HYPER-DEMIX: BLIND SOURCE SEPARATION OF HYPERSPECTRAL IMAGES USING LOCAL ML ESTIMATES

Simon Arberet

EPFL, Signal Processing Lab.

ABSTRACT

We propose a new method to unmix hyperspectral images. Our method exploits the structure of the material abundance maps by assuming that in some regions of the spatial dimension, only one material is present. Such regions provide a local estimate of the endmember spectrum of the corresponding material. Our main contribution is a new clustering algorithm called Hyper-DEMIX to estimate the endmember spectrum of each material based on such local estimates. The abundance map of each material is then recovered with a binary masking technique. Experimental results over noisy hyperspectral images show the effectiveness of the proposed approach.

Index Terms— Blind source separation, hyperspectral images

1. INTRODUCTION

Hyperspectral imaging spectrometers provide a set of "images" (called hyperspectral image cube) taken contiguously over a large wavelength region of the electromagnetic spectrum. The spectral range obtained through hyperspectral images (HSI) gives each material a spectral signature by which the material can be directly identified. HSI is used in applications such as environmental mapping, geological research, plant and mineral identification.

We assume that the signal measured at each pixel consists of a linear combination of individual material signatures called *endmembers* weighted by their percentage area in the pixel. The percentage area of a pixel for a particular material is called the materials *abundance*. The problem of estimating the spectral signatures of each material and the abundance of each material in each pixel is a typical multichannel blind source separation (BSS) problem.

A mixed pixel containing N spectrally distinct materials, denoted by a $M \times 1$ vector $\mathbf{x}(i,j) \in \mathbb{R}_+$ can be described by the linear model

$$\mathbf{x}(i,j) = \mathbf{A}\mathbf{s}(i,j) + \mathbf{n}(i,j) \tag{1}$$

This research was supported by the SMALL project funded by the EU FP7 FET-Open program.

where M is the number of spectral bands, (i,j) is the spatial position of the pixel, $\mathbf{A} \in \mathbb{R}_+^{M \times N}$ is a nonnegative matrix called *endmembers matrix*, commonly known as *mixing matrix* in the BSS community and $\mathbf{s}(i,j)$ is a $N \times 1$ vector where the n-th element $s_n(i,j) \in [0\ 1]$ represents the material abundance of the n-th endmember present in the pixel (i,j). $s_n(i,j), \forall (i,j)$ is commonly known as the n-th *source* in the BSS community. $\mathbf{n}(i,j)$ is a $M \times 1$ random vector of noise which is assumed to be an independent, identically distributed (i.i.d.) Gaussian process with zero mean and covariance matrix $\sigma^2\mathbf{I}$.

The hyperspectral unmixing problem can be formulated as the procedure by which given the hyperspectral image cube $\mathbf{X} = [\mathbf{x}(i,j)]_{i,j}$, we want to estimate the endmembers matrix $\mathbf{A}$, and the abundance vectors of each pixel $\mathbf{S} = [\mathbf{s}(i,j)]_{i,j}$.

Independent Component Analysis (ICA) has been used in the context of hyperspectral unmixing. The ICA Endmember Extraction Algorithm (ICA-EEA) [1] finds endmembers by decomposing the data into M independent components (using Fast-ICA algorithm). Another approach is the Automatic Target Generation Procedure ATGP [2] which works by iteratively selecting the largest magnitude vector $\mathbf{x}(i,j)$, and then projecting the data into the orthogonal space spanned by the selected endmembers.

None of the above mentioned methods (ICA, ATPG) exploit the spatial structure of the material abundance maps, i.e. the fact that the pixels corresponding to a given material are clustered in the spatial dimension. We propose in this paper a new algorithm, which exploits the spatial structure of the material abundance maps $\mathbf{S}$ to recover them from the observation of the hyperspectral image cube $\mathbf{X}$. We describe our approach in section 2. The main contribution of our unmixing approach is a feature extraction step, described in section 3, which exploits the spatial structures of the material abundance maps, and a clustering algorithm, described in section 4, which estimates the endmembers matrix. Section 5 describes the Dimension Reduction (DR) step which is used as a pre-processing step so as to reduce the computational complexity of the algorithm. In section 6, we describe the estimation of the material abundance maps $\mathbf{S}$ using a sparse approximation algorithm [3] assuming that the abundance vectors have sparsity one, i.e. only one material is present in each

pixel[1].

