Unmixing Prior to Supervised Classification of Urban Hyperspectral Images

Inmaculada Dópido and Antonio Plaza
Hyperspectral Computing Laboratory
Department of Technology of Computers and Communications
Escuela Politécnica de Cáceres, University of Extremadura, Cáceres, Spain
E-mail: {inmaculadadopido, aplaza}@unex.es

***Abstract*— Supervised classification of urban hyperspectral images is a very challenging task due to the generally unfavorable ratio between the number of spectral bands and the number of training samples available *a priori*, which results in the Hughes phenomenon. Training samples are particularly challenging to be collected in urban environments. A possible solution is to reduce the dimensionality of the data to the right subspace without losing the original information that allows for the separation of classes. In this paper, we propose a new strategy for feature extraction prior to supervised classification of urban hyperspectral data which is based on spectral unmixing concepts. The proposed strategy includes the sub-pixel information that can be obtained with spectral unmixing techniques into the classification process, and does not penalize classes which are not relevant in terms of variance or signal-to-noise ratio (SNR) as it is the case with other transformations such as principal component analysis (PCA) or the minimum noise fraction (MNF). Experiments using urban hyperspectral image data collected by the reflective optics spectrographic imaging system (ROSIS) over the city of Pavia in Italy are discussed, using the support vector machine (SVM) classifier as a baseline for demonstration purposes.**

I. Introduction

Hyperspectral imaging is concerned with the measurement, analysis, and interpretation of spectra acquired from a given scene (or specific object) at a short, medium or long distance by an airborne or satellite sensor [1]. In many studies, hyperspectral analysis techniques are divided into full-pixel and mixed-pixel classification techniques [2], [3], [4], where each pixel vector defines a *spectral signature* or *fingerprint* that uniquely characterizes the underlying materials at each site in a scene. Full-pixel classification techniques assume that each pixel vector measures the response of one single underlying material. Often, however, this is not a realistic assumption. If the spatial resolution of the sensor is not fine enough to separate different pure signature classes at a macroscopic level, these can jointly occupy a single pixel, and the resulting spectral signature will be a composite of the individual pure spectra, often called *endmembers* in hyperspectral imaging terminology [5].

Under the full-pixel assumption, several techniques have been applied to extract relevant information from urban hyperspectral data [6], [7]. The generally unfavorable ratio between the small number of available training samples in this context (due to the complexity of obtaining labeled samples in urban environments) and the high number of features available *a priori* results in the Hughes phenomenon [8], which makes reliable estimation of statistical class parameters a very challenging goal. One possible approach to handle the high-dimensional nature of hyperspectral data sets is to consider the geometrical rather than the statistical properties of the classes [9], [10]. The good classification performance demonstrated by successful machine learning techniques such as the support vector machine (SVM) [2], [11], [12], using spectral signatures as input features, can be improved by the incorporation of intelligent feature extraction strategies which can reduce the dimensionality of the data to the right subspace without losing the original information that allows for the separation of classes [13]. Techniques used for this purpose include principal component analysis (PCA) [4] or the minimum noise transform (MNF) [14]. These transforms respectively maximize the amount of data variance and signal-to-noise ratio (SNR) yielding a transformed dataset in a new uncorrelated coordinate system. However, both PCA and MNF focus upon the signal variation of the data set *as a whole* by maximizing the variation contained in the first transformed components, relegating variations of less significant size to low order components. This can result in lower classification accuracies for small classes. To address this issue, inclusion of spatial features has been used to improve classification in urban areas [6], [7], using techniques such as morphological profiles [10], [13].

In this paper, we propose a new strategy for feature extraction prior to supervised classification of urban hyperspectral data which is based on mixed-pixel classification concepts. The proposed unmixing-based feature extraction strategy consists of applying a standard spectral unmixing processing chain to extract relevant features prior to classification. A variation of the proposed approach also tested in this work consists of using a partial unmixing strategy to address the difficulty in estimating the number of endmembers in the input data. In both cases, spectral unmixing provides additional information for classification in urban hyperspectral analysis scenarios, since the sub-pixel composition of training samples can be used as part of the learning process of the classifier. In addition, a distinguishing feature of the proposed approach is that it does not penalize classes which are less relevant than others in terms of variance or SNR.

