



Unmixing Prior to Supervised Classification of Remotely Sensed Hyperspectral Images

Inmaculada Dopido, Maciel Zortea, Alberto Villa, Antonio Plaza, Paolo Gamba

► To cite this version:

Inmaculada Dopido, Maciel Zortea, Alberto Villa, Antonio Plaza, Paolo Gamba. Unmixing Prior to Supervised Classification of Remotely Sensed Hyperspectral Images. IEEE Geoscience and Remote Sensing Letters, 2011, 8 (4), pp.760-764. 10.1109/LGRS.2011.2109367 . hal-00602477

HAL Id: hal-00602477

<https://hal.science/hal-00602477>

Submitted on 22 Jun 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Unmixing Prior to Supervised Classification of Remotely Sensed Hyperspectral Images

Inmaculada Dórido, Maciel Zortea, Alberto Villa, *Student Member, IEEE*,
Antonio Plaza, *Senior Member, IEEE*, and Paolo Gamba, *Senior Member, IEEE*

Abstract—Supervised classification of 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. For this purpose, several feature extraction methods have been investigated in order to reduce the dimensionality of the data to the right subspace without significant loss of the original information that allows for the separation of classes. In this letter, we explore the use of spectral unmixing for feature extraction prior to supervised classification of hyperspectral data using support vector machines. The proposed feature extraction strategy has been implemented in the form of four different unmixing chains and evaluated using two different scenes collected by National Aeronautics and Space Administration Jet Propulsion Laboratory's Airborne Visible/Infrared Imaging Spectrometer. The experiments suggest competitive results but also show that the definition of the unmixing chains plays an important role in the final classification accuracy. Moreover, differently from most feature extraction techniques available in the literature, the features obtained using linear spectral unmixing are potentially easier to interpret due to their physical meaning.

Index Terms—Endmember extraction, feature extraction, hyperspectral image classification, spectral unmixing, support vector machines (SVMs).

I. INTRODUCTION

IN MANY studies, hyperspectral analysis techniques are divided into full- and mixed-pixel classification techniques [1]–[3], where each pixel vector defines a *spectral signature* or *fingerprnt* that uniquely characterizes the underlying materials at each site in a scene. Full-pixel classification techniques assume that each spectral signature comprises 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 hyperspec-

tral imaging terminology [4]. Let us denote a remotely sensed hyperspectral scene with n bands by $\mathbf{I}$, in which each pixel is represented by a vector $\mathbf{X} = [x_1, x_2, \dots, x_n] \in \mathbb{R}^n$, where $\mathbb{R}$ denotes the set of real numbers in which the pixel's spectral response x_k at sensor channels $k = 1, \dots, n$ is included. Under the linear mixture model assumption, each pixel vector can be modeled using

$$\mathbf{X} = \sum_{z=1}^p \Phi_z \cdot \mathbf{E}_z + \mathbf{n} \quad (1)$$

where $\mathbf{E}_z$ denotes the spectral response of endmember z , Φ_z is a scalar value designating the fractional abundance of the endmember z at the pixel $\mathbf{X}$, p is the total number of endmembers, and $\mathbf{n}$ is a noise vector. Two physical constraints can be imposed into the model described in (1): abundance nonnegativity constraint, i.e., $\Phi_z \geq 0$, and abundance sum-to-one constraint, i.e., $\sum_{z=1}^p \Phi_z = 1$ [5].

Several machine learning techniques have been applied, under the full-pixel assumption, to extract relevant information from hyperspectral data. The good classification performance exhibited by the support vector machine (SVM) [1], [6], [7] using spectral signatures as input features can be improved by applying suitable feature extraction strategies that are able to reduce the dimensionality of the data to a subspace without losing the original information [8]. Techniques used for this purpose include principal component analysis (PCA) [3], minimum noise fraction (MNF) [9], or independent component analysis (ICA) [10]. PCA and MNF maximize the amount of data variance and signal-to-noise ratio (SNR), respectively, yielding a transformed data set in a new uncorrelated coordinate system, while ICA tries to find components as statistically independent as possible. However, all these methods maximize the information contained in the first transformed components, relegating variations of less significant size to low-order components. If such low-order components are not preserved, small classes may be affected. The inclusion of spatial features such as morphological profiles can be used to address this issue [8], [11], [12].

