# Unsupervised Clustering and Spectral Unmixing for Feature Extraction Prior to Supervised Classification of Hyperspectral Images

Inmaculada Dópido[a], Alberto Villa[a] and Antonio Plaza[a]

[a]Hyperspectral Computing Laboratory
Department of Technology of Computers and Communications
University of Extremadura, Avda. de la Universidad s/n
10071 Cáceres, Spain

## ABSTRACT

Classification and spectral unmixing are two very important tasks for hyperspectral data exploitation. Although many studies exist in both areas, the combined use of both approaches has not been widely explored in the literature. Since hyperspectral images are generally dominated by mixed pixels, spectral unmixing can particularly provide a useful source of information for classification purposes. In previous work, we have demonstrated that spectral unmixing can be used as an effective approach for feature extraction prior to supervised classification of hyperspectral data using support vector machines (SVMs). Unmixing-based features do not dramatically improve classification accuracies with regards to features provided by classic techniques such as the minimum noise fraction (MNF), but they can provide a better characterization of small classes. Also, these features are potentially easier to interpret due to their physical meaning (in spectral unmixing, the features represent the abundances of real materials present in the scene). In this paper, we develop a new strategy for feature extraction prior to supervised classification of hyperspectral images. The proposed method first performs unsupervised multidimensional clustering on the original hyperspectral image to implicitly include spatial information in the process. The cluster centres are then used as representative spectral signatures for a subsequent (partial) unmixing process, and the resulting features are used as inputs to a standard (supervised) classification process. The proposed strategy is compared to other classic and unmixing feature extraction methods presented in the literature. Our experiments, conducted with several reference hyperspectral images widely used for classification purposes, reveal the effectiveness of the proposed approach.

**Keywords:** Hyperspectral image analysis, classification, spectral unmixing, partial unmixing, unsupervised clustering.

## 1. 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] The concept of hyperspectral imaging originated at NASA's Jet Propulsion Laboratory in California with the development of the Airborne Visible Infra-Red Imaging Spectrometer (AVIRIS), able to cover the wavelength region from 0.4 to 2.5 $\mu$m using more than two hundred spectral channels, at nominal spectral resolution of 10 nm.[2] As a result, each pixel vector collected by a hyperspectral instrument can be seen as a *spectral signature* or *fingerprint* of the underlying materials within the pixel.[3]

The special characteristics of hyperspectral datasets pose different processing problems,[4] which must be necessarily tackled under specific mathematical formalisms, such as classification, segmentation or spectral mixture analysis, of which linear spectral unmixing has been one of the most successful approaches.[5] In many studies, techniques are divided into full-pixel and mixed-pixel classification techniques,[6,7] where each pixel vector defines a *spectral signature* or *fingerprint* that uniquely characterizes the underlying materials at each site in a scene.

Send correspondence to Antonio J. Plaza:
E-mail: aplaza@unex.es; Telephone: +34 927 257000 (Ext. 51662); URL: http://www.umbc.edu/rssipl/people/aplaza

Satellite Data Compression, Communications, and Processing VII, edited by Bormin Huang, Antonio J. Plaza, Carole Thiebaut, Proc. of SPIE Vol. 8157, 81570M · © 2011 SPIE · CCC code: 0277-786X/11/$18 · doi: 10.1117/12.892469

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.[8] Mixed pixels can also result when distinct materials are combined into a homogeneous or intimate mixture, which occurs independently of the spatial resolution of the sensor. To address these issues, several endmember extraction[9] and spectral unmixing approaches,[6,10–12] among many others, have been developed in the literature under the assumption that each pixel vector measures the response of multiple underlying materials.

It is important to emphasize that there have been only a few efforts in the literature towards the integration of full-pixel and mixed-pixel techniques for classification of hyperspectral data.[13] Previous efforts towards performing unmixing-based feature extraction for improving unsupervised classification were presented in[14,15] but, to the best of our knowledge, the topic of whether spectral unmixing can be used as feature extraction for improving supervised classification of hyperspectral data via machine learning techniques is a novel contribution. In particular, the good classification performance demonstrated by successful machine learning techniques such as the support vector machine (SVM),[4,16,17] using spectral signatures as input features,[18] can still be improved by the incorporation of intelligent feature extraction strategies which can effectively deal with mixed pixels.