978-1-4244-8657-1/11/$26.00

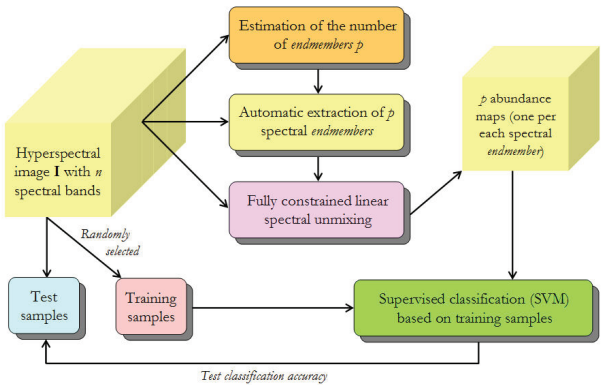


Fig. 1. Unmixing-based feature extraction chain #1.

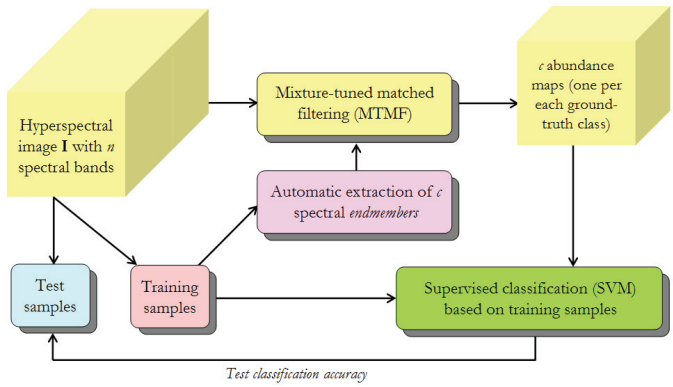


Fig. 2. Unmixing-based feature extraction chain #2.

II. Unmixing-based Feature Extraction

A. Unmixing-based Feature Extraction Chain #1

In this subsection we describe our first approach to design an unmixing-based feature extraction chain which can be summarized by the flowchart in Fig. 1. First, we estimate the number of endmembers, p, directly from the original n-dimensional hyperspectral image $\mathbf{I}$. For this purpose, we use in this work two standard techniques widely used in the literature such as the HySime method [15] and the virtual dimensionality (VD) concept [16]. Once the number of endmembers p has been estimated, the correct determination of a set of endmembers from the original hyperspectral image is a challenging issue. Several algorithms have been developed over the last decade for accomplishing this goal in automatic or semi-automatic fashion [17]. Here, we use an orthogonal subspace projection (in OSP) technique in [18] which has been shown in previous work to provide a very good trade-off between the signature purity of the extracted endmembers and the computational time to obtain them. Finally, a fully unconstrained linear spectral unmixing estimate (with abundance non-negativity and sum-to-one constraints) is obtained for each endmember in each pixel of the scene [19]. The resulting information (i.e. a p-dimensional stack of fractional abundance maps) is provided to a standard SVM classifier for classification purposes.

B. Unmixing-based Feature Extraction Chain #2

Our main motivation for introducing a second unmixing-based feature extraction chain is the fact that the estimation of the number of endmembers p in the original image is a very challenging issue. In order to address this issue, Fig. 2 describes a new unmixing-based feature extraction chain in which the endmembers are extracted from the available (labeled) training samples instead of from the original image. This introduces two main properties with regards to the other unmixing-based feature extraction chains presented in this section: 1) the number of endmembers to be extracted is given by the total number of different classes, c, in the labeled samples available in the training set, and 2) the endmember searching process is conducted only on the training set, which reduces computational complexity. The increase in computational performance comes at the expense of introducing an additional consideration: it is likely that the actual number of endmembers in the original image, p, is larger than the number of different classes comprised by available labeled training samples, c. Therefore, in order to unmix the original image we need to address a *partial unmixing* problem (in which not all endmembers may be available *a priori*). A successful technique to estimate abundance fractions in such partial unmixing scenarios is mixture-tuned matched filtering (MTMF) [20], also known in the literature as constrained energy minimization (CEM), which combines linear spectral unmixing and statistical matched filtering. From matched filtering, it inherits the ability to map a single known target without knowing the other background endmember signatures, unlike the standard linear unmixing model. From spectral mixture modeling, it inherits the leverage arising from the mixed pixel model and the constraints on feasibility.