In this letter, we explore an alternative strategy focused on the use of spectral unmixing for feature extraction prior to classification. Previous efforts in this direction were presented in [13] and [14], but the analysis of whether spectral unmixing can replace standard feature extraction transformations remains an unexplored topic. Although classification techniques often neglect the impact of mixed pixels in the provision of a set of final class labels, widely used benchmark data sets in the literature—e.g., the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) Indian Pines scene—are known to be

Manuscript received November 20, 2010; revised January 17, 2011; accepted January 19, 2011. This work was supported in part by the European Community's Marie Curie Research Training Networks Programme under Grant MRTN-CT-2006-035927 and in part by the Spanish Ministry of Science and Innovation under Grant AYA2008-05965-C04-02.

I. Dórido, M. Zortea, A. Villa, and A. Plaza are with the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, University of Extremadura, 10071 Cáceres, Spain (e-mail: inmaculadadopido@unex.es; maciel.zortea@hyperinet.eu; alberto.villa@hyperinet.eu; aplaza@unex.es).

P. Gamba is with the Telecommunications and Remote Sensing Laboratory, University of Pavia, 27100 Pavia, Italy (e-mail: paolo.gamba@unipv.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/LGRS.2011.2109367

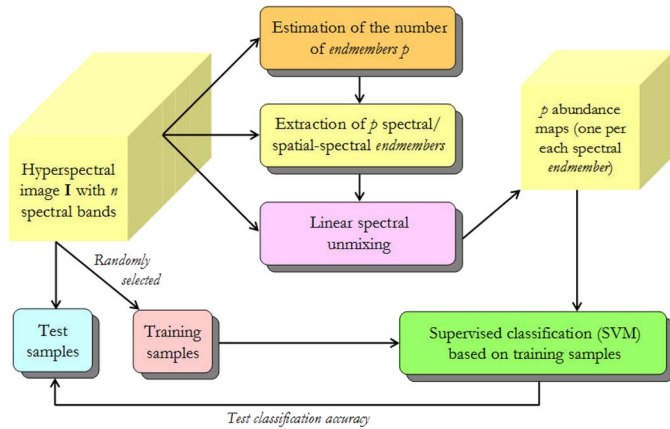


Fig. 1. Unmixing-based feature extraction chains #1 (spectral endmembers) and #2 (spatial-spectral endmembers).

dominated by mixed pixels, even if the associated ground-truth information is only available in full-pixel form. Hence, the use of spectral unmixing presents distinctive features with regard to other approaches such as PCA, MNF, or ICA. First, it provides additional information for classification in hyperspectral analysis scenarios with moderate spatial resolution, since the subpixel composition of training samples can be used as part of the learning process of the classifier. Second, the components estimated by spectral unmixing can be physically explained as the abundances of spectral endmembers. Third, spectral unmixing does not penalize classes which are not relevant in terms of variance or SNR. Here, we design different unmixing processing chains with the goal of addressing three specific research questions.

- 1) Is spectral unmixing a feasible strategy for feature extraction prior to classification?
- 2) Does the inclusion of spatial information at the endmember extraction stage lead to better classification results?
- 3) Is it really necessary to estimate pure spectral endmembers for classification purposes?

We have structured the remainder of this letter as follows. Section II describes the considered spectral unmixing chains. Section III presents different experiments specifically designed to address the aforementioned research questions and provide a comparison between the proposed unmixing-based strategy and other feature extraction approaches in the literature. Section IV concludes with some remarks and future research avenues.

II. UNMIXING-BASED FEATURE EXTRACTION

A. Unmixing Chain #1

In this section, 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 I . For this purpose, we use in this letter 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, we apply an automatic algorithm to extract a set of endmembers from the original hyperspectral image [17]. Here, we use an orthogonal subspace projection technique [18] which has