We build an experiment in section 7 with various level of noise so as to evaluate the proposed algorithm with the ATPG and ICA-EEA algorithms.

2. PROPOSED APPROACH

The hyperspectral unmixing is generally done in three steps as shown in Figure 1.

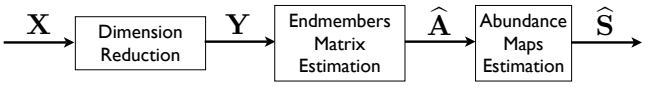


Fig. 1. Three steps approach of hyperspectral unmixing.

The first step is dimensionality reduction (DR) of the spectral dimension (i.e. the dimension of the M spectral bands) to reduce the computational complexity, the second step is the endmembers matrix estimation (in the reduced dimension), and the third step is the abundance maps estimation.

In this paper we propose a new method called HyperDEMIX, which first reduces the dimension with a nonegative PCA (NNPCA) [4] instead of a classic PCA so as to take into account and preserve the positiveness of the data. Then the endmembers are estimated locally with a corresponding *confidence value* $\widehat{\mathcal{T}}$ by exploiting the structure of the materials/sources in the spatial domain. Then the endmembers matrix is estimated by a proposed clustering algorithm which merges the local endmember estimations according to their confidence value. This clustering algorithm is inspired by DEMIX [5] and the Orthogonal Subspace Projection (OSP) approach [6]. Finally the material abundance maps are estimated using the binary masking technique [3].

3. EXPLOITING THE SPATIAL STRUCTURE OF THE SOURCES/MATERIALS

Our approach relies on the assumption that for each material n, there is at least one local region Ω_n, i.e. some neighborhood pixels, where only this material is present. A similar assumption has been already used in the audio source separation problem [7, 8, 5].

3.1. Local endmember estimation

The pixels $\mathbf{x}(i', j')$ in a region $\Omega_{i,j}$ where one material is dominant should be very similar if the noise is not too strong and the other sources are weak enough. Indeed, according to the mixture model (1), $\mathbf{x}(i', j') \approx \mathbf{a}_{n(i,j)} + \mathbf{n}(i', j')$ where $\mathbf{a}_{n(i,j)}$ is the endmember of the material $n(i, j)$ that is dominant in $\Omega_{i,j}$. Then, the mixture in such a neighborhood has a Gaussian distribution $\mathbf{x}(i', j') \sim \mathcal{N}(\mathbf{a}_{n(i,j)}, \sigma^2_{\Omega_{i,j}}\mathbf{I})$, Thus the Maximum Likelihood (ML) estimate of $\mathbf{a}_{n(i,j)}$ is:

$$\mathbf{u}_{\Omega_{i,j}} := \sum_{i',j'} \omega(i' - i, j' - j) \cdot \mathbf{x}(i', j') \tag{2}$$

where ω is a bi-dimensional window specifying the shape of the neighborhood Ω and such that $\sum_{i,j} \omega(i, j) = 1$ and $\omega(i, j) = 0, \forall (i, j) \notin \Omega$.

3.2. Confidence measure

To have an idea of how likely it is that the ML estimate $\mathbf{u}_{\Omega_{i,j}}$ corresponds to an endmember spectrum, we need to know with what confidence we can trust the fact that a single material is present in the corresponding local region. We propose a measure that represent the signal to noise ratio (SNR) between the abundance of the dominant material and the contribution of the other ones together with the noise. The ML estimate of the noise in the neighborhood $\Omega_{i,j}$ is:

$$\widehat{\sigma}^2_{\Omega_{i,j}} := \frac{1}{M} \operatorname{Tr}\left(\mathbf{C}_{\Omega_{i,j}}\right). \tag{3}$$

$\mathbf{C}_{\Omega_{i,j}}$ is the empirical covariance matrix defined by $\mathbf{C}_{\Omega_{i,j}} := \sum_{i',j'} \omega(i'-i, j'-j) \left(\mathbf{x}(i', j') - \mathbf{u}_{\Omega_{i,j}}\right) \cdot \left(\mathbf{x}(i', j') - \mathbf{u}_{\Omega_{i,j}}\right)^T$.

