CLASSIFICATION OF HYPERSPECTRAL IMAGES WITH EXTENDED ATTRIBUTE PROFILES AND FEATURE EXTRACTION TECHNIQUES

Mauro Dalla Mura[a,b], *Jon Atli Benediktsson*[b] *and Lorenzo Bruzzone*[a]

[a] Department of Information Engineering and Computer Science, University of Trento
Via Sommarive, 14 I-38123, Povo, Trento, Italy,
E-mail: dallamura@disi.unitn.it, lorenzo.bruzzone@ing.unitn.it
[b] Faculty of Electrical and Computer Engineering, University of Iceland,
Hjardarhaga 2-6, 101 Reykjavik, Iceland,
E-mail: benedikt@hi.is

1. INTRODUCTION

It is well known that a proper exploitation of the spatial information is essential in the analysis of remote sensing images of very high geometrical resolution (VHR), especially when dealing with urban areas. Approaches based on mathematical morphology have already proved to be effective for the processing of VHR images [1]. In particular, techniques that employ morphological connected filters [2], (e.g., opening and closing by reconstruction, attribute filters, etc.) are able to process an image without distorting the geometrical characteristics of the structures in the scene, which is a fundamental feature of this kind of imagery. Moreover, these operators can be used for performing a multi-level analysis of the images which permits the extraction of structural information of the objects in the scene. In [3], Pesaresi and Benediktsson defined a data structure, called morphological profile (MP), generated by a sequence of the morphological operators of opening and closing by reconstruction of increasing size computed on a VHR panchromatic image. Since the MPs perform a multi-scale decomposition of the original image, they provide features which are useful for inferring the size of the objects in the image. In [4], feature extraction techniques were proposed in order to reduce the high redundancy generated by the MPs. The concept of MP was extended to the analysis of hyperspectral data in [5], by concatenating in a single data structure, called extended morphological profile (EMP), the MPs computed on the first principal components of the original hyperspectral data. Moreover, the morphological features extracted by the EMP were considered along with the original hyperspectral data in [6]. Although the features extracted by the morphological profiles generated with operators by reconstruction provide important information on the geometry of the structures, they only model the size of the structures in the image. Since, other structural and spectral features might be interesting rather than only the size of the objects, in [2] we recently proposed to use morphological attribute filters for the analysis of VHR panchromatic images. Attribute filters can process the image according to many different attributes not only related to the concept of scale [7]. In [8] attribute profiles (APs), a generalization of the MP based on attribute filters, were proposed. Thanks to the introduction of attribute filters in the analysis, APs can model different spatial features of the image. Analogously, for hyperspectral data, the concept of EAP was generalized by introducing the definition of extended attribute profiles (EAPs) generated by attribute filters [9]. The use of multiple profiles built on different attributes has proved to provide richer descriptions of the spatial information contained in the scene with respect to an approach bases on the conventional profiles built by the operators by reconstruction [8]-[9]. Nevertheless, the presence of redundant information already noticeable in MPs and EMPs is increased when considering multiple APs/EAPs computed on the same images. When a classification task is considered, the high dimensionality and the high redundancy that can be reached by the features generated by the morphological processing can decrease the performaces of the classification, especially if a low number of training sample is available. Furthermore, the use of feature extraction approaches for the extended attribute profiles becomes very important in order to make the best use of the available computational resources in the classification.

In this paper we propose an approach for the classification of VHR hyperspectral images based on extended profiles, computed with attribute filters, and feature extraction techniques. In particular, at first, morphological attribute profiles are computed on the first principal components of a VHR hyperspectral image leading to the generation of EAPs based on different attributes. Subsequently, the dimensionality of the generated profiles is reduced by the discriminant analysis feature extraction (DAFE) and the decision boundary feature extraction (DBFE) techniques. The features extracted are then classified by a random forest classifier [10].

2. PROPOSED CLASSIFICATION APPROACH

In this section the concepts at the basis of the proposed approach are presented.

2.1. Extended Attribute Profiles

We recall the definition of AP as the concatenation of two profiles, Π_{ϕ^T} and Π_{γ^T} generated by an extensive (i.e., closing or thickening) and anti-extensive (i.e., opening or thinning) morphological attribute operator respectively:

$$AP(f) = \Pi_i : \left\langle \begin{array}{llll} \Pi_i = \Pi_{\phi^{T_\lambda}}, & with & \lambda = (n - 1 + i) & \forall \lambda \in [1, n]; \\ \Pi_i = \Pi_{\gamma^{T_\lambda}}, & with & \lambda = (i - n - 1) & \forall \lambda \in [n + 1, 2n + 1]. \end{array} \right. \tag{1}$$

with ϕ^{T_λ} and γ^{T_λ} denoting a morphological attribute closing and opening respectively for an increasing criterion T and λ a set of scalar values used as reference in the filtering procedure. If the attribute considered in the analysis is not increasing, then the criterion is also not increasing. This leads to have operators of thickening and thinning instead of closing and opening respectively [7].