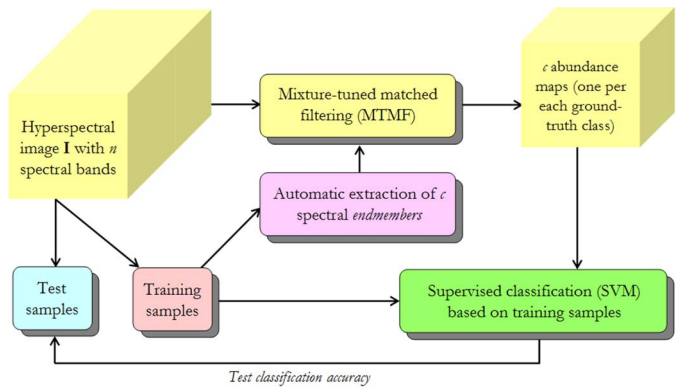


Fig. 2. Unmixing-based feature extraction chain #3 (chain #4 replaces end-member extraction with averaging of the signatures associated to each labeled class in the training set).

been shown in previous work to provide a very good tradeoff between the signature purity of the extracted endmembers and the computational time to obtain them. Preliminary experiments conducted with other endmember extraction techniques, such as vertex component analysis [19] and N-FINDR [20], have shown very similar results in terms of classification accuracy. Finally, linear spectral unmixing (either unconstrained or constrained) can be used to estimate the abundance of each endmember in each pixel of the scene, providing a set of p abundance maps. Then, standard SVM classification is performed on the stack of abundance fractions using randomly selected training samples.

B. Unmixing Chain #2

In this section, we introduce a variation of the unmixing-based feature extraction chain which includes spatial preprocessing prior to endmember extraction in order to guide the endmember searching process to those areas which are more spatially homogeneous. This approach is shown in Fig. 1. The spatial preprocessing strategy adopted in this letter is described in detail in [21]. As in the previous chain, the features resulting from the proposed (spatially enhanced) unmixing process are used to train an SVM classifier with a few randomly selected labeled samples. The classifier is then tested using the remaining labeled samples.

C. Unmixing Chain #3

Our main motivation for introducing a third 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. Fig. 2 shows a new chain in which the endmembers are extracted from the set of available (labeled) training samples instead of from the original image. This chain introduces two important variations: 1) As a simplification to the challenging estimation problem, the number of endmembers to be extracted is set as the total number of different classes c in the training set, and 2) the endmember searching process is conducted only on the training set, which reduces computational complexity. However, the number of endmembers in the original image p is probably different than c , the number of labeled classes. 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

TABLE I

CLASSIFICATION ACCURACIES (PERCENTAGE) AND STANDARD DEVIATION OBTAINED AFTER APPLYING THE CONSIDERED SVM CLASSIFICATION SYSTEM (WITH GAUSSIAN AND POLYNOMIAL KERNELS) TO THREE DIFFERENT TYPES OF FEATURES (ORIGINAL, REDUCED, AND UNMIXING BASED) EXTRACTED FROM THE AVIRIS INDIAN PINES AND KENNEDY SPACE CENTER SCENES (TEN RANDOMLY CHOSEN TRAINING SETS)

Image		AVIRIS Indian Pines					
Type of feature	# of features	Polynomial kernel			Gaussian kernel		
		5%	10%	15%	5%	10%	15%
Original spectral information	202	75.23±1.23	81.55±0.86	83.58±0.78	75.78±1.06	82.11±0.43	84.49±0.53
PCA	18	77.07±1.46	81.66±0.88	83.11±0.52	77.12±1.29	81.68±0.61	82.96±0.58
MNF	18	82.97±1.93	87.41±0.31	88.38±0.57	84.04±0.75	87.66±0.52	89.34±0.43
ICA	18	76.63±1.27	81.00±0.71	82.94±0.36	76.92±0.72	81.27±0.61	82.95±0.71
Chain #1	18	74.56±1.04	79.20±1.12	80.97±0.50	74.65±0.99	79.45±0.40	80.91±0.39
Chain #2	18	71.93±0.96	77.58±0.92	79.31±0.33	72.31±0.98	77.36±0.72	79.17±0.28
Chain #3	16	81.32±0.84	85.56±0.84	86.83±0.55	81.78±0.62	86.12±0.66	87.40±0.82
Chain #4	16	82.36±1.09	86.87±0.59	87.97±0.57	82.72±1.04	87.59±0.57	88.92±0.80
Image		AVIRIS Kennedy Space Center					
Type of feature	# of features	Polynomial kernel			Gaussian kernel		
		1%	3%	5%	1%	3%	5%
Original spectral information	176	70.97±3.32	82.53±1.63	85.71±1.40	72.26±2.42	82.91±1.38	85.50±1.35
PCA	15	73.52±3.69	83.26±1.26	86.11±1.16	74.66±2.94	82.54±1.70	86.28±1.46
MNF	15	77.01±3.77	86.85±2.19	89.59±1.89	77.94±3.48	87.43±2.11	90.01±1.52
ICA	15	70.09±2.91	80.28±1.73	84.59±1.50	70.39±1.58	80.79±1.60	84.58±1.58
Chain #1	15	69.41±2.64	78.62±1.58	82.84±1.17	69.02±5.40	79.08±1.46	83.53±1.25
Chain #2	15	67.91±3.98	78.61±3.56	84.26±1.41	68.56±4.70	83.86±1.89	83.86±1.22
Chain #3	12	74.28±3.23	85.37±1.30	87.88±1.57	75.02±4.13	84.92±1.97	88.47±1.38
Chain #4	12	76.10±2.49	86.38±1.40	87.84±1.28	77.53±2.58	86.57±0.97	87.72±1.13