Recently, we have investigated this issue by developing a new set of feature extraction techniques based on spectral unmixing concepts.[19–22] These techniques are intended to take advantage of spectral unmixing models[6] in the characterization of training samples, thus including additional information about sub-pixel composition at the feature extraction stage that can be then exploited at the classification stage. An advantage of unmixing-based techniques over statistical transformations commonly used for feature extraction purposes such as principal component analysis (PCA),[23] the minimum noise fraction (MNF),[24] or independent component analysis (ICA),[25] is the fact that the features derived by spectral unmixing are physically meaningful since they can be interpreted as the abundance of spectrally pure endmembers. The study in[19] also suggested that partial unmixing[26,27] could be an effective solution to deal with the likely fact that not all pure spectral constituents in the scene (needed for spectral unmixing purposes) are known *a priori*. In this case, a more exhaustive investigation of partial unmixing (particularly in combination with the use of spatial information, which may be helpful to better model the image features) is needed.

In this paper, we develop a new feature extraction technique exploiting sub-pixel information is proposed. This approach includes spatial information implicitly using unsupervised clustering in order to define spatially homogeneous regions prior to a partial spectral unmixing stage. The remainder of the paper is organized as follows. Section 2 describes the proposed method. Section 3 first describes the hyperspectral scene used in experiments, collected by AVIRIS over the region of Indian Pines, Indiana, and then provides an experimental comparison of the proposed feature extraction chain with regards to other classic approaches. Section 4 concludes the paper with some remarks and hints at plausible future research lines.

## 2. PROPOSED FEATURE EXTRACTION METHOD

This section is organized as follows. In subsection 2.1 we describe some general concepts about linear spectral unmixing with particular attention to partial unmixing problems that will be addressed in this work. Subsection 2.2 describes a new unsupervised feature extraction strategy based on spectral unmixing concepts.

### 2.1 Linear Spectral Unmixing and Partial Unmixing

Let us denote a remotely sensed hyperspectral scene with $n$ bands by $\mathbf{I}$, in which a given pixel is represented by a vector $\mathbf{x} = [x_1, x_2, \cdots, x_n] \in \Re^n$, where $\Re$ denotes the set of real numbers in which the pixel's spectral response $x_k$ at sensor channels $k = 1, \ldots, n$ is included. Under the linear mixture model assumption, each pixel vector in the original scene can be modeled using the following expression:

$$\mathbf{x} \approx \sum_{z=1}^{p} \Phi_z \cdot \mathbf{e}_z + \mathbf{n}, \qquad (1)$$

where $\mathbf{e}_z$ denotes the spectral response of endmember $z$, $\Phi_z$ 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. An unconstrained solution to Eq. (1) is simply given by the following expression:[28]

$$\hat{\Phi}_{\mathrm{UC}} = (\mathbf{E}^T\mathbf{E})^{-1}\mathbf{E}^T\mathbf{x}, \tag{2}$$

where $\mathbf{E} = \{\mathbf{e}_z\}_{z=1}^{p}$ is a full set of endmembers. Two physical constrains are generally imposed into the model described in Eq. (1), these are the abundance non-negativity constraint (ANC), i.e., $\Phi_z \geq 0$, and the abundance sum-to-one constraint (ASC), i.e., $\sum_{z=1}^{p} \Phi_z = 1$.[11] As indicated in,[11] a fully constrained (i.e. ASC-constrained and ANC-constrained) estimate can be obtained in practice. Such fully constrained linear spectral unmixing estimate is generally referred to in the literature by the acronym FCLSU. In order for such estimate to be meaningful, it is required that the spectral signatures of all endmembers, i.e., $\mathbf{E} = \{\mathbf{e}_z\}_{z=1}^{p}$, are available *a priori*, which is not always possible in practice. In such case, partial unmixing has emerged as a suitable alternative to solve the linear spectral unmixing problem.

