New Hyperspectral Unmixing Techniques in the Framework of the Earth Observation Optical Data Calibration and Information Extraction (EODIX) Project

Antonio Plaza, Javier Plaza, Inmaculada Dópido, Gabriel Martín, Marian-Daniel Iordache, and Sergio Sánchez
Hyperspectral Computing Laboratory
Department of Technology of Computers & Communications
University of Extremadura, Avda. de la Universidad s/n, 10071 Cáceres, Spain
Emails: {aplaza, jplaza, inmaculadadopido, gamahefpi, diordache, sersanmar}@unex.es

ABSTRACT- *The main objective of the Earth Observation Optical Data Calibration and Information Extraction (EODIX) project is to develop advanced ground segment methodologies for optical data calibration and information extraction from a broad family of imaging instruments currently in orbit. EODIX also has an important component related with advanced information extraction from remotely sensed data sets. In this paper, we describe some of the advances carried out in one of the tasks of the EODIX project, focused on the area of spectral unmixing of remotely sensed data. It is well known that most pixels in remotely sensed images are characterized by their mixed nature and can be modelled as the combination of elementary components (called endmembers) with variable per-pixel fractional abundances. The techniques described in this paper comprise several new algorithms for endmember extraction (including not only spectral information but also spatial information) and abundance estimation (with particular focus on the sparsity of the abundance estimation problem, which can be used to increase the accuracy of the estimations). The paper also describes computationally efficient implementations of some of the discussed algorithms.*

1. INTRODUCTION

Hyperspectral imaging instruments are capable of collecting hundreds of images, corresponding to different wavelength cha-nnels, for the same area on the surface of the Earth (Goetz, 1985). For instance, NASA is continuously gathering imagery data with instruments such as the Jet Propulsion Laboratory's Airborne Visible-Infrared Imaging Spectrometer (AVIRIS), able to record the visible and near-infrared spectrum (wavelength region from 0.4 to 2.5 micrometers) of the reflected light of an area 2 to 12 kilometers wide and several kilometers long using 224 spectral bands, thus allowing spectral signature-based analyses as shown by Fig. 1. On the other hand, Fig. 2 illustrates the problem of mixed pixels challenging hyperspectral data interpretation.

Spectral mixture analysis (or spectral unmixing) has been an alluring exploitation goal since the earliest days of imaging spectroscopy (Keshava and Mustard, 2002). No matter the spatial resolution, in natural environments, spectral signatures in hyperspectral data are invariably a mixture of the signatures of the various materials found within the spatial extent of the ground instantaneous field view. The mixing systematics can be inherently linear or nonlinear. On the one hand, the linear model assumes that pure spectral components (*endmembers*) are sitting side-by-side within the field of view of the imaging instrument, as illustrated in Fig. 3(top). On the other hand, the nonlinear model assumes that multiple scattering effects dominate the interaction between incident radiation and the response measured at the imaging spectrometer, as illustrated in Fig. 3(bottom). The linear model is generally adopted in practice due to its simplicity and independence of physical properties, but it may provide more accurate results (Keshava and Mustard, 2002).

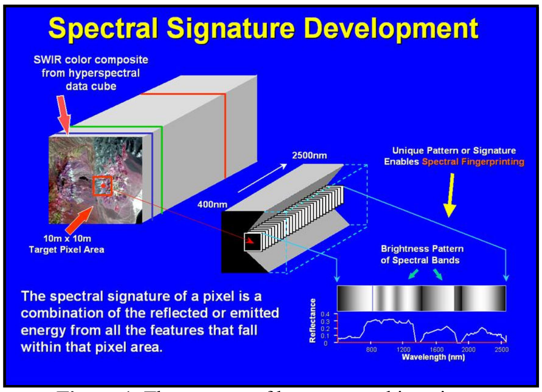

Figure 1. The concept of hyperspectral imaging.