technique for this purpose is mixture-tuned matched filtering (MTMF) [22], also known as constrained energy minimization [23], 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. From spectral mixture modeling, it inherits the leverage arising from the mixed pixel model and the constraints on feasibility.

D. Unmixing Chain #4

The fourth unmixing chain tested in our experiments represents a slight variation of unmixing chain #3 in which the spectral signatures used for unmixing purposes are not obtained via endmember extraction but through averaging of the spectral signatures associated to each labeled class in the training set. To keep the number of estimated components low, only one component is allowed for each class. This averaging strategy produces c signatures, each representative of a labeled class, which are then used to partially unmix the original hyperspectral scene using MTMF.

III. EXPERIMENTAL RESULTS

A. Hyperspectral Data

The first data set used in our experiments was collected by the AVIRIS sensor over the Indian Pines region. The scene¹ comprises 145 lines by 145 samples and 220 spectral channels with the wavelength range from 0.4 to 2.5 μm , nominal spectral resolution of 10 nm, and spatial resolution of 20 m by pixel. After removing noisy and water absorption bands, 202 channels were left. The ground truth contains 16 land cover classes. The number of pixels in the smallest class is 20, while the number of pixels in the largest class is 2468.

The second data set was collected by the AVIRIS sensor over the Kennedy Space Center,² Florida, in March 1996. The portion of this scene used in our experiments has dimensions

of 292×383 pixels. After removing water absorption and low-SNR bands, 176 bands were used for the analysis. The spatial resolution is 20 m by pixel. Twelve ground-truth classes were available, where the number of pixels in the smallest class is 134 while the number of pixels in the largest class is 761.

B. Experiments

1) *Experiment 1. Use of Unmixing as a Feature Extraction Strategy:* In this experiment, we use the AVIRIS Indian Pines and Kennedy Space Center data sets to analyze the impact of imposing nonnegativity and sum-to-one constraints in abundance estimation prior to classification. For the AVIRIS Indian Pines image, we construct ten small training sets by randomly selecting 5%, 10%, and 15% of the ground-truth pixels. For the AVIRIS Kennedy Space Center, since the size of the smaller classes is bigger, we decided to reduce the training sets even more and selected 1%, 3%, and 5% of the available ground-truth pixels. Then, the three considered types of input features (original, reduced, and unmixing based) are built for the selected training samples and used to train an SVM classifier in which two types of kernels, i.e., polynomial and Gaussian, are used. 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 and standard deviation accuracy values were reported. Kernel parameters were optimized by a grid search procedure, and the optimal parameters were selected using tenfold cross validation. The LIBSVM library³ was used for the experiments.

Table I summarizes the overall classification accuracies obtained after applying the considered SVM classification system (with polynomial and Gaussian kernels) to the features extracted after applying unmixing chain #1 (see Fig. 1) to the AVIRIS scenes. The dimensionality of the input data, as estimated by a consensus between the HySime and the VD concept, was $p = 18$ for the Indian Pines scene and $p = 15$ for the Kennedy Space Center scene. Chain #1 was implemented using two different linear spectral unmixing algorithms [5]: unconstrained and fully constrained; due to better accuracy and