The most widely used partial unmixing technique is mixture-tuned matched filtering (MTMF)[27] –also known in the literature as constrained energy minimization (CEM)[26,28]– which combines the best parts of the linear spectral unmixing model and the statistical matched filter model while avoiding some drawbacks of each parent method. 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 including the ASC and ANC requirements. It is essentially a target detection algorithm designed to identify the presence (or absence) of a specified material by producing a score of 1 for pixels wholly covered by the material of interest, while keeping the average score over an image as small as possible. It uses just one endmember spectrum (that of the target of interest) and therefore behaves as a partial unmixing method that suppresses background noise and estimates the sub-pixel abundance of a single endmember material without assuming the presence of all endmembers in the scene, as opposed to FCLSU. If we assume that $\mathbf{e}_z$ is the endmember to be characterized, MTMF estimates the abundance fraction $\Phi$ of $\mathbf{e}_z$ in a specific pixel vector $\mathbf{x}$ of the scene as follows:

$$\hat{\Phi}_{\mathrm{MTMF}} = ((\mathbf{e}_z^T\mathbf{R}^{-1}\mathbf{e}_z)^{-1}\mathbf{R}^{-1}\mathbf{e}_z)^T\mathbf{x}, \tag{3}$$

where $\mathbf{R}^{-1}$ is the inverse of the sample correlation matrix.[28]

### 2.2 Unmixing-Based Feature Extraction

A new spectral unmixing-based feature extraction technique which implicitly includes the spatial information is summarized in Fig. 1. First, we apply the $k$-means clustering algorithm[29] to the original hyperspectral image. Its goal is to determine a set of $c$ points, called centers, so as to minimize the mean squared distance from each pixel vector to its nearest center.[30] The algorithm is based on the observation that the optimal placement of a center is at the centroid of the associated cluster. It starts with a random initial placement. At each stage, the algorithm moves every center point to the centroid of the set of pixel vectors for which the center is a nearest neighbor according to the spectral angle (SA),[6] and then updates the neighborhood by recomputing the SA from each pixel vector to its nearest center. These steps are repeated until the algorithm converges to a point that is a minimum for the distortion.[29] The output of $k$-means is a set of spectral clusters, each made up of one or more spatially connected regions.

The centroids of each cluster resulting from the application of $k$-means are now used as the spectral endmembers for a subsequent partial unmixing process. Our main motivation for using a partial unmixing technique at this point is the fact that the estimation of the number of endmembers in the original image is a very challenging issue. It is possible that the actual number of endmembers in the original image, $p$, may actually be larger than the number of clusters derived by $k$-means. In this case, in order to unmix the original image we need to address a situation in which not all endmembers may be available *a priori*. It has been shown in previous work that the FCLSU technique does not provide accurate results in this scenario.[19] In turn, it is also possible that $p \leq c$. In this case, partial unmixing has shown great success[27] in abundance estimation. Following this line of reasoning, we resort to the MTMF partial unmixing technique in this work. As shown by Fig. 1, the features resulting from

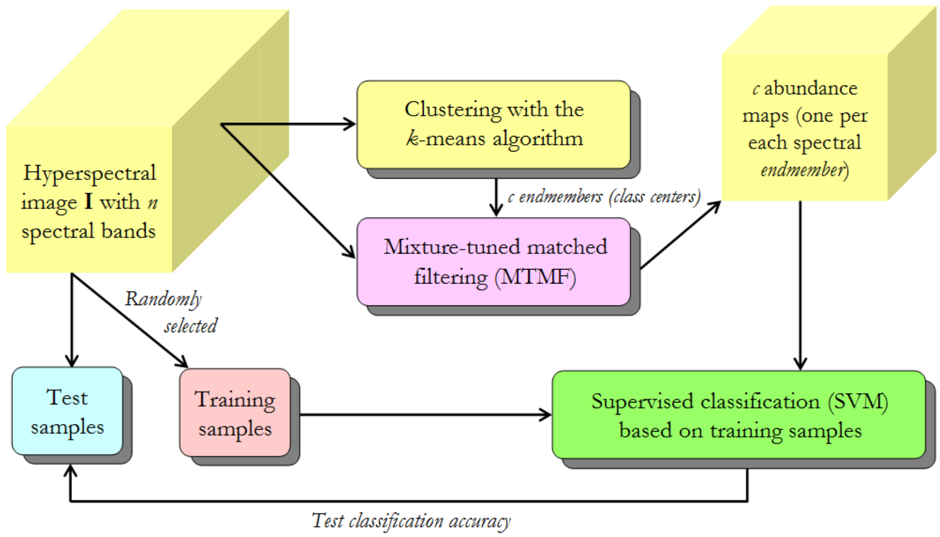



Figure 1. Block diagram illustrating an unsupervised clustering followed by MTMF (CMTMF) technique for unmixing-based feature extraction.

