UNMIXING SPARSE HYPERSPECTRAL MIXTURES

Marian-Daniel Iordache[1], José Bioucas-Dias[1] and António Plaza[2]

[1]Instituto de Telecomunicações, Instituto Superior Técnico, Av. Rovisco Pais, 1049-001, Lisbon, Portugal
[2]Department of Technology of Computers and Communications, University of Extremadura, Avda. de la Universidad S/N, 10.071, Cáceres, Spain

ABSTRACT

Spectral unmixing is the process of estimating the number of materials present in a scene, their spectral signatures and their corresponding fractional abundances.

The hyperspectral sensors are able to sample the electromagnetic spectrum (from the visible to the near-infrared region) in tens or hundreds of contiguous narrow spectral bands, with a spectral resolution which can reach less than 10 nm. This is one of the most valuable characteristics of these sensors, as it allows the analysis at sub-pixel level. As the spatial resolution of the hyperspectral sensors is typically tens of meters (e.g., 30 m for Hyperion), the spectrum acquired by the sensor for one pixel is a combination of the spectral signatures of the materials present in the respective pixel (also called *endmembers*). We will refer at this type of pixel as *mixed pixel* (a pixel containing only one material is called a *pure pixel*).

There are two main scenarios used in spectral unmixing problems:

i) the incident radiation only interacts with one component; in this scenario, the measured spectrum is a linear combination of the spectral signatures of materials present in the scene, weighted by their corresponding fractional abundances. This is called *the Linear Mixing Model (LMM)*.

ii) the endmembers form an intimate mixture inside the respective pixel; the incident radiation is affected by multiple scattering effects (it interacts with more than one component). This scenario corresponds to *the Non-Linear Mixing Model*.

The non-linear mixing models are much more complex than their linear counterparts; they depend, very often, on scene parameters, which are hard and expensive to obtain; for these reasons, the non-linear mixing models are usually applied in particular cases. LMMs, on the contrary, are of widespread use in hyperspectral unmixing applications. They provide an accurate modeling of checkerboard type scenes, where the endmembers occupy distinct regions of the area inside the pixel.

Mathematically, the LMM is expressed as follows:

$$R = S \cdot f + n, \quad (1)$$

where R is an L-by-1 column vector (the measured spectra of the pixel), S is an L-by-m matrix called "the mixing matrix" (the spectral signatures of the endmembers), f is an m-by-1 column-vector (the respective fractional abundances of the endmembers), n is an L-by-1 column-vector (the noise at each band), L is the number of spectral bands, and m is the total number of endmembers present in the pixel.

The fractional abundances of the constituent endmembers are subject to two constraints:

1) the non-negativity constraint: the fractional abundances can not be negative (a material can not occupy a "negative area" inside the pixel);
2) the sum-to-one constraint: the fractional abundances of the endmembers should sum to one (the total area occupied by all the endmembers inside the pixel is equal to the area represented by the respective pixel).

A semi-supervised approach to deal with the linear spectral unmixing problem consists of assuming that the observed spectral vectors are linear combinations of a small number of spectral signatures. Unmixing amounts, then, to find a small number of materials in the spectral library that best represent the observed data. Finding a small number of signatures in a large library is a combinatorial problem which calls for efficient sparse regression techniques.

The unmixing problem is an ill-posed inverse problem, as the mixing matrix S is not square, thus it is not invertible. Moreover, the properties of the mixing matrix affect dramatically the unmixing process, because, if it has small singular values, these are sources of instabilities in the computational process.

This paper studies the efficiency of the sparse regression techniques in the spectral unmixing problem, by conducting a comparison between four different algorithms: Moore-Penrose Pseudoinverse, Orthogonal Matching Pursuit (OMP, Y. C. Pati, P. Rezahfar and P. S. Krishnaprasad – 1993), Iterative Spectral Mixture Analysis (ISMA, Derek M. Rogge, Benoit Rivard, Jinkai Zhang, Jilu Feng - 2006) and Two-Step Iterative Shrinking/Thresholding (TwIST) algorithm (José M. Bioucas-Dias and Mário A. T. Figueiredo – 2007).