¹ Available online: <http://dynamo.ecn.purdue.edu/~biehl/MultiSpec>

² Available online: <http://www.csr.utexas.edu/hyperspectral/data/KSC/>

³ <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>

TABLE II
STATISTICAL DIFFERENCES EVALUATED USING MCNEMAR'S TEST (POLYNOMIAL KERNEL)

AVIRIS Indian Pines	5%		10%		15%	
	Chain #3	Chain #4	Chain #3	Chain #4	Chain #3	Chain #4
PCA	-9.52 (0/0/10)	-11.88 (0/0/10)	-9.24 (0/0/10)	-12.40 (0/0/10)	-8.86 (0/0/10)	-11.40 (0/0/10)
MNF	5.22 (8/1/1)	2.05 (6/2/2)	6.13 (10/0/0)	1.72 (5/4/1)	5.23 (10/0/0)	1.35 (3/7/0)
ICA	-10.45 (0/0/10)	-12.85 (0/0/10)	-17.37 (0/0/10)	-20.26 (0/0/10)	-9.28 (0/0/10)	-11.87 (0/0/10)
AVIRIS Kennedy Space Center	1%		3%		5%	
	Chain #3	Chain #4	Chain #3	Chain #4	Chain #3	Chain #4
PCA	-2.09 (1/2/7)	-5.73 (0/2/8)	-2.61 (1/2/7)	-5.14 (0/0/10)	-3.78 (0/1/9)	-2.56 (0/4/6)
MNF	3.08 (6/3/1)	-0.58 (3/4/3)	3.48 (6/2/2)	0.78 (5/1/4)	2.42 (6/2/2)	4.04 (7/3/0)
ICA	-5.32 (0/0/10)	-7.90 (0/0/10)	-7.14 (0/0/10)	-8.82 (0/0/10)	-5.08 (0/1/9)	-5.29 (0/0/10)

faster computation, only results for the unconstrained case are presented. The results after applying the classification system to the original spectral features and to those extracted using PCA, MNF, and ICA are also reported.

As shown in Table I, the classification accuracy is correlated with the training set size (the larger the training set, the higher the classification accuracy). The good generalization ability exhibited by SVMs is demonstrated by the classification results reported for the original spectral information, even with very limited training sets. The fact that MNF is more effective than PCA and ICA for feature extraction purposes is also remarkable, since the MNF has been more widely used in the context of spectral unmixing rather than classification. Table I also reveals that the use of unmixing chain #1 as a feature extraction strategy cannot improve the classification results provided by PCA, MNF, ICA, or the original spectral information. This is because endmember extraction is generally sensitive to outliers and anomalies; hence, a strategy for directing the endmember searching process to spatially homogeneous areas could improve the final classification results.

2) *Experiment 2. Impact of Including Spatial Information at the Endmember Extraction Stage:* In this experiment, we apply unmixing chain #2 for feature extraction prior to classification. As shown in Table I, spatial preprocessing prior to endmember extraction cannot lead to improved classification results with regard to chain #1 and the original spectral information. This is due to the spectral similarity of the most spatially representative classes in our considered scenes. For instance, in the AVIRIS Indian Pines scene, the corns and soybeans were very early in their growth cycle at the time of data collection, which resulted in low coverage of the soil ($\approx 5\%$) [24]. Given this low canopy ground cover, the variation in spectral response among different classes is very low, and spatial information cannot significantly increase discrimination between different classes. In order to address this issue, a possible solution is to conduct the endmember extraction process in a supervised fashion, taking advantage of the information contained in the available labeled samples in order to guarantee that a highly representative endmember is selected per class.

3) *Experiment 3. Impact of Endmember Purity on the Final Classification Results:* In a supervised endmember extraction framework, our first experiment is based on applying unmixing chain #3 to select endmembers only from the available training samples. Apart from reducing computational complexity (which, in this case, involves a search for c endmembers in the pixels belonging to the training set), Table I reveals that this strategy improves the classification results reported for chains #1 and #2. However, in order to make sure that only one endmember per labeled class is used for unmixing purposes, we also apply unmixing chain #4 in which spectral averaging of the available training pixels in each class is conducted in

order to produce a final set of c spectral signatures. Although averaging of endmembers can lead to degradation of spectral purity, it can also reduce the effects of noise and/or average out the subtle spectral variability of a given class, thus obtaining a more representative endmember for the class as a whole. This is illustrated by the classification results for unmixing chain #4 in Table I, which outperform those reported for most other tested methods except the MNF. This indicates that, in a supervised unmixing scenario, the use of spectrally pure signatures is not as important as the choice of signatures which are representative of the available training samples.