The proposed confidence measure is then defined by the SNR in the deciBel scale:

$$\widehat{\mathcal{T}}_{\Omega_{i,j}} := 10 \log_{10} \left(\frac{\|\mathbf{u}_{\Omega_{i,j}}\|_2^2}{\widehat{\sigma}^2_{\Omega_{i,j}}} \right). \tag{4}$$

3.3. Local spatial regions

So as to capture the possible shapes of the regions of the image where only one material is present, we define a family of neighborhood shapes Ω^s having weight ω_s and indexed by s. Figure 2 depicts the different neighborhood shapes used in our experiments, but other shapes are possibles.

The size of the neighborhood has to be small enough so that it exists regions of this size where only this material is present, and big enough so that the covariances matrices $\mathbf{C}_\Omega$ are estimated with enough samples. Finally, for each pixel (i, j) we merge the estimated confidence values $\widehat{\mathcal{T}}_{\Omega^s_{i,j}}$ and local endmembers estimates $\mathbf{u}_{\Omega^s_{i,j}}$ by keeping the one that had the largest confidence value, i.e.

$$\begin{aligned} s^*(i, j) &= \arg\max_s \widehat{\mathcal{T}}_{\Omega^s_{i,j}} \\ \mathbf{u}(i, j) &= \mathbf{u}_{\Omega^{s^*(i,j)}_{i,j}} \\ \widehat{\mathcal{T}}(i, j) &= \widehat{\mathcal{T}}_{\Omega^{s^*(i,j)}_{i,j}} \end{aligned}$$

Figure 3(b) shows the map of confidence values $\widehat{\mathcal{T}}(i, j)$ obtained from the hyperspectral image used in section 7 with an SNR of 50 dB.

[1] It is also possible to assume that more than one material can be present in each pixel.

Fig. 2. The 17 different neighborhood shapes Ω^s used in our experiments. The shades of gray represent the value of the weight $\omega_s(i,j)$, varying from white at the weakest intensity to black at the strongest.

4. CLUSTERING

We propose an iterative clustering algorithm which begins with an empty endmember matrix $\widehat{\mathbf{A}}_0 := []$ and estimates one new endmember at each iteration. The algorithm works in a similar way as the Automatic Target Generation Procedure (ATGP) [2] algorithm, which iteratively selects the largest magnitude vector $\mathbf{x}(i,j), \forall (i,j)$, and then projects the data in the orthogonal space spanned by the selected endmembers. In the contrary to the ATPG algorithm, the selection is done according to a weight $\pi_n(i,j)$ which is re-estimated at each iteration and which depends on the confidence value $\widehat{\mathcal{T}}(i,j)$.

Below, we detail each step of the algorithm at iteration n. At the first iteration the weights are the confidence values in the log scale $\pi_1(i,j) = \widehat{\mathcal{T}}(i,j)$.

1. Select the normalized vector $\widehat{\mathbf{v}}(i,j) := \widehat{\mathbf{u}}(i,j)/\|\widehat{\mathbf{u}}(i,j)\|_2$ having the largest weight $\pi_n(i,j), \forall (i,j)$.
2. Augment the endmembers matrix with the selected vector $\widehat{\mathbf{v}}(i,j)$: $\widehat{\mathbf{A}}_n := [\widehat{\mathbf{A}}_{n-1}, \widehat{\mathbf{v}}(i,j)]$.
3. Update the weight $\pi_n(i,j)$ with the rule $\pi_n(i,j) := \widehat{\mathcal{T}}(i,j)\|\mathbf{P}_n\widehat{\mathbf{v}}(i,j)\|_2$, where $\mathbf{P}_n = (\mathbf{I} - \widehat{\mathbf{A}}_n\widehat{\mathbf{A}}_n^{\dagger})$ is the operator that project the data $\widehat{\mathbf{v}}(i,j)$ in the orthogonal subspace spanned by the selected endmembers and where $\mathbf{A}^{\dagger} := (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$ is the Moore-Penrose pseudoinverse.

The algorithm stops after N iterations.