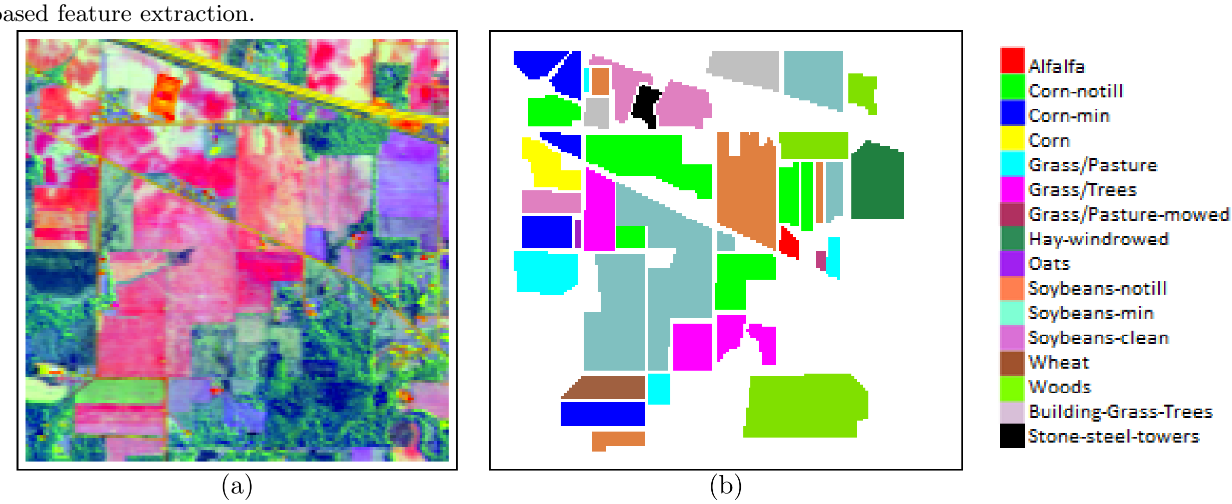



Figure 2. (a) False color composition of the AVIRIS Indian Pines scene. (b) Ground truth-map containing 15 mutually exclusive land-cover classes (right).

the proposed unmixing-based technique, referred to hereinafter as unsupervised clustering followed by MTMF (CMTMF), are used to train an SVM classifier with a few randomly selected labeled samples. The classifier is then tested using the remaining labeled samples.

## 3. EXPERIMENTAL RESULTS

In this section we evaluate our proposed unmixing-based feature extraction technique with regards to other traditional approaches. Before describing the results obtained in experimental validation, we briefly describe the hyperspectral data used in our experiments.

### 3.1 Hyperspectral Data Set

The hyperspectral data set used in our experiments was collected by the AVIRIS sensor over the Indian Pines region in Northwestern Indiana, in 1992. This scene, with a size of 145 lines by 145 samples, was acquired over a mixed agricultural/forest area, early in the growing season. The scene comprises 202 spectral channels in the wavelength range from 0.4 to 2.5 $\mu$m, nominal spectral resolution of 10 nm, moderate spatial resolution of 20 meters by pixel, and 16-bit radiometric resolution. For illustrative purposes, Fig. 2(a) shows a false color

Table 1. Overall classification accuracy (in percentage) obtained by different features extracted from the AVIRIS Indian Pines scene. Only the best case is reported for each considered feature extraction technique (with the optimal number of features given in the parentheses).

| Feature extraction | 5% Training | 15% Training |
|---|---|---|
| None | 75.78% (202) | 84.49% (202) |
| PCA | 77.25% (20) | 83.86% (20) |
| ICA | 76.84% (20) | 83.52% (20) |
| MNF | 86.67% (10) | 91.35% (10) |
| CMTMF | 87.18% (30) | 91.61% (25) |

rendition of the AVIRIS Indian Pines scene while Fig. 2(b) shows the available ground-truth map, displayed in the form of a class assignment for each labeled pixel with 15 mutually-exclusive ground-truth classes. These data, including ground-truth information, are available online*, a fact which has made this scene a widely used benchmark for testing the accuracy of hyperspectral image classification algorithms.