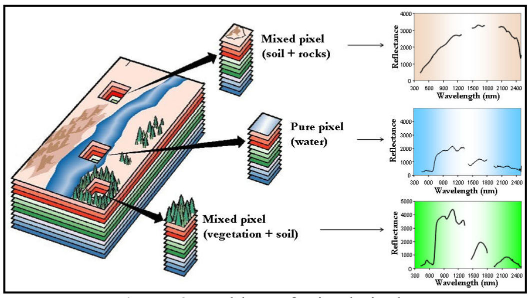

Figure 2. Problem of mixed pixels.

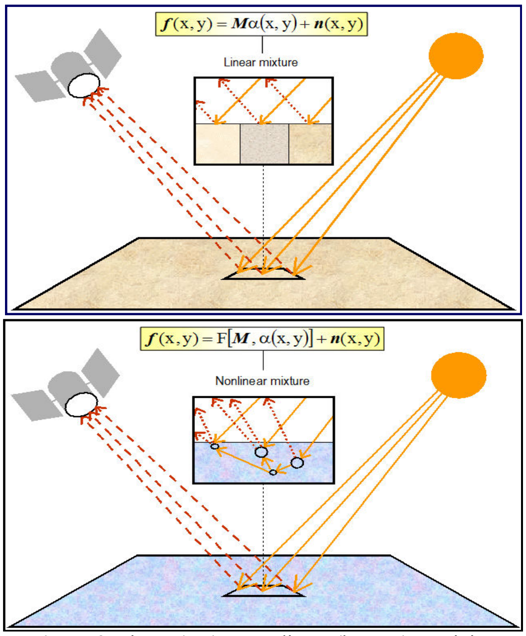

Figure 3. Linear (top) vs nonlinear (bottom) unmixing.

In the context of linear spectral unmixing, the estimated endmember fractional abundances are often required to satisfy two constraints. First, all abundances must be non-negative. Second, the sum of abundances for a given pixel must be unity. However, it is the derivation and validation of the correct suite of endmembers that has remained a challenging and elusive goal for the past several years. Several approaches have been developed for this purpose over the last few years. Many available approaches have been focused on analyzing the data in spectral terms only (Chang, 2003), i.e. trying to find the endmembers as the extreme pixels (corners) in the multi-dimensional data cloud. Examples of this convex geometry-based approach to identification of image endmembers include the pixel purity index (PPI) algorithm, the orthogonal subspace projection (OSP), the N-FINDR algorithm, or the iterative error analysis (IEA) algorithm, among others (Plaza et al., 2009).

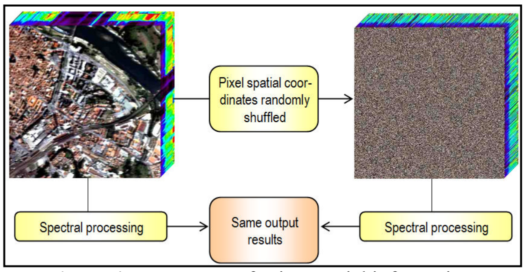

Figure 4. Importance of using spatial information.

Although these methods have shown considerable promise, they are exclusively based on the spectral information of the data. However, most endmember extraction algorithms could benefit from an integrated framework in which both the spectral information and the spatial arrangement of pixel vectors are taken into account. An example is given in Fig. 4 in which the spatial locations of pixel vectors in an urban hyperspectral scene are randomly shuffled. In this case, the application of convex geometry-based endmember extraction in the original scene and the one without spatial correlation would be the same, meaning that the rich spatial information available was not taken into account during the endmember searching process.

In this paper, we address new trends in spectral unmixing specifically developed in the framework of the Earth Observation Optical Data Calibration and Information Extraction (EODIX) project. These comprise the incorporation of spatial constraints into spectral unmixing, the use of spectral unmixing for feature extraction purposes, and the use of sparse regression-based approaches to spectral unmixing. Experimental results using a real hyperspectral data set are given at the end of the paper for illustration purposes. Combined, these topics reflect the maturity of a field that currently represents one of the most active research areas in hyperspectral image analysis.