5. DIMENSION REDUCTION

Since it requires intensive computation to process HSI, and as the images are very redundant across the spectral dimension, dimension reduction (DR) like PCA is performed before the endmembers estimation step. So as to deal with the positivity constraint on the data, we might better use a Non-Negative PCA (NNPCA) [4] instead of the classical PCA.

6. ABUNDANCE MAPS ESTIMATION

Once the endmembers matrix is estimated, the abundance maps of each material $\mathbf{S}_n = [s_n(i,j)]_{i,j}$ can be estimated. We propose to use the binary masking technique [3]. The binary masking technique assumes that only one material is present in a pixel[2]. Thus, as we also have the constraint $s_n(i,j) \in [0,1]$ which can be imposed by a normalization step after the binary masking technique, the abundance maps estimation consists of a binary classification problem. Then our abundance maps estimation consists, for each pixel $\mathbf{x}(i,j)$, of setting to one the material n that maximizes the orthogonal projection of the estimated endmember spectrum $\widehat{\mathbf{a}}_n$ on $\mathbf{x}(i,j)$.

7. EXPERIMENT

So as to evaluate the performance of the proposed approach, we used a ground truth map image of size 640×152 of the Washington DC Mall shown in Figure 3(a) where each pixel corresponds to one material among six, and synthesize the HSI using (1). The endmember spectra $\mathbf{a}_n$ are choosen at random form the USGS digital spectral library[3]. We then add a Gaussian white noise such that the Signal to Noise Ratio (SNR) as defined in [10] vary from -5 dB to 50 dB. For each SNR level we generated 50 instances of the noise resulting in 50 instances of the HSI $\mathbf{X}$. We measure the performance of the tested algorithms in term of the mean, among these instances, of the accuracy which is defined by the proportion of true classification results (both true positives and true negatives) in the total number of pixels.

7.1. Tested algorithms

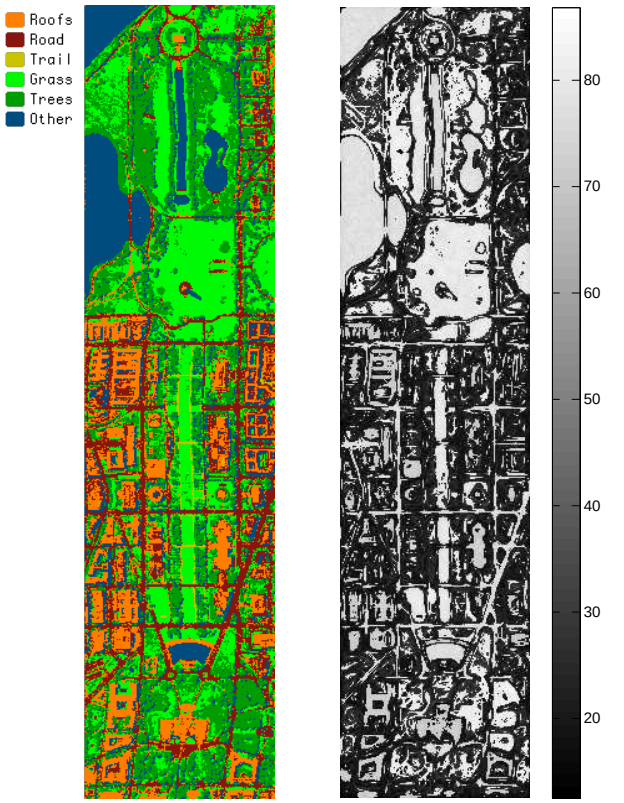
We test our Hyper-DEMIX algorithm with the ATPG method using a nonnegative least square approach (NNLS) [11] to estimate the abundance maps and the ICA-EEA algorithm with a normalization of the resulting abundance maps into the interval $[0,1]$. Both of these approaches are described in [10] where a PCA is used for the DR. We also tested the ATPG and the ICA-EEA approach with a NNPCA instead of the PCA.

7.2. Results

Results depicted in figure 4 show that the proposed Hyper-DEMIX method performs better than the other tested methods and that the classification task is perfect as soon as the SNR is higher than 15 dB. The results also show that it is more

[2]The binary masking techique is a very fast method used in audio source separation whereas the disjointness assumption is not perfectly valid. However the results are quite convincing. If we want to allow more than one active source per pixel, we can use another sparse approximation technique like the l_1 norm minimization [9].