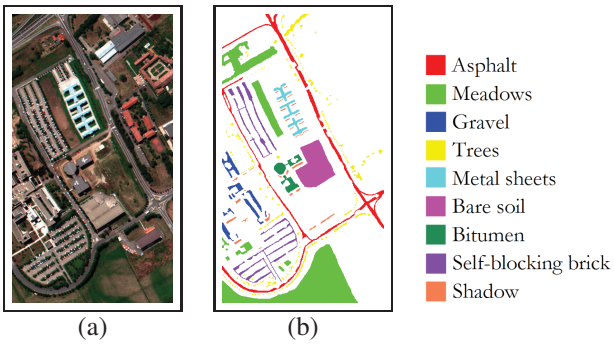
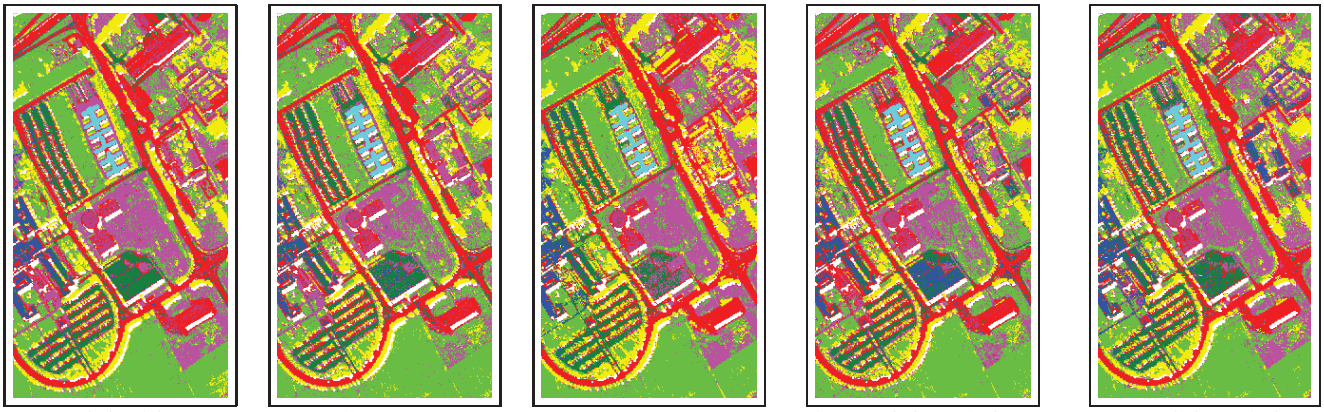


Fig. 3. (a) False color composition of the ROSIS Pavia scene. (b) Ground truth-map containing 9 mutually exclusive land-cover classes (right).

III. Experimental Results

The data set used in experiments was collected by the ROSIS optical sensor over an urban area centered at the University of Pavia, Italy. The flight was operated by the Deutschen Zentrum for Luftund Raumfahrt (DLR, the German Aerospace Agency) in the framework of the HySens project, managed and sponsored by the European Union. The image size in pixels is 610×340, with very high spatial resolution



(a) Original image (b) PCA (c) MNF (d) Unmixing chain #1 (e) Unmixing chain #2

Fig. 4. Classification results for the ROSIS Pavia scene (using an SVM classifier with Gaussian kernel, trained with 5% of the available samples).

of 1.3 meters per pixel. The number of data channels in the acquired image is 115 (with spectral range from 0.43 to 0.86 μm). Fig. 3(a) shows a false color composite of the image, while Fig. 3(b) shows nine ground-truth classes of interest, which comprise urban features, as well as soil and vegetation features.

Before describing the results obtained in experimental validation, we first briefly describe the adopted supervised classification system. Firstly, relevant features for classification are extracted from the original image. The five types of input features considered in the classification experiments conducted in this work can be summarized as follows:

1) *Original*. In this case, we use the (full) original spectral signatures available in the hyperspectral data as input to the proposed classification system. The dimensionality of the input features used for classification equals n, the number of spectral bands in the original data set.
2) *Reduced*. Here, we apply a dimensionality transformation (such as the MNF or the PCA) to the original input data so that the dimensionality of the input data is reduced and information is packed in the first components resulting after the transformation. In this case, we use a consensus between the HySime and the VD methods to estimate the dimensionality of the hyperspectral data set and then retain the first p components of the data after the dimensional transformation. As a result, the dimensionality of the input features used for classification in this particular case is p.
3) *Unmixing-based*. Finally, in this case we use the three considered unmixing-based feature extraction chains to reduce the dimensionality of the input data. The dimensionality of the input features resulting from chain #1 is p, the number of extracted endmembers, while in the case of chain #2 the dimensionality of the input features is c, the number of different classes available in the labeled training set.