### 3.2 Experimental Validation

Different types of input features were extracted from the AVIRIS Indian Pines hyperspectral image prior to classification. In addition to the considered feature extraction techniques, we also use the (full) original spectral information available in the hyperspectral data as input to the proposed classification system. In this case, the dimensionality of the input features used for classification equals $n$, the number of spectral bands in the original data set. When using feature extraction techniques such as PCA, MNF, ICA and the proposed CMTMF, the number of features was varied empirically in our experiments and only the best results are reported. In all cases, a supervised classification process was performed using the SVM with Gaussian kernel (observed empirically to perform equal or better than other tested kernels, such as polynomial or linear). Kernel parameters were optimized by a grid search procedure, and the optimal parameters were selected using 10-fold cross-validation. The LIBSVM library† was used for implementation purposes. In order to evaluate the ability of the tested methods to perform under training sets with different number of samples, we randomly selected 5% and 15% of the pixels in each ground-truth class in Fig. 2(b) and used them to build the training set. The remaining pixels were used as test pixels. Based on the aforementioned training sets, the overall (OA) and average (AA) classification accuracies[7] were computed over the remaining test samples for each data set. Each experiment was repeated ten times to guarantee statistical consistency, and the average results are reported. Fig. 3 shows the classification maps obtained by using PCA, MNF, ICA and the proposed CMTMF in some of the conducted experiments.

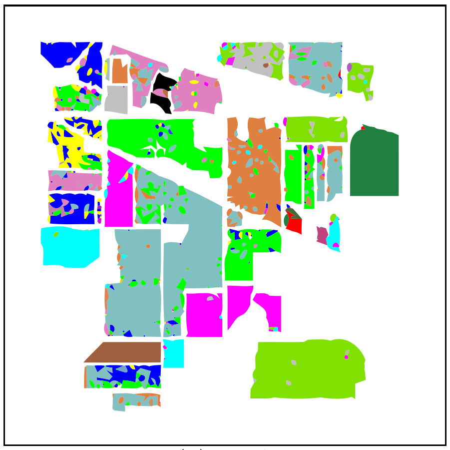
Tables 1 and 2 respectively show the OA and AA (in percentage) obtained by the considered classification system for the AVIRIS hyperspectral image using the original spectral information as input feature, and also the features provided by PCA, MNF, ICA and the proposed CMTMF. It is important to emphasize that, in the tables, we only report the best case (meaning the one with highest OA) after testing numbers of extracted features ranging from 5 to 40. In all cases, this range was sufficient to observe a decline in classification OA after a certain number of features, so the number given in the parentheses can be considered a good approximation to the optimal number of features for each considered technique (in the case of the original spectral information, the number in the parentheses corresponds to the number of bands of the original hyperspectral image). From Tables 1 and 2, we can observe that the proposed technique leads to results which are comparable to those provided by the MNF and superior to those reported for the PCA and ICA, as well as to the use of the full spectral information available. An important aspect to be outlined is that the MNF needs comparatively less features than CMTMF to achieve its optimal classification accuracy, as indicated by Tables 1 and 2.

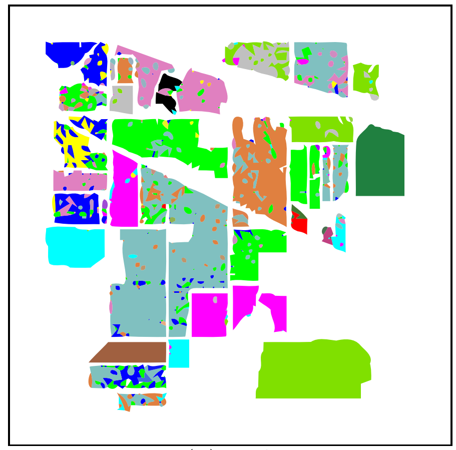
An arising question at this point is whether there is any advantage of using the proposed CMTMF versus the MNF transform. Since both methods are unsupervised and leading to similar classification results, with the MNF requiring less components than CMTMF to produce its optimal result, it is reasonable to argument if there exists any advantage of using an unmixing-based technique such as CMTMF over a well-known, statistical