[3]The USGS digital spectral library is available at the following URL: `http://speclab.cr.usgs.gov/spectral.lib06`.



(a) Ground truth. (b) Map of the confidence value $\widehat{\mathcal{T}}$.

Fig. 3. Map of the Washington DC Mall. SNR=50 dB.

convenient to use NNPCA instead of the traditional PCA for the DR. Indeed results of ICA-EEA are always better when using NNPCA instead of PCA, and results of ATPG using PCA are unstable whereas they monotically increase with the SNR when using NNPCA and perform perfect classification as soon as the SNR is higher than 30 dB. It is also interesting to note that the ICA-EEA method (with NNPCA) works better than ATPG (with NNPCA) when the SNR is lower than 20 dB, and vice versa.

8. CONCLUSION

We proposed a new method to unmix hyperspectral images which exploits the spatial structure of the material abundance map in the spatial domain. The proposed clustering algorithm uses a confidence measure and an orthogonal subspace approach to iteratively estimate the endmember spectra. The abundance maps of each material are then estimated using the binary masking technique. The results show that the proposed method outperforms state-of-the-art approaches whatever is the SNR. Further works include evaluation of the algorithm on real hyperspectral images and low spatial resolution images where multiple materials might be captured in a single pixel.

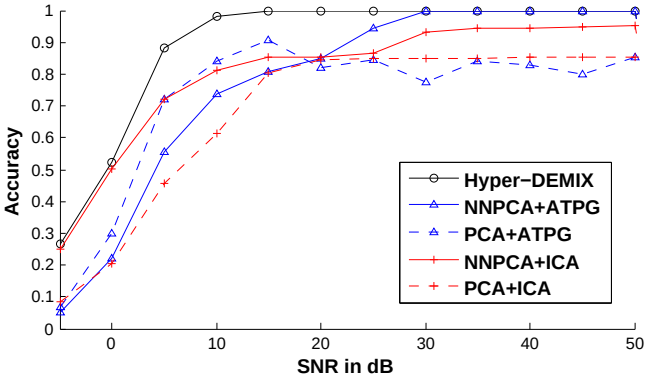


Fig. 4. Accuracy of the different unmixing algorithms.

9. REFERENCES

[1] J. Wang and C.I. Chang, "Applications of independent component analysis (ICA) in endmember extraction and abundance quantification for hyperspectral imagery," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 44, pp. 2601–2616, 2006.

[2] H. Ren and C.I. Chang, "Automatic spectral target recognition in hyperspectral imagery," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1232–1249, 2003.

[3] O. Yilmaz and S. Rickard, "Blind separation of speech mixtures via time-frequency masking," *IEEE Trans. on Sig. Proc.*, vol. 52, no. 7, pp. 1830–1847, July 2004.

[4] C.D. Sigg and J.M. Buhmann, "Expectation-maximization for sparse and non-negative PCA," in *ICML*. ACM, 2008, pp. 960–967.

[5] S. Arberet, R. Gribonval, and F. Bimbot, "A robust method to count and locate audio sources in a multichannel underdetermined mixture," *IEEE Transactions on Signal Processing*, vol. 58, no. 1, January 2010.

[6] JC Harsanyi and C.I. Chang, "Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach," *IEEE Trans. Geosci. and Remote Sens*, vol. 32, pp. 779–785, 1994.

[7] F. Abrard and Y. Deville, "Blind separation of dependent sources using the "time-frequency ratio of mixtures" approach," in *ISSPA*, July 2003.

[8] S. Arberet, R. Gribonval, and F. Bimbot, "A robust method to count and locate audio sources in a stereophonic linear instantaneous mixture," in *ICA*, 2006.

[9] P. Bofill and M. Zibulevsky, "Underdetermined blind source separation using sparse representations," *Signal Processing*, vol. 81, no. 11, 2001.

[10] Isaac D. Gerg, "An Evaluation of Three Endmember Extraction Algorithms: ATGP, ICA-EEA, AND VCA," M.S. thesis, PennState University, USA, May 2008.

[11] R. Bro and S. De Jong, "A fast non-negativity-constrained least squares algorithm," *Journal of Chemometrics*, vol. 11, no. 5, pp. 393–401, 1997.