The resulting features are used to train an SVM classifier using three types of kernels: polynomial, Gaussian and sigmoid. Specifically, the SVM was trained with each of these training subsets and then evaluated with the remaining test set. Each experiment was repeated ten times, and the mean accuracy values were reported to guarantee the statistical significance of the results. Kernel parameters were optimized in all experiments by a grid search procedure. Random samples of 5%, 10% and 15% of the pixels were chosen from the known ground-truth of the nine land-cover classes in Fig. 3(b). Then, the three types of input features mentioned at the beginning of this section were constructed for the selected training samples. The dimensionality of the ROSIS data, as estimated by a consensus between the HySime and the VD concept, was $p = 10$.

Table I summarizes the overall classification accuracies obtained by the SVM classifier (using the three considered kernels) after processing the ROSIS Pavia scene. As shown by Table I, the classification accuracies obtained by MNF-based features and the proposed unmixing-based strategy #2 offer the best compromise in terms of classification accuracy across the three considered kernels, with the unmixing-based strategy outperforming most other methods for the sigmoid kernel and the MNF-based strategy outperforming all other approaches for the polynomial and Gaussian kernels. If we analyze in detail the classification results provided for the SVM classifier with Gaussian kernel and trained with 5% of the available samples, reported in Fig. 4, we can see that the MNF-based feature extraction approach can reduce the confusion between classes such as *Meadows* and *Bare_soil* observed when using the PCA-based feature extraction and the original spectral information. However, the confusion between classes such as *Gravel* and *Self-blocking brick*, or between *Asphalt* and *Bitumen* is more effectively resolved by the proposed unmixing-based feature extraction methods, and in particular by strategy #2. As a result, unmixing-based feature extraction can assist in characterizing subtle spectral differences between the classes while maintaining the dimensionality of input features low and retaining the capacity to provide good general characterization of the classes, as indicated by Fig. 4(e). This can also be observed in Fig. 5 which compares the individual class accuracies by the Unmixing Chain #2 and the MNF for

TABLE I

OVERALL CLASSIFICATION ACCURACIES (IN PERCENTAGE) OBTAINED AFTER APPLYING THE CONSIDERED SVM CLASSIFICATION SYSTEM (WITH POLYNOMIAL, GAUSSIAN AND SIGMOID KERNELS) TO $p = 10$ FEATURES EXTRACTED AFTER APPLYING THE UNMIXING CHAIN #1, AND TO $c = 9$ FEATURES EXTRACTED AFTER APPLYING THE UNMIXING CHAIN #2 TO THE ROSIS PAVIA SCENE WITH $n = 115$ SPECTRAL BANDS. THE RESULTS OBTAINED FOR THE ORIGINAL SPECTRAL INFORMATION AND FOR $p = 10$ FEATURES EXTRACTED USING PCA AND MNF ARE ALSO REPORTED.

	Polynomial kernel			**Gaussian kernel**		
Training set size	5%	10%	15%	5%	10%	15%
Original	93.98±0.24	93.98±0.23	94.29±0.11	93.38±0.21	94.40±0.21	94.87±0.21
PCA	89.02±0.31	90.42±0.23	90.86±0.19	89.08±0.24	90.30±0.22	90.75±0.15
MNF	92.27±0.18	93.40±0.21	93.98±0.09	92.40±0.19	93.51±0.23	94.06±0.09
Unmixing chain #1	89.25±0.21	90.21±0.16	90.53±0.23	89.34±0.29	90.43±0.21	90.95±0.26
Unmixing chain #2	90.75±0.51	91.98±0.73	93.05±0.23	90.94±0.49	92.05±0.76	92.95±0.31

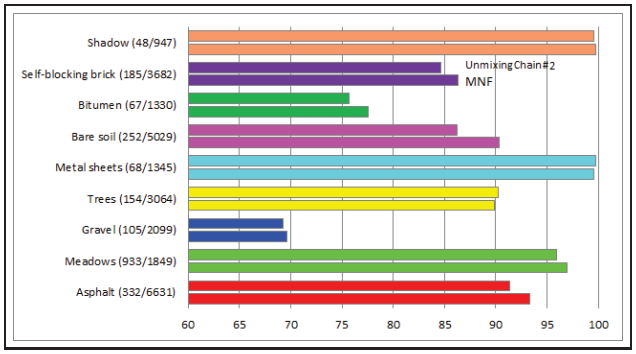


Fig. 5. Class accuracies (MNF vs. Chain #2)

the polynomial kernel, using 5% of the samples for training.

