POTENTIAL AND LIMITATIONS OF BAND SELECTION AND LIBRARY PRUNING IN SPARSE HYPERSPECTRAL UNMIXING

Marian-Daniel Iordache[1], *José M. Bioucas-Dias*[2], *Antonio Plaza*[3]

[1] Flemish Institute for Technological Research (VITO), Boeretang 200, BE-2400 Mol, Belgium
[2]Instituto de Telecomunicações and Instituto Superior Técnico,
Universidade de Lisboa, 1049-001, Lisbon, Portugal
[3]Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications,
University of Extremadura, E-10071 Cáceres, Spain

ABSTRACT

Sparse regression using spectral libraries is nowadays a widely used technique for hyperspectral data unmixing. Assuming that the potential endmembers are collected in a large database of spectra, sparse unmixing finds the fractional abundances of a reduced set of constituent materials by solving convex optimization problems which target, at the same time, low data reconstruction errors. The large amount of data, jointly with other limitations related to the internal characteristics of the large spectral libraries (such as spectra similarity), affect the performance of sparse unmixing algorithms in terms of accuracy and running time. Recently, an efficient method based on a multi-measurement vector (MMV) approach was proposed to select, from a large library, suitable spectra for unmixing. Many research efforts have also been devoted to data dimensionality reduction techniques, of which band selection is very popular. In this work, we investigate the effects on the unmixing performance when the two techniques are applied simultaneously. Our experiments show that important improvements can be achieved, depending on several factors, such as: data noise, number of endmembers present in the dataset, number of spectra retained from the library, the order of operations, accuracy of data subspace estimation.

***Index Terms*—** MMV, sparse unmixing, band selection, spectral libraries, data collaborativity

1. INTRODUCTION

In hyperspectral unmixing, the aim is to find the constituent materials of a scene, jointly with their corresponding spectral signatures and fractional abundances (or areas occupied in each pixel by the endmembers) [1, 2]. Most of pixels in hyperspectral scenes are mixed, *i.e.*, they contain more than one material. Sparse unmixing [3, 4] was proposed as an alternative to the classical approach relying on endmember extraction, in order to overcome the lack of pure pixels in the data. It assumes that the endmembers are present in a large collection of pure spectra, called spectral library. Convex optimization problems are employed to infer which library members are present in the scene (the endmembers) and what are their fractional abundances in each pixel. It was shown that one of the limitations acting on the performances of sparse unmixing is the size of the spectral libraries, as an increasing number of library spectra leads to lower unmixing performances. To overcome this difficulty, a method to select suitable spectra for unmixing was described in [5]. Based on the MMV approach ((see [6–8] and references therein)), this method is able to retain a reduced set of endmembers to be used in unmixing. Then, it applies a so-called collaborative sparse regression (CSR) to infer fractional abundances.

Apart from the fractional abundances inference, CSR has much more applicability in the hyperspectral data processing. In [9], we have shown that a complete hyperspectral unmixing chain (including endmember extraction, band selection and inversion) can be compiled based on an unique CSR optimization function. However, while it was shown that hyperspectral datacubes can be reconstructed with high accuracy from a reduced set of spectral band using appropriate coefficients, the impact of the band selection on the unmixing output was not investigated. This paper exploits a combination of band selection and library spectra selection in order to analyze the quality of the unmixing results in this scenario. Several questions are to be taken into account:

- should band selection be performed before or after library member selection?
- what is the influence of noise when band selection and library pruning are simultaneously used in unmixing?
- how many spectral bands can be dropped from the dataset without significantly degrading the unmixing quality?
- what other factors influence the unmixing performance?

The remainder of the paper is organized as follows. Section 2 presents the employed methodology. In Section 3, experiments with simulated data are presented. Section 4 concludes the paper with observations and conclusions on the presented work.

2. PROPOSED UNMIXING METHODOLOGY WITH BAND SELECTION

Let $\mathbf{Y}$ and $\mathbf{A}$ be the observed dataset and the available spectral library, respectively. Let L denote the number of spectral bands, n – the number of pixels in $\mathbf{Y}$ (each column contains one pixel) and m – the number of library spectra stored in $\mathbf{A}$. Under the linear mixing model [2], the observed data $\mathbf{Y}$ can be expressed as a linear combination of spectra from $\mathbf{A}$ as follows:

$$\mathbf{Y} = \mathbf{AX} + \mathbf{N}, \tag{1}$$

where $\mathbf{X} := [\mathbf{x}_1, \ldots, \mathbf{x}_n]$ is the abundance fraction matrix and $\mathbf{N} := [\mathbf{n}_1, \ldots, \mathbf{n}_n]$ is the noise matrix. Because the abundance fractions are nonnegative and sum to one in each pixel, the constraints

This work has been supported by the Portuguese Science and Technology Foundation under Projects UID/EEA/50008/2013 and PTDC/EEI-PRO/1470/2012.

