PhD Work Plan

Phd Student: *Marian-Daniel Iordache*

Advisor: *José Bioucas-Dias,* IST, Portugal

Co-adviser: *Antonio J. Plaza,* UEX, Espanha

1. *Introduction*

1.1 Hyper-i-net RTN

Marian-Daniel Iordache has been appointed as an Early Stage Researcher (ESR) in the Hyperspectral Imaging Network (HYPER-I-NET), which is a four-year (2007-2010) FP6 Marie Curie Research Training Network (RTN) designed to build an interdisciplinary European research community focusing on hyperspectral imaging activities.

While pursuing a PhD degree, the appointed ESRs will be exposed to the following network-wide training activities provided by the RTN**:** a) researcher exchange programme; b) summer schools; c) summer camps; d) round-robin calibration experiments; e) general-purpose research visits; f) mid-term and final network workshop. For further details see Hyper-i-net at: http://hyperinet.multimediacampus.it/.

Each ESR will develop specific training activities defined on the respective contract, which involves at least two partners. In the case of Marian-Daniel Iordache, the partners are the Instituto Superior Técnico and the University of Extremadura (UEX), Cáceres, Spain.

.

2. *PhD Theme: Efficient endmember extraction and spectral unmixing algorithms*

Hyperspectral unmixing is a source separation problem [1-5]. Compared with the canonical source separation scenario, the sources in hyperspectral unmixing (i.e., the materials present in the scene) are statistically dependent and may combine in a nonlinear fashion. These characteristics, together with the high dimensionality of the hyperspectral vectors, place the unmixing of hyperspectral mixtures beyond the reach of most source separation algorithms, thus fostering active research in this theme.

The research to be carried out in this PhD thesis will address the following aspects of hyperspectral unmixing: a) geometrical based unmixing in the absence of pure pixels; b) sparse signature pursuit denoising; and c) efficient implementation of the developed algorithms. These topics are addressed below.

2.1 Minimum volume transform (MVT) unmixing

Given a set of mixed spectral (multispectral or hyperspectral) vectors, linear spectral mixture analysis, or linear unmixing, aims at estimating the number of reference substances, also called endmembers, their spectral signatures, and their abundance fractions. Under the linear mixing model, the spectral vectors acquired by a hyperspectral sensor from a scene are in a simplex whose vertices correspond to the pure materials present in the scene (endmembers). Several approaches [6],[7],[8] have exploited this geometric feature of hyperspectral mixtures, under the additional assumption that at least one pure pixel of per material is present in the scene. This assumption does not hold, however, in many data sets. In these cases, although the endmembers may still correspond to the vertices of the data simplex, there are no pixels in the data corresponding to the vertexes. Therefore, the pure pixel based algorithms yield poor results and we have to resort to minimum volume transform (MVT) algorithm [8].

As the name suggests, MVT algorithms determine the simplex of minimum volume containing the data. The MVT type approaches are complex from the computational point of view. Usually, these algorithms first find the convex hull defined by the observed data and then fit a minimum volume simplex to it. For example, the *gift wrapping algorithm* [9] computes the convex hull of data points in a d-dimensional space with a computational complexity of $O(n^{d/2+1})$, where n is the number of samples.

This PhD work shall address the MVT approach to unmixing. A possible starting point is a modification of the algorithm recently introducing in [11] to make it more efficient from the computational point of view. In fact, this algorithms yields very good results for a low number of endmembers, say 10. If the number of endmenmber is larger, the algorithm time complexity is simply too high to be used in practical applications.

Another direction is the exploitation of spatial continuity in a statistical sense [12], [13]: it is likely that neighboring pixels display close mixtures. This information can be inserted into the unmixing criteria in order to obtain better unmixing results.

2.2 Signature pursuit denoising

Very often, the hyperspectral mixed vectors can not be modeled as a linear combination of a few enmembers. The reasons may be, e.g., spectral variability or nonlinear scattering phenomema. In this cases, the canonical linear unmixing does not produce reasonable results. A line of attack for the unmixing consists in modeling each observed vector as a linear combination of only a few spectral vectors from a large dictionary.

These vectors are chosen as to minimize the regression error. Examples of this methodology are published in [14], [15].

The concept described above is exactly that of “basis puisuit denoising” [16], where a signal is reconstructed from a small set of atomic signals taken from an overcomplete dictionary. The reconstruction is computed by solving an inverse problem with an l1-norm regularizer, thus inducing sparsity in the solution.

This PhD work shall attach the multiple endmembers rationale with the “basis puisuit denoising” tools. In this way, the exponential complexity of selecting the smallest number of spectral vector from the dictionary is overcame. We term this research direction “signature pursuing denoising”.

2.3 Efficient implementation of the developed algorithms

Hyperspectral sensors sample the reflected solar radiation from the Earth surface in the portion of the spectrum extending from the visible region through the near-infrared and mid-infrared (wavelengths between $0.3\,\mu m$ and $2.5\,\mu m$) in hundreds of narrow (on the order of $10\,nm$) contiguous bands. This high spectral resolution yields large amounts of data. For example, AVIRIS collects a 512 (along track) $\times 614$ (across track) $\times 224$ (bands) $\times 12$ (bits) data cube in 43 seconds, corresponding to more than 700 Mbits; Hyperion collects 4 Mbits in 3 seconds, corresponding to 366Kbytes/Km^2. Such huge data volumes put stringent requirements in what concerns communications, storage, and processing. Therefore, the development of efficient (from the computational point of view) algorithms plays a key rule in the exploitation of the enormous potential of hyperspectral imagery in a vast plethora of remote sensing applications.