*http://dynamo.ecn.purdue.edu/biehl/MultiSpec
†http://www.csie.ntu.edu.tw/cjlin/libsvm/

Table 2. Average classification accuracy (in percentage) obtained by different features extracted from the AVIRIS Indian Pines scene. Only the best case is reported for each considered feature extraction technique (with the optimal number of features given in the parentheses).

| Feature extraction | 5% Training | 15% Training |
|---|---|---|
| None | 66.37% (202) | 79.52% (202) |
| PCA | 70.59% (15) | 80.37% (15) |
| ICA | 70.03% (15) | 80.03% (15) |
| MNF | 83.31% (10) | 89.04% (10) |
| CMTMF | 82.17% (20) | 89.55% (20) |


(a) PCA


(b) ICA


(c) MNF


(d) CMTMF

Figure 3. Classification results after applying different feature extraction methods to the AVIRIS Indian Pines scene (using an SVM classifier with Gaussian kernel, trained with 5% of the available samples).

method such as the MNF. To answer this question, we emphasize that the components provided by CMTMF can be interpreted in a physical manner (as the abundances of representative spectral constituents in the scene). This suggests that unmixing-based feature extraction techniques can provide an alternative strategy to classic feature extraction chains such as the MNF with a main difference: while standard feature extraction techniques such as the MNF do not incorporate the pure/mixed nature of the pixels in hyperspectral data, unmixing-based techniques such as the CMTMF incorporate information about mixed pixels, which are the dominant type of pixel in hyperspectral images. An additional aspect is that the CMTMF implicitly includes spatial information (accomplished in this work through a simple clustering strategy for endmember extraction). However, the true fact is that our comparative assessment only indicates a moderate improvement (or comparable performance) of CMTMF with regards to the MNF and further improvements are needed in future developments of this method.

## 4. CONCLUSIONS AND FUTURE LINES

In this paper, we have developed a new unmixing-based feature extraction technique based on clustering followed by mixture-tuned matched filtering (CMTMF) that implicitly includes spatial information through a combination of unsupervised clustering and partial spectral unmixing. We have compared our newly developed technique with other classic techniques for feature extraction. Our quantitative assessment, conducted using a well-known hyperspectral scene collected by AVIRIS over the Indian Pines region in Indiana, indicates that our newly developed technique provides results which are superior to other standard techniques such as the PCA or ICA, and comparable to those provided by the MNF, with the advantage that it provides physically meaningful components that can be interpreted as the abundance of spectrally pure materials materials in the scene. In turn, the MNF needs to derive comparatively less components to achieve optimal performance. Future developments of this work will include an investigation of additional techniques for feature extraction from a spectral unmixing point of view, in order to fully substantiate the advantages that can be gained at the feature extraction stage by including additional information about mixed pixels (which are predominant in hyperspectral images) to improve classification results. Another research line deserving future attention is the determination of automatic procedures able to determine *a priori* the optimal number of features to be extracted from the original hyperspectral image.

## ACKNOWLEDGEMENT

This work has been supported by the European Community's Marie Curie Research Training Networks Programme under reference MRTN-CT-2006-035927 (HYPER-I-NET). Funding from the Spanish Ministry of Science and Innovation (HYPERCOMP/EODIX project, reference AYA2008-05965-C04-02) and Junta de Extremadura (PRI09A110 and GR10035 projects) are also gratefully acknowledged.

## REFERENCES

1. A. F. H. Goetz, G. Vane, J. E. Solomon, and B. N. Rock, "Imaging spectrometry for earth remote sensing," *Science* **228**, pp. 1147–1153, 1985.
2. R. O. Green, "Imaging spectroscopy and the airborne visible-infrared imaging spectrometer (AVIRIS)," *Remote Sensing of Environment* **65**, pp. 227–248, 1998.
3. M. E. Schaepman, S. L. Ustin, A. Plaza, T. H. Painter, J. Verrelst, and S. Liang, "Earth system science related imaging spectroscopy – an assessment," *Remote Sensing of Environment* **113**, pp. 123–137, 2009.
4. 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* **113**, pp. 110–122, 2009.
5. J. J. Settle and N. A. Drake, "Linear mixing and estimation of ground cover proportions," *International Journal of Remote Sensing* **14**, pp. 1159–1177, 1993.
6. N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine* **19**(1), pp. 44–57, 2002.
7. J. A. Richards, "Analysis of remotely sensed data: the formative decades and the future," *IEEE Trans. Geosci. Remote Sens.* **43**, pp. 422–432, 2005.
8. 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* **91**, pp. 8098–8112, 1986.
9. 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.* **42**(3), pp. 650–663, 2004.
10. C. A. Bateson and B. Curtiss, "A Method for Manual Endmember Selection and Spectral Unmixing," *Remote Sensing of Environment* **55**(3), pp. 229–243, 1996.
11. D. Heinz and C.-I. Chang, "Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.* **39**, pp. 529–545, 2001.
12. D. M. Rogge, B. Rivard, J. Zhang, and J. Feng, "Iterative spectral unmixing for optimizing per-pixel endmember sets," *IEEE Trans. Geosci. Remote Sens.* **44**(12), pp. 3725–3736, 2006.