IV. CONCLUSIONS AND FUTURE RESEARCH LINES

In this paper, we have investigated several strategies to use spectral unmixing concepts to extract relevant features from urban scenes for classification purposes. Our experimental results, conducted using ROSIS hyperspectral data, reveal that unmixing-based techniques can provide a suitable alternative to the PCA or MNF transforms for feature extraction prior to classification. Future lines of research should comprise the development of more intelligent approaches to construct the training sets, which in this work are randomly selected from the available labeled samples. Additional scenes (from urban and other areas) should also be used in future developments.

REFERENCES

[1] A. F. H. Goetz, G. Vane, J. E. Solomon, and B. N. Rock, "Imaging spectrometry for earth remote sensing," *Science*, vol. 228, pp. 1147–1153, 1985.

[2] A. Plaza, J. A. Benediktsson, J. Boardman, J. Brazile, L. Bruzzone, G. Camps-Valls, J. Chanussot, M. Fauvel, P. Gamba, J. Gualtieri, M. Marconcini, J. C. Tilton, and G. Trianni, "Recent advances in techniques for hyperspectral image processing," *Remote Sensing of Environment*, vol. 113, pp. 110–122, 2009.

[3] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 44–57, 2002.

[4] J. A. Richards, "Analysis of remotely sensed data: the formative decades and the future," *IEEE Trans. Geoscience and Remote Sensing*, vol. 43, pp. 422–432, 2005.

[5] J. B. Adams, M. O. Smith, and P. E. Johnson, "Spectral mixture modeling: a new analysis of rock and soil types at the viking lander 1 site," *Journal of Geophysical Research*, vol. 91, pp. 8098–8112, 1986.

[6] P. Gamba, F. Dell'Acqua, A. Ferrari, J. A. Palmason, and J. A. Benediktsson, "Exploiting spectral and spatial information in hyperspectral urban data with high resolution," *IEEE Geoscience and Remote Sensing Letters*, vol. 1, pp. 322–326, 2004.

[7] J. B. J. Chanussot and M. Fauvel, "Classification of remote sensing images from urban areas using a fuzzy possibilistic model," *IEEE Geoscience and Remote Sensing Letters*, vol. 3, pp. 40–44, 2006.

[8] D. A. Landgrebe, *Signal theory methods in multispectral remote sensing*. Hoboken, NJ: John Wiley and Sons, 2003.

[9] G. M. Foody and A. Mathur, "Toward intelligent training of supervised image classifications: directing training data acquisition for svm classification," *Remote Sensing of Environment*, vol. 93, pp. 107–117, 2004.

[10] J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Trans. Geoscience and Remote Sensing*, vol. 42, pp. 480–491, 2005.

[11] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. Vila-Frances, and J. Calpe-Maravilla, "Composite kernels for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 3, pp. 93–97, 2006.

[12] L. Bruzzone, M. Chi, and M. Marconcini, "A novel transductive SVM for the semisupervised classification of remote sensing images," *IEEE Trans. Geoscience and Remote Sensing*, vol. 44, pp. 3363–3373, 2006.

[13] A. Plaza, P. Martinez, J. Plaza, and R. Perez, "Dimensionality reduction and classification of hyperspectral image data using sequences of extended morphological transformations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 466–479, 2005.

[14] A. A. Green, M. Berman, P. Switzer, and M. D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, pp. 65–74, 1988.

[15] J. M. Bioucas-Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 8, pp. 2435–2445, 2008.

[16] C.-I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 3, pp. 608–619, 2004.

[17] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 3, pp. 650–663, 2004.

[18] J. C. Harsanyi and C.-I. Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, no. 4, pp. 779–785.

[19] D. Heinz and C.-I. Chang, "Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, pp. 529–545, 2001.

[20] J. Boardman, "Leveraging the high dimensionality of AVIRIS data for improved subpixel target unmixing and rejection of false positives: mixture tuned matched filtering," *Proceedings of the 5th JPL Geoscience Workshop*, pp. 55–56, 1998.