Algorithm efficiency shall always be present in this PhD work, namely by designing and implementing parallel algorithms in line with work [17].

3 Plan of the Activities

This section schedules the PhD activities for a time period of three years. The first year will be mostly devoted to take courses from PhD doctoral program in ECE. The core of the research and scientific contributions will take place in the second year and in the first half of the third year. The second half part of the third year will be devoted to the writing of the thesis. These activities and a few expected outcomes are detailed below.

First Year

- **Courses from the PhD doctoral program in ECE**

The selected courses are consistent with Iordache's background and are appropriate to her general research area.

Course	Year/Semester
Optimization	1º/1º
Dynamic Stochastic Estimation, Prediction and Smoothing	1º/1º
Inverse Problem in Signal and Image Processing	1º/2º
Learning	1º/2º
Detection, Estimation, and Filtering	1º/2º

- A report presenting a taxonomy of the hyperspectral unmixing, dimensionality reduction, and parallel implementation of these algorithms.

- Development of programming skills on the matlab and IDL/ENVI platforms. Devlop an application in matlab and IDL oriented to compare and evaluate unmixing and dimensionality reduction algorithms.

- Publishing two conference papers.

Second Year

- Research on minimum volume transform unmixing and on signature pursuit denoising.

- Writing of the thesis proposal.
- Three conference papers and a journal paper.

Third Year

- Research on signature pursuit denoising and on efficient implementation.
- Writing of the PhD thesis.
- Three conference paper and a journal paper.

References

[1] N. Keshava, J. Kerekes, D. Manolakis, and G. Shaw, "An Algorithm taxonomy for hyperspectral unmixing", Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VI. Sylvia S. Shen and Michael R. Descour, Eds., p.42, SPIE vol. 4049, 2000.

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

[3] A. Plaza, P. Martinez, R.M. 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, March 2004.

[4] J. Nascimento and J. Dias, "Vertex component analysis: a fast algorithm to unmix hyperspectral data", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 4, pp. 898-910, 2005.

[5] L. Miao and H. Qi, " Endmember Extraction From Highly Mixed Data Using Minimum Volume Constrained Nonnegative Matrix Factorization", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 3, pp. 765-777, 2007.

[6] A. Ifarraguerri and C.-I Chang, “Multispectral and hyperspectral image analysis with convex cones,” *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 2, pp. 756–770, 1999.

[7] J. Boardman, “Automating spectral unmixing of AVIRIS data using convex geometry concepts,” in *Summaries 4th Annu. JPL Airborne Geoscience Workshop*, vol. 1, 1993, JPL Pub. 93-26, pp. 11–14.

[8] M. D. Craig, “Minimum-volume transforms for remotely sensed data,” *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 1, pp. 99–109, 1994.

[9] R. Seidel, *Convex Hull Computations*. Boca Raton, FL: CRC, ch. 19, pp. 361-375, 1997.

[10] C. Bateson, G. Asner, and C. Wessman, “Endmember bundles: A new approach to incorporating endmember variability into spectral mixture analysis,” *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 2, pp. 1083–1094, Mar. 2000.

[11] J. Li and J. Bioucas-Dias, "Minimum volume simplex analysis: a fast algorithm to unmix hyperspectral data", in *Proceedings of the International Geoscience and Remote Sensing Symposium- IGARSS’08*, 2008.

[12] . A. Zymnis, S.-J. Kim, J. Skaf, M. Parente, and S. Boyd Hyperspectral Image Unmixing via Alternating Projected Subgradients, in Proceedings Asilomar Conference, 2007.

[13] J. Borges, J. Bioucas-Dias and A. Marçal, “Bayesian hyperspectral image segmentation with discriminative class learning,” in *Pattern Recognition and Image Analysis: 3rd Iberian Conference*, *IbPRIA 2007*, Girona, Spain, 2007, Lecture Notes in Computer Science, vol. 1, pp. 22-29, Springer-Verlag Heidelberg.

[14] A. Roberts, M. Gardner, R. Church, S. Ustin, G. Scheer, and R. Green, “Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models”, in *Remote Sensing of Environment*, vol. 65, pp. 267 – 279, 1998.

[15] M. Bachmann, A. Ller, M. Habermeyer, M. Schmidt, S. Dech “Iterative MESMA unmixing for fractional cover estimates - evaluating the portability”, in *Proceedings of the 4th EARSeL Workshop on Imaging Spectroscopy*, Warsaw, 2005

[16] S. Chen, D. Donoho, and M. Saunders, “Atomic Decomposition by basis pursuit” SIAM Journal on Scientific Computing, 1995.

[17] A. Plaza “Parallel processing of remotely sensed hyperspectral imagery: full-pixel versus mixed-pixel classification”. Concurrency and Computation: Practice & Experience, 2008 (accepted for publication).

April 20, 2008

José Manuel Bioucas Dias

(Prof. of the Instituto Superior Técnico)

António Plaza

(Prof. of the University of Extremadura)

Marian-Daniel Iordache