2. INCORPORATION OF SPATIAL INFORMATION INTO SPECTRAL UNMIXING

We have recently developed a region-based spatial preprocessing technique for endmember extraction algorithms intended to exploit spectral information more effectively by adequately incorporating spatial context. Our proposed approach first adaptively searches for spectrally pure and spatially homogeneous regions by using a hybrid procedure that combines unsupervised clustering and orthogonal subspace projections, thus selecting a set of representative regions in spatial-spectral terms. This spatial preprocessing is followed by a standard endmember extraction process using the pixels located in such regions, providing a set of spatially

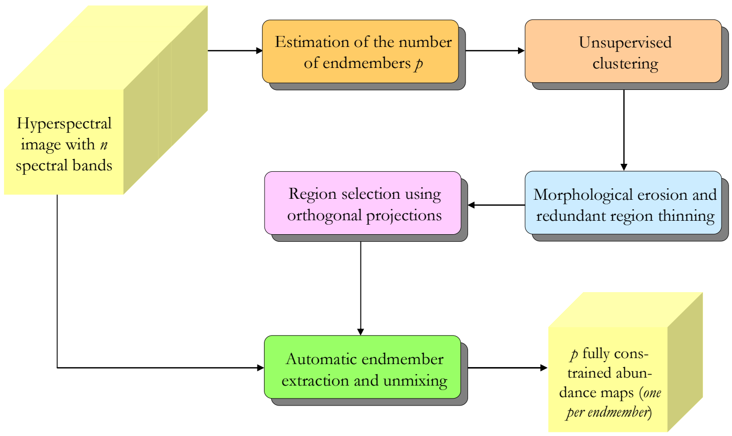

Figure 5. Region-based spatial preprocessing for endmember extraction.

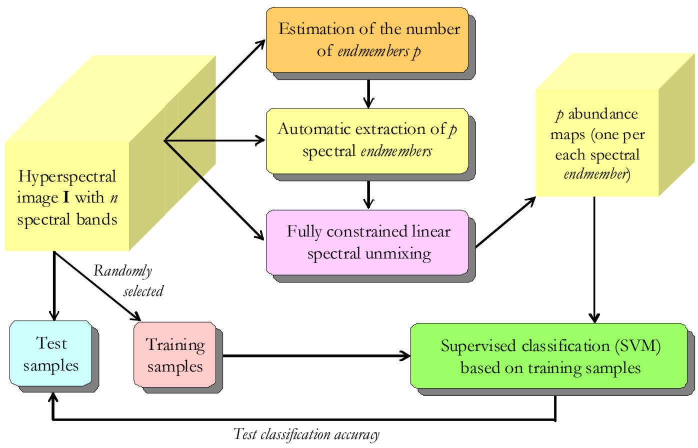

Figure 6. Spectral unmixing for feature extraction prior to supervised classification.

representative endmembers with the potential to accurately characterize large homogeneous areas in the original hyperspectral scene (Martin and Plaza, 2010). The concept of region-based spatial preprocessing, illustrated in Fig. 5, can be readily combined with algorithms that use only the spectral information contained in the original hyperspectral data set, as it is the case of the N-FINDR (Winter, 1999) in Fig. 7.

3. USING SPECTRAL UNMIXING FOR FEATURE EXTRACTION PURPOSES

Recently, we have developed a new strategy for feature extraction prior to supervised classification of hyperspectral data which is based on spectral unmixing concepts (Rojas et al., 2010). This unmixing-based feature extraction approach presents some distinctive features with regards to classic approaches commonly used in the framework of classification (such as principal component analysis) or unmixing (such as the minimum noise fraction):

- First, it provides additional information for classification in hyperspectral analysis scenarios with moderate spatial resolution, since the sub-pixel composition of training samples can be used as part of the learning process of the classifier.
- Second, it can effectively model the non-stationary behavior of the spectral signatures of land-cover classes in the spatial domain of the scene, since (possibly disjoint) regions belonging to the same class are represented by the same spectral signature, and the variations related with different cover proportions or illumination conditions are modeled via the abundance estimation process inherent in spectral unmixing.
- Third, the components estimated by the proposed feature extraction strategy exhibit physical meaning as opposed to those obtained by principal component analysis or the minimum noise transform.