Table II shows the statistical differences (average value of ten comparisons) between traditional dimensionality reduction methods and unmixing chains #3 and #4, computed using McNemar's test [25] for the case of the polynomial kernel. The differences are statistically significant at a confidence level of 95% if $|Z| > 1.96$. For each couple of compared feature extraction chains, we report also how many times each chain wins/ties/loses after comparing the thematic maps obtained using the same training set. If the value of Z reported for each entry in Table II is positive and larger than 1.96, the first compared chain wins. By convention, the comparison is always performed with the first chain in a line in Table II and the second chain in a column in Table II. It can be noticed that unmixing chains #3 and #4 always perform significantly better than PCA and ICA. MNF performs better than chain #3, while the differences with chain #4 are in general not significant.

To conclude this section, Fig. 3 shows the best classification results (out of ten runs) obtained after applying the SVM—trained with 10% of the available training samples—to each feature extraction strategy considered for the AVIRIS Indian Pines scene. As shown in Fig. 3, both the MNF in Fig. 3(d) and chain #4 in Fig. 3(h) provide the best classification scores, with less confusion in heavily mixed classes.

IV. CONCLUSION AND FUTURE RESEARCH LINES

In this letter, we have investigated several strategies to extract relevant features from hyperspectral scenes prior to classification. For classification scenarios using SVMs trained with relatively small subsets of labeled samples, our experimental results reveal that the MNF greatly improves accuracies when compared to the more well-known PCA and ICA transformations, used as an unsupervised feature reduction tool prior to classification. Due to the reduced dimensionality, classification using both the MNF and PCA subspaces generally improved the overall accuracy when compared to using all the original pixel's spectral curves. The results indicate that the proposed unmixing-based feature extraction chains can provide an alternative strategy to the PCA or MNF by incorporating information about the (possibly) mixed nature of the training

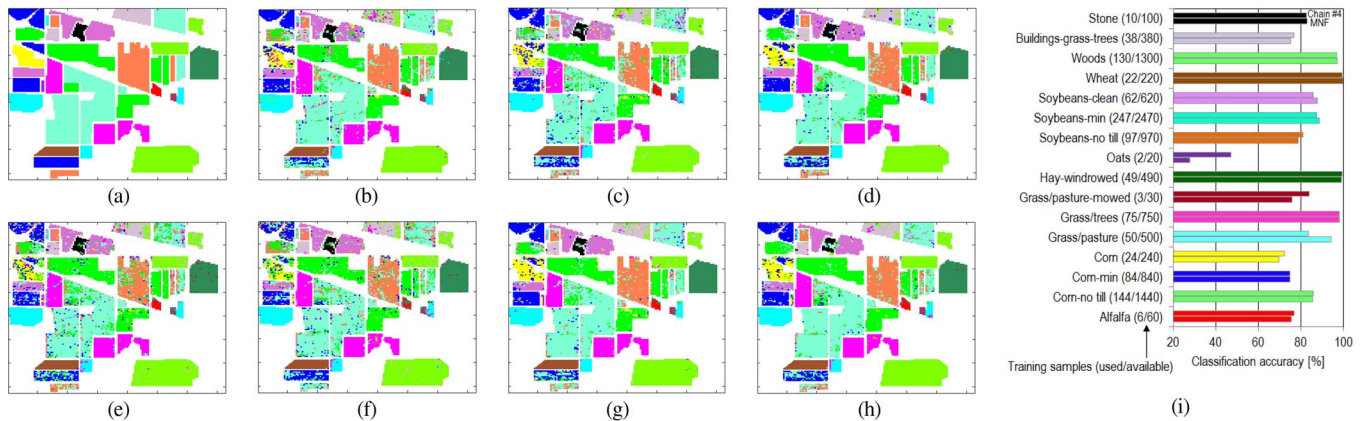


Fig. 3. Best classification results for AVIRIS Indian Pines (using SVM classifier with Gaussian kernel, trained with 10% of the available samples per class). (a) Ground-truth. (b) Original image (84.27%). (c) PCA (83.33%). (d) MNF (89.41%). (e) Chain #1 (81.29%). (f) Chain #2 (79.64%). (g) Chain #3 (87.99%). (h) Chain #4 (89.26%). (i) Class accuracies (MNF versus Chain #4).

samples during the learning stage, with the potential advantage of improved interpretability of features due to the physical nature of the extracted abundance maps. Although final classification accuracies are likely to be dependent on the particular data set considered, the chains tested suggest higher accuracies with respect to traditional methods, such as PCA and ICA, and comparable accuracies related to MNF.

