation approach for both data sets.

We compare the results of the proposed HAP-SVM with those obtained by i) the SVM classification applied to original image features (denoted as the SVM); and ii) the SVM classification applied to the standard APs (denoted as the AP-SVM). The SVM and the AP-SVM are defined by using Radial Basis Function (RBF) kernel. This choice is done due to the fact that the RBF is much effective than HI kernel if the features are not histograms (also note that in the papers on standard AP the RBF kernel is used). In the experiments the spread of the RBF kernel parameter is chosen by a grid-search model selection.

Table 2 shows the classification results obtained by the standard SVM, the AP-SVM and the proposed HAP-SVM when using different attributes. From the table, one can observe that the accuracies obtained by the proposed HAP-SVM are significantly higher than those yielded by the SVM and the AP-SVM, independently from the selected attributes. As an example, the proposed HAP-SVM yields an accuracy of 80.5% when the standard deviation attribute is used, whereas the standard AP-SVM provides an accuracy of 56.6% with the same attribute. For a qualitative analysis, Fig. 2 shows the related classification maps. By analyzing the figure, 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 demonstrates that the proposed HAP-SVM characterizes more accurately the spatial/texture information compared to the AP-SVM. It is worth noting that the good performances of the proposed HAP are relying on: i) capturing the complex texture from APs based on histogram based texture descriptors; and also ii) high generalization ability of SVMs with HI kernel in high dimensional spaces.

Table 2: Overall classification accuracy (in %) obtained by the standard SVM, the AP-SVM and the proposed HAP-SVM

Features/Attributes	SVM	AP-SVM	**HAP-SVM**
Only Panchromatic	55.5	-	**-**
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**
Area+Moment of Inertia +Standard Deviation	-	86.1	**90.7**

4. CONCLUSION

In this paper, we propose histogram based APs (HAPs) for the classification of remote sensing images with SVMs. The proposed HAPs provide a more complete and detailed characterization of the spatial information of the scene with respect to the standard APs, representing efficiently the texture information. The HAPs have high dimensional feature vectors given as input to an SVM classifier with HI kernel that is robust to the Hughes phenomenon for histogram features. Experimental results show that the proposed HAPs significantly outperform the standard APs, due to an effective modeling of the marginal local distribution of attribute filter responses. The proposed method is promising as it significantly improves the classification accuracy without being affected by Hughes phenomenon. As a future development of this work, we plan to apply histogram based dimension reduction methods to HAPs to remove the redundancy present in their components.

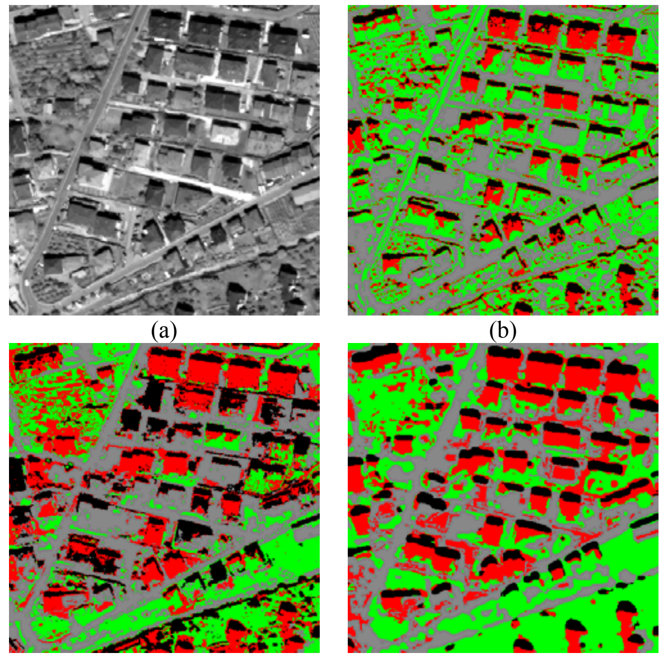

Fig.2.Trento data set: (a) the panchromatic image; and the classification maps obtained by (b) the SVM; (c) the AP-SVM and (d) the HAP-SVM when the standard deviation attribute is used (gray: roads; red: buildings; black: shadow; green: vegetation)

REFERENCES

[1] L. Bruzzone, B. Demir, A Review of Modern Approaches to Classification of Remote Sensing Data, in Land Use and Land Cover Mapping Europe, Practices and Trends, Eds: I. Manakos, M. Braun, EARSeL Book Series, Springer Verlag, Chapter 9, 2014, pp. 127-143.

[2] M. Dalla Mura, J.A. Benediktsson, B. Waske, and L. Bruzzone, “Morphological attribute profiles for the analysis of very high resolution images”, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 10, pp. 3747-3762, 2010.

[3] P. Ghamisi, M. Dalla Mura, J. A. Benediktsson, “A survey on spectral-spatial classification techniques based on attribute profiles”, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 5, May 2015.

[4] M. Pedergnana, P. Reddy Marpu, M. Dalla Mura, J. A. Benediktsson, L. Bruzzone, “A novel technique for optimal feature selection in attribute profiles based on genetic algorithms”, *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, pp. 3514-3528, 2013.

[5] Chapelle, S. P. Haffner, and V. Vapnik. “Support vector machines for histogram-based image classification”, *IEEE Transactions on Neural Networks*, vol. 10, no. 5, pp. 1055–1064, May 1999.

[6] S. Maji, A.C. Berg, and J. Malik, “Classification using intersection kernel support vector machines is efficient,” Proc. *IEEE Conf. Computer Vision and Pattern Recognition*, Anchorage, AK, USA, 2008, pp.1-8.