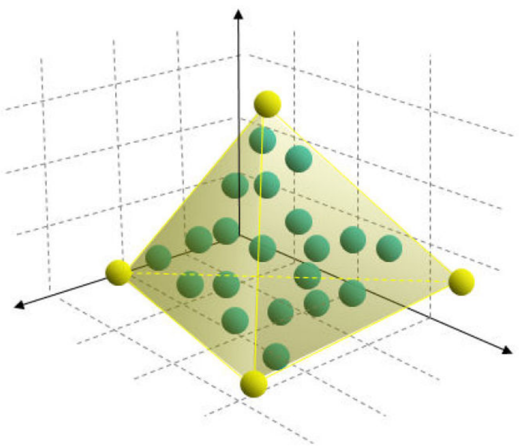

Figure 7. The N-FINDR algorithm inflates a simples with maximum volume using the pixels available in the input hyperspectral data. After a random initialization, it tries all pixels in each endmember position until the combination with maximum volume is found.

- A final advantage of the proposed approach is that it does not penalize classes which are not relevant in terms of variance or signal-to-noise ratio (SNR).

The proposed unmixing-based feature extraction strategy has been implemented in the form of a standard unmixing processing chain prior to supervised classification using a traditional method such as the support vector machine (SVM) as illustrated in Fig. 6, and also as a modified unmixing chain in which spatial information is used to guide the selection of endmembers to spatially relevant areas by means of a spatial preprocessing framework in Fig. 8.

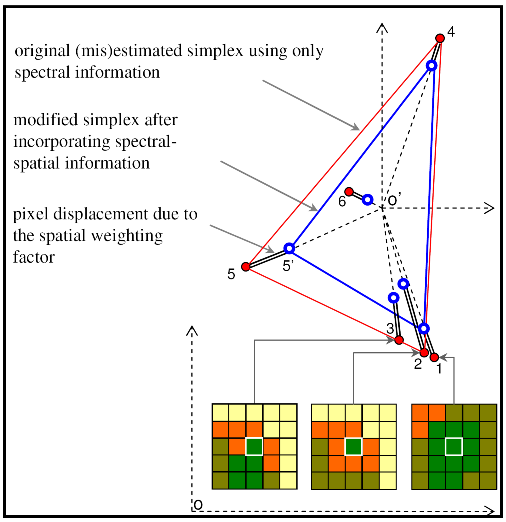

Figure 8. Spatial preprocessing to guide endmember extraction to spatially representative areas.

4. SPARSE HYPERSPECTRAL UNMIXING

The spectral unmixing problem has been recently been approached in a semi-supervised fashion, by assuming that the observed image signatures can be expressed in the form of linear combinations of a number of pure spectral signatures known in advance (e.g., spectra collected on the ground by a field spectro-radiometer). Unmixing then amounts to finding the optimal subset of signatures in a (potentially very large) spectral library that can best model each mixed pixel in the scene (Iordache et al., 2010). In practice, this is a combinatorial problem which calls for efficient linear sparse regression techniques based on sparsity-inducing regularizers, since the number of endmembers participating in a mixed pixel is usually very small compared with the (ever-growing) dimensionality and availability of spectral libraries to spatially relevant areas (Bioucas-Dias and Plaza, 2010).

5. PARALLEL IMPLEMENTATION OF A FULL HYPERSPECTRAL UNMIXING CHAIN

The unmixing techniques introduced in previous sections can be extremely time consuming when applied to real hyperspectral data sets. At the same time, these techniques exhibit inherent parallelism at multiple levels: across pixel vectors (coarse grained pixel-level parallelism), across spectral information (fine grained spectral-level parallelism), and even across tasks (task-level parallelism). As a result, they map nicely to massively parallel systems such as graphics processing units (GPUs) or field programmable gate arrays (FPGAs). The latter are particularly suitable to on-board processing due to low power consumption and tolerance to radiation in space. We have developed parallel implementations of a full hyperspectral unmixing chain for both types of platforms: GPUs (Sanchez et al., 2010) and FPGAs (Gonzalez et al., 2010). These platforms are respectively illustrated in Figs. 9 and 10.