13. A. Villa, J. Chanussot, J. A. Benediktsson, and C. Jutten, "Spectral unmixing for the classification of hyperspectral images at a finer spatial resolution," *IEEE Journal of Selected Topics in Signal Processing* **5**(3), pp. 1932–4553, 2011.
14. B. Luo and J. Chanussot, "Hyperspectral image classification based on spectral and geometrical features," *Proceedings of the IEEE International Workshop on Machine Learning for Signal Processing* , pp. 1–6, 2009.
15. B. Luo and J. Chanussot, "Unsupervised classification of hyperspectral images by using linear unmixing algorithm," *Proceedings of the IEEE International Conference on Image Processing* , pp. 2877–2880, 2009.
16. 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.* **3**, pp. 93–97, 2006.
17. L. Bruzzone, M. Chi, and M. Marconcini, "A novel transductive SVM for the semisupervised classification of remote sensing images," *IEEE Trans. Geosci. Remote Sens.* **44**, pp. 3363–3373, 2006.
18. J. A. Gualtieri and R. F. Cromp, "Support vector machines for hyperspectral remote sensing classification," in *Advances in Computer-Assisted Recognition*, R. J. Mericsko, ed., *Proceedings of SPIE* **3584**, pp. 221–232, 1999.
19. I. Dopido, M. Zortea, A. Villa, A. Plaza, and P. Gamba, "Unmixing prior to supervised classification of remotely sensed hyperspectral images," *IEEE Geosci. Remote Sens. Lett.* **8**, pp. 760–764, 2011.
20. I. Dopido, A. Villa, A. Plaza, and P. Gamba, "A comparative assessment of several processing chains for hyperspectral image classification: What features to use?," *Proceedings of the IEEE/GRSS Workshop on Hyperspectral Image and Signal Processing: Evolution in remote Sensing* **1**, 2011.
21. I. Dopido and A. Plaza, "Unmixing prior to supervised classification of urban hyperspectral images," *Proceedings of the IEEE GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas* **1**, 2011.
22. M. Rojas, I. Dopido, A. Plaza, and P. Gamba, "Comparison of support vector machine-based processing chains for hyperspectral image classification," *Proceedings of the SPIE Optics and Photonics, Satellite Data Compression, Communication, and Processing Conference* **1**, 2010.
23. J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis: An Introduction*, Springer, 2006.
24. 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.* **26**, pp. 65–74, 1988.
25. P. Comon, "Independent component analysis, a new concept?," *Signal Processing* **36**(3), pp. 287 – 314, 1994.
26. C.-I. Chang, J.-M. Liu, B.-C. Chieu, C.-M. Wang, C. S. Lo, P.-C. Chung, H. Ren, C.-W. Yang, and D.-J. Ma, "Generalized constrained energy minimization approach to subpixel target detection for multispectral imagery," *Optical Engineering* **39**, pp. 1275–1281, 2000.
27. 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.
28. C.-I. Chang, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*, Kluwer Academic/Plenum Publishers: New York, 2003.
29. J. A. Hartigan and M. A. Wong, "Algorithm as 136: A k-means clustering algorithm," *Journal of the Royal Statistical Society, Series C (Applied Statistics)* **28**, pp. 100–108, 1979.
30. D. Tuia and G. Camps-Valls, "Urban image classification with semisupervised multiscale cluster kernels," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **4**, pp. 65–74, 2011.