An EAP can be computed by concatenating the APs computed by considering the same attribute on the c principal components, PCs, extracted from the original hyperspectral data. Thus, the EAP can be formally defined as:

$$EAP = \{AP(PC_1), AP(PC_2), \ldots, AP(PC_c)\}. \tag{2}$$

The attribute chosen for building the APS and the corresponding EAP, determines the type of processing that is performed on the image and the related characteristics of the structures that can be modeled. For example, if the area of the regions is considered as an attribute, then the multi-level analysis performed on the image leads to the extraction of information on the size of the structures. Conversely, by considering attributes such as, the shape factor, the moments invariant, the eccentricity, etc. the geometry of the regions is characterized. Furthermore, with attributes related to the contrast or to the spectral homogeneity of the regions, other different features can be extracted. In order to perform an analysis of the image suitable for discriminating among different thematic objects present in the scene, the number of the levels selected in the definition of the profile has to be sufficient to extract the different characteristics of the data. On the other hand, the selection of too many levels in the analysis would lead to both a great presence of redundant information and the increase of the dimensionality of the features. For this reason, considering feature extraction techniques for reduce the effect of noisy features can be of help in the analysis.

2.2. Feature Extraction Techniques

Two well known feature extraction techniques are considered: DAFE and DBFE. The technique of discriminant analysis feature extraction performs the Fisher discriminant analysis [11], which is a linear projection of the multivariate data on the $C - 1$ orthogonal directions that are the most discriminant for the C Gaussian distributions estimated on the thematic classes in the image. We remind that if the number of features is less than the number of classes, then the number of features becomes the maximum number of discriminant components that can be extracted. The most discriminant features are found as those which maximize a separability criterion $J = tr(\Sigma_W^{-1}\Sigma_B)$ where Σ_W denotes the within-scatter matrix which gives a measure of the overlapping of the different distributions and Σ_B the between-scatter matrix which indicates the separability of the means of the distributions.

The Decision Boundary Feature Extraction (DBFE) technique, in contrast to DAFE, does not take into account the statistics of the class distributions, since it is based on the analysis of the boundary that separates the different classes. In greater detail, discriminant features in the DBFE are extracted as directions orthogonal to the decision boundary.

3. EXPERIMENTAL RESULTS

The data set considered in the experiments is a 610×340 pixels very high resolution hyperspectral image acquired by the airborne sensor ROSIS-03 (Reflective Optics Systems Imaging Spectrometer) on the University campus of the city of Pavia, Italy. The geometrical resolution of the image is 1.3 m. From the 115 spectral bands acquired on the range 0,43 μm to 0,86 μm 12 were discarded due to noise. The geospatial objects surveyed by this image are belonging to nine thematic classes: Trees, Asphalt, Bitumen, Gravel, Metal sheets, Shadows, Self-blocking Bricks, Meadows, and Bare soil. For those, a training and reference data, used as test in the experiments, were available.

From the hyperspectral data four PCs were considered for the analysis in order to explain more than the 99% of the total variance of the data. Subsequently, four EAPs were computed with different attributes: i) a, area of the regions; ii) d, length of the diagonal of the box bounding the region; iii) i, first moment invariant of Hu, or moment of inertia, [12]; and iv) s, standard deviation of the gray-level values of the pixels in the regions. The area and the length of the diagonal of the bounding box are increasing attributes that are useful to perform a multi-scale analysis of the data. The moment of inertia attribute is also a purely geometric descriptor that results sensitive to measure the elongation of the regions. Since it is scale invariant it is not increasing and can be employed for extracting information on the geometry of the regions regardless their scale. Finally, the standard deviation attribute measures the homogeneity of the intensity values of the pixels belonging to each region in the image and it gives information that is not dependent on the geometry of the regions but on the spectral contrast of the pixels. Four reference values, λ in 1, were considered for building each of the four EAPs, leading to 36-dimensional profiles (composed by four APs of 9 levels computed on the PCs). The λs are listed in the following:

1. EAP$_a$: λ_a = [100 500 1000 5000];
2. EAP$_d$: λ_d = [10 25 50 100];
3. EAP$_i$: λ_i = [0.2 0.3 0.4 0.5];
4. EAP$_s$: λ_s = [20 30 40 50].

Table 1. Overall Accuracy (OA), Average Accuracy (AA) and Kappa coefficient of the results obtained without feature extraction.

	Spectr.	PC_s	EAP_a	EAP_d	EAP_i	EAP_s	EAP_{all}
Features	103	4	36	36	36	33	129
OA (%)	71.31	70.04	**92.75**	86.58	76.26	77.63	89.64
AA (%)	82.00	78.88	**92.73**	87.43	84.30	85.58	89.74
Kappa	0.65	0.63	**0.90**	0.82	0.70	0.72	0.86

Table 2. Overall Accuracy (OA), Average Accuracy (AA) and Kappa coefficient of the results obtained with DAFE.