Figure 9. NVidia commodity graphic processing unit.

6. CONCLUSION AND SUMMARY

In this paper, we have described several new techniques for spectral unmixing of hyperspectral data developed in the framework of the Earth Observation Optical Data Calibration and Information Extraction (EODIX) project, comprising new methods for endmember extraction (including not only spectral information but also spatial information) and abundance estimation (with particular focus on the sparsity of the abundance estimation problem, which can be used to increase the accuracy of the estimations). The paper also described computationally efficient implementations of some of the discussed algorithms. Although the techniques have been functionally described, experimental evidence of their success can be found in the provided references. We would like to emphasize that this work is part of a much larger strive to fully incorporate the advantages that can be obtained by spectral unmixing into the analysis of remotely sensed hyperspectral data sets. In this regard, complementary lines of research that we are planning to address in the near future comprise the refinement of some of the presented techniques (most notably, the inclusion of spatial information into sparse unmixing methods) and also the development of additional parallel implementations for some of the discussed algorithms, which are often dominated by regular computations.

ACKNOWLEDGEMENT

This work has been supported by the European Community's Marie Curie Research Training Networks Programme under reference MRTN-CT 2006-035927, Hyperspectral Imaging Network (HYPER-I-NET) coordinated at University of Extremadura (UEX) by Antonio Plaza. Daniel Iordache is sponsored by a research fellowship associated to this project. This work has also been supported by the Spanish Ministry of Science and Innovation (EODIX project, reference AYA2008-05965-C04-00) coordinated at University of Valencia by José Sobrino and at UEX by Antonio Plaza. Inmaculada Dópido, Gabriel Martín and Sergio Sánchez are sponsored by research fellowships associated to this project. Funding from Junta de Extremadura (local government) under project PRI09A110, coordinated at UEX by Javier Plaza, and from the European Cooperation in Science and Technology (COST) programme through the project entitled Open European Network for High Performance Computing on Complex Environments (ComplexHPC) coordinated at UEX by Antonio Plaza and led by Emmanuel Jeannot at INRIA, are also very gratefully acknowledged.

Figure 10. Xilinx field programmable gate array.

REFERENCES

Bioucas-Dias, J., and Plaza, A., 2010, Hyperspectral unmixing: geometrical, statistical and sparse regression-based approaches. *Proceedings of SPIE*, in press.

Chang, C.-I, 2003, Hyperspectral Imaging: Techniques for Spectral Detection and Classification (New York, USA: Kluwer).

Goetz, A., 1985, Imaging spectrometry for Earth remote sensing. *Science*, 228, 1147–1153.

Gonzalez, C., Resano, C., Mozos, D., Plaza, A., and Valencia, D., 2010, FPGA implementation of the pixel purity index algorithm for remotely sensed hyperspectral image analysis. *EURASIP Journal on Advances in Signal Processing*, 969806, 1–13.

Iordache, M. D., Bioucas-Dias, J., and Plaza, A., 2010, Sparse unmixing of hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, under review.

Keshava, N., and Mustard, J. F., 2002, Spectral unmixing. *IEEE Signal Processing Magazine*, 19, 44–57.

Martin, G. and Plaza, A., 2010, Region-based spatial preprocessing for endmember extraction and spectral unmixing. *IEEE Geoscience and Remote Sensing Letters*, under review.

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

Rojas, M., Dopido, I., Plaza, A., and Gamba, P., 2010, Comparison of support vector machine-based processing chains for hyperspectral image classification. *Proceedings of SPIE*, 7810, 1–10.

Sanchez, S., Paz, A., Martin, G., and Plaza, A., 2010, Parallel unmixing of remotely sensed hyperspectral images on commodity graphics processing units. *Concurrency and Computation: Practice & Experience*, under review.

Winter, M. E., 1999, N-FINDR: An algorithm for fast autonomous spectral endmember determination in hyperspectral data. *Proceedings of SPIE*, 3753, 266–277.