Further research is needed to define an optimality criterion to design unmixing chains as a feature reduction tool for classification purposes. A start point might be chain #4 which indicates that, in the context of a supervised unmixing scenario, the use of spectrally pure signatures is not as important as the choice of signatures which are highly representative of the available training samples.

ACKNOWLEDGMENT

The authors would like to thank D. Landgrebe and M. Crawford for sharing the hyperspectral data and the two reviewers for their comments.

REFERENCES

- [1] 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 Sens. Environ.*, vol. 113, pp. 110–122, Sep. 2009.
- [2] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.
- [3] J. A. Richards, "Analysis of remotely sensed data: The formative decades and the future," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 422–432, Mar. 2005.
- [4] 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," *J. Geophys. Res.*, vol. 91, no. B8, pp. 8098–8112, 1986.
- [5] D. Heinz and C.-I. Chang, "Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, pp. 529–545, Mar. 2001.
- [6] G. Camps-Valls, L. Gomez-Chova, J. Munoz-Mari, J. Vila-Frances, and J. Calpe-Maravilla, "Composite kernels for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 93–97, Jan. 2006.
- [7] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote-sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778–1790, Aug. 2004.
- [8] A. Plaza, P. Martinez, J. Plaza, and R. Perez, "Dimensionality reduction and classification of hyperspectral image data using sequences of extended morphological transformations," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 466–479, Mar. 2005.
- [9] 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 Trans. Geosci. Remote Sens.*, vol. 26, no. 1, pp. 65–74, Jan. 1988.
- [10] P. Comon, "Independent component analysis, a new concept?" *Signal Process.*, vol. 36, no. 3, pp. 287–314, Apr. 1994.
- [11] J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 480–491, Mar. 2005.
- [12] 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 Geosci. Remote Sens. Lett.*, vol. 1, no. 4, pp. 322–326, Oct. 2004.
- [13] B. Luo and J. Chanussot, "Hyperspectral image classification based on spectral and geometrical features," in *Proc. IEEE Int. Workshop Mach. Learn. Signal Process.*, 2009, pp. 1–6.
- [14] B. Luo and J. Chanussot, "Unsupervised classification of hyperspectral images by using linear unmixing algorithm," in *Proc. IEEE Int. Conf. Image Process.*, 2009, pp. 2877–2880.
- [15] J. M. Bioucas-Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 8, pp. 2435–2445, Aug. 2008.
- [16] C.-I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, pp. 608–619, Mar. 2004.
- [17] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, pp. 650–663, Mar. 2004.
- [18] J. C. Harsanyi and C.-I. Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 4, pp. 779–785, Jul. 1994.
- [19] J. M. P. Nascimento and J. M. Bioucas-Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, Apr. 2005.
- [20] M. E. Winter, "N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data," in *Proc. SPIE Image Spectrometry V*, 2003, vol. 3753, pp. 266–277.
- [21] M. Zortea and A. Plaza, "Spatial preprocessing for endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2679–2693, Aug. 2009.
- [22] J. Boardman, "Leveraging the high dimensionality of AVIRIS data for improved subpixel target unmixing and rejection of false positives: Mixture tuned matched filtering," in *Proc. 5th JPL Geosci. Workshop*, 1998, pp. 55–56.
- [23] Q. Du, H. Ren, and C.-I. Chang, "A comparative study for orthogonal subspace projection and constrained energy minimization," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 6, pp. 1525–1529, Jun. 2003.
- [24] D. A. Landgrebe, *Signal Theory Methods in Multispectral Remote Sensing*. Hoboken, NJ: Wiley, 2003.
- [25] G. Foody, "Thematic map comparison: Evaluating the statistical significance of differences in classification accuracy," *Photogramm. Eng. Remote Sens.*, vol. 70, no. 5, pp. 627–633, May 2004.