The *Moore-Penrose Pseudoinverse* is an unconstrained method to unmix pixels, using the pseudoinverse of matrix *S* to find the abundance fraction of the endmembers.

Orthogonal Matching Pursuit is an iterative technique derived from *Matching Pursuit* algorithm (Mallat and Zhang - 1992). This algorithm computes, at each step, a residual that is orthogonal to all endmembers contained by the model at the respective iteration, and stops when a stop criteria is satisfied.

ISMA determines, first, an unconstrained solution using the entire spectral library. After computing this first solution, it removes, at each iteration, the endmember with the lowest fractional abundance and updates the fractional abundances vector, until one endmember remains (excluding shade). The optimal endmember set is found by examining the change in root-mean-squared (RMS) error for every two successive iterations.

TwIST is an iterative method which combines the advantages of IST (Iterative Shrinkage/Thresholding) and IRS (Iterative Re-Weighted Shrinkage) algorithms. *TwIST* solves efficiently a sparse regression problem by minimizing an *l2-l1* objective function. The *l2* term measures the data misfit and the *l1* term measures the lack of sparseness. The update equation depends on the two previous estimates. *TwIST* is an efficient method with general applicability in linear inverse problems.

The algorithms are tested for different spectral libraries *S* (each of them with a different condition number), different SNR (signal-to-noise ratios) and different number of endmembers present in the pixel. The study is underlying the strong and the weak points of each algorithm and provides a quantitative and qualitative comparison between them.

Fig. 1 shows an example of graphical comparison between the four methods, for a bad-conditioned mixing matrix *S*, SNR=40dB and 6 endmembers present out of 20 contained in the spectral library.

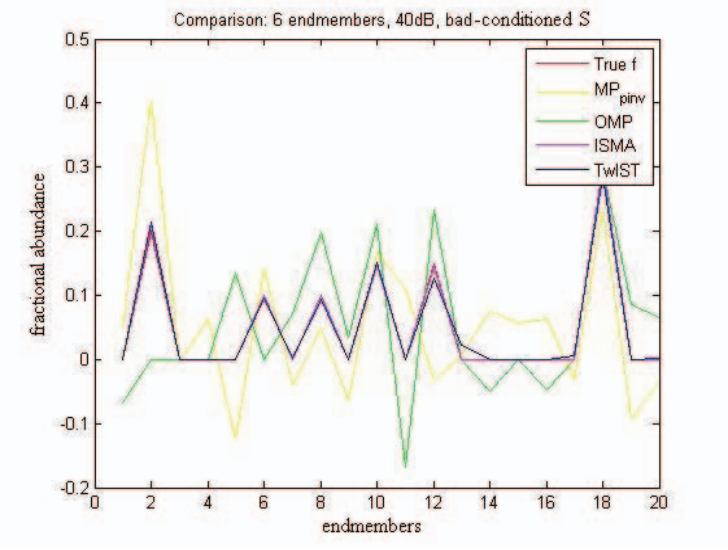

Fig. 1 Graphical comparison between the four methods

As it can be seen in fig. 1, the *Moore-Penrose Pseudoinverse* is useless when the mixing matrix is bad-conditioned, as the result is unrealistic (negative abundance fractions) because of the unconstrained character of the method. *Orthogonal Matching Pursuit* leads to accurate results when the matrix *S* is medium or well-conditioned and the noise is low, but they can be totally wrong in bad conditions: bad-conditioned matrix *S* and high level of the noise. *ISMA* finds good results most of the time, but it encounters problems for medium- or bad-conditioned *S* and high noise, because of the difficulty in choosing the appropriate stop criteria. In this last case, *TwIST* returns the most accurate results. In hard conditions, *TwIST* is the most powerful tool, as it combines in an efficient way the advantages of IST and IRS algorithms: it has a very fast convergence rate for ill-conditioned problems, like IRS, and keeps the efficiency of IST for high noise and not too bad conditioned problems.