	EAP_a	EAP_d	EAP_i	EAP_s	EAP_{all}
Features	8	8	8	8	32
OA (%)	**87.65**	86.95	71.27	71.70	84.29
AA (%)	89.48	89.48	83.93	85.44	**90.13**
Kappa	**0.84**	0.83	0.65	0.66	0.80

From each of the profiles considered the feature extraction techniques DAFE and DBFE were applied. The estimation of the covariance matrices of the different thematic distributions, required for the computation of DAFE was done by the leave-one-out covariance estimator since the high redundancy present in the profiles led to singular matrices by using the conventional estimator.

Preliminary results were obtained by classifying the data with a random forest classifier created with 200 trees [10] and considering as the number of variables, investigated in each split of the trees, the square root of the number of the features of the data, as suggested in [10]. The classification results were quantitatively evaluated by measuring the Overall Accuracy (OA), the Average Accuracy (AA) and the Kappa coefficient (K) on the reference data.

In Table 1, the accuracies obtained without the feature extraction techniques are presented. It is possible to notice how similar results are obtained by considering only the hyperspectral values (Spectr) and the first PCs. As expected, by considering the EAPs, higher accuracies are obtained. In particular, the EAP$_a$ performed extremely well with respect to the EAPs with other attributes. When considering all the profiles together, the obtained accuracies are higher than all the single EAPs except the EAP$_a$.

The results obtained by considering the DAFE technique are presented in Table 2. The dimensionality was reduced to 8 features for each EAP. The EAP$_{all}$ considered all the features extracted by the four profiles. The achieved accuracies are comparable to those obtained by the profiles without feature extraction but, in general, slightly lower. However, the EAP$_a$ achieved the best OA and K accuracies while the best AA accuracy was obtained by the EAP$_{all}$.

In Table 3 are reported the results obtained by considering the DBFE technique. The number of feature extracted by the analysis was selected in order to achieve more than the 99% of the total variance of the data. The best accuracies were obtained by considering all the features extracted from each single EAP together.

When comparing the results in Table 2 and 3 it is possible to notice that by performing DAFE the accuracy achieved is higher than by considering the DBFE technique for EAP$_a$, EAP$_d$ and EAP$_i$. Conversely, better results were obtained by applying DBFE instead of DAFE for the EAP$_s$ and EAP$_{all}$.

From the preliminary obtained results it can be noticed how the use of feature extraction techniques in most of the cases led to lower, even if still comparable, values of the accuracies with respect to considering the data in its original dimensions. This behaviour can be due to i) the availability of a set of training samples that can fully describe the class variabily and ii) the use of a random forest classifier which has proved to be robust in handling redundant features.

In conclusion based on the preliminary results carried out, it is possible to notice how by using feature extraction techniques the dimensionality of the data is significantly reduced (i.e., requiring less computational resources) while achieving results which are comparable to those obtained by the data with full dimensionality. In the full paper further results of our on-going research aimed at exploring the role of feature extraction techniques with EAPs will be presented.

Table 3. Overall Accuracy (OA), Average Accuracy (AA) and Kappa coefficient of the results obtained with DBFE.

	EAP_a	EAP_d	EAP_i	EAP_s	EAP_{all}
Features	17	18	20	17	72
OA (%)	86.22	80.34	70.48	76.66	**88.58**
AA (%)	89.35	85.94	82.99	84.73	**90.40**
Kappa	0.82	0.75	0.64	0.71	**0.85**

4. REFERENCES

[1] P. Soille and M. Pesaresi, "Advances in mathematical morphology applied to geosciences and remote sensing," *IEEE Trans. Geosci. Remote Sensing*, vol. 40, pp. 2042–2055, 2002.

[2] M. Dalla Mura, J. A. Benediktsson, B. Waske, and L. Bruzzone, "Morphological attribute filters for the analysis of very high resolution images," in *IEEE International Geoscience & Remote Sensing Symposium, Cape Town, South Africa, 13-17 July, in press*, 2009.

[3] M. Pesaresi and J. A. Benediktsson, "A new approach for the morphological segmentation of high-resolution satellite imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 2, pp. 309–320, 2001.

[4] J. A. Benediktsson, M. Pesaresi, and K. Amason, "Classification and feature extraction for remote sensing images from urban areas based on morphological transformations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 9, pp. 1940–1949, 2003.

[5] J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 480–491, 2005.

[6] M. Fauvel, J. A. Benediktsson, J. Chanussot, and J. R. Sveinsson, "Spectral and spatial classification of hyperspectral data using svms and morphological profiles," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 11, pp. 3804–3814, 2008.

[7] Edmond J. Breen and Ronald Jones, "Attribute openings, thinnings, and granulometries," *Computer Visual and Image Understanding*, vol. 64, no. 3, pp. 377–389, 1996.

[8] 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*, submitted.

[9] M. Dalla Mura, J. A. Benediktsson, B. Waske, and L. Bruzzone, "Extended profiles with morphological attribute filters for the analysis of hyperspectral data," *International Journal of Remote Sensing*, submitted.

[10] Leo Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[11] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed. London: Academic Press, 1990.

[12] Ming-Kuei Hu, "Visual pattern recognition by moment invariants," *IRE Transactions on Information Theory*, vol. 8, no. 2, pp. 179–187, 1962.