$\mathbf{X} \geq \mathbf{0}$, to be understood in the component-wise sense, and $\mathbf{1}_m^T \mathbf{x_i} = 1$ ($\mathbf{1}_m$ stands for a column vector with m ones; $i = 1, \ldots, n$) called abundance non-negativity constraint (ANC) and abundance sum-to-one constraint (ASC), respectively, are often imposed into the model (1). In this work, we disregard ASC and only consider ANC.

2.1. Band selection using collaborative sparse regression (CSR)

Given that in a given hyperspectral dataset the number of active materials is usually small compared with the number of columns of $\mathbf{A}$ [2], then the matrix of fractional abundances $\mathbf{X}$ is row-wise sparse, *i.e* it has many zero rows. A plethora of well-established methods are available in the literature to perform unmixing taking sparsity into account, considering per–pixel processing, spatial regularization, structured sparsity or data collaborativity, among others. In this work, we use the *Collaborative Sparse Unmixing via variable Splitting and Augmented Lagrangian* (CSUnSAL) algorithm [10], which takes into account the fact that all pixels share a common set of pixels, thus they should collaborate together in inferring an unmixing solution. CLSUnSAL solves the following collaborative sparse regression (CSR) optimization problem:

$$\begin{aligned} \min_{\mathbf{X}} \quad & \|\mathbf{AX} - \mathbf{Y}\|_F^2 + \lambda \sum_{k=1}^{m} \|\mathbf{x}^k\|_2 \qquad (2) \\ \text{subject to:} \quad & \mathbf{X} \geq 0. \end{aligned}$$

where $\mathbf{x}^k$ denotes the k-th line of $\mathbf{X}$ and λ is a regularization parameter which weights the two terms of the objective function. The convex term $\sum_{k=1}^{m} \|\mathbf{x}^k\|_2$ is the so-called $\ell_{2,1}$ mixed norm which promotes sparsity among the rows of $\mathbf{X}$, *i.e.*, it promotes solutions of (1) with small number of nonzero lines of $\mathbf{X}$. In all cases presented in this work, the reported unmixing performances correspond to the optimal parameter λ, out of the following values: 0 (no sparsity imposed; equivalent to the classical *non-negative least squares* solution), 0.01, 0.03, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5 and 0.6.

Note that CSR, solved by CLSUnSAL, apart from the fractional abundances retrieval (the so-called *inversion* step in unmixing), can serve as an endmember extraction algorithm by simply replacing the library $\mathbf{A}$ with $\mathbf{Y}$ in (2); also, a band selection algorithm is obtained if $\mathbf{A}$ is replaced with $\mathbf{Y}^T$ (the transpose of $\mathbf{Y}$) [9]. We will exploit this possibility in our experiments, in which the number of selected spectral bands will be varied.

2.2. Library pruning using an MMV approach

We defined, in the previous subsection, a way to perform band selection. We now shortly describe the methodology to perform library pruning.

In our previous work [5], the library pruning was established as part of a complete unmixing algorithm called MUSIC-CSR. Here, we use the same approach, based on the *MUltiple SIgnal Classification* (MUSIC) algorithm [11, 12], which consists in the following steps: 1) *Signal Subspace Identification*: Infers the subspace in which the hyperspectral data $\mathbf{Y}$ lives using the HySime algorithm [13]; 2) *Projection errors*: Computes the Euclidean distance from each member of the library to the estimated subspace; 3) *Active set detection*: Sorts the normalized projection errors by increasing order and retain the indexes of first k_f in the set of m library members.

The logic behind these operations is the following: if one library member is a data endmember, it should lie in the data subspace, such that the distance to this subspace is ideally zero. Otherwise, it should be placed at a non-zero distance from the data subspace. In real applications, due to noise and non-linearities, it is unlikely to obtain distances equal to zero. However, the closest library members to the inferred subspace should still be endmembers with a higher probability than the others. In this work, we run our experiments for different sets of retained spectra: $k_f = k$, $k_f = k + 10$, $k_f = k + 20$ and $k_f = k + 30$, where k is the true number of endmembers in the considered dataset.

The MUSIC-CSR algorithm uses the retained library members as input to CLSUnSAL. Here, the same concept is used. The difference between the analysis performed in [5] and this work consists in the fact that, here, a band selection is performed such that not all L spectral bands are used in unmixing. The goal is to identify the advantages and weaknesses that such a band reduction might introduce into the unmixing accuracy. In all cases, the unmixing performance is measured by the so-called *signal-to-reconstruction error*: SRE $\equiv E[\|\mathbf{x}\|_2^2]/E[\|\mathbf{x} - \widehat{\mathbf{x}}\|_2^2]$, expressed in dB: SRE(dB) $\equiv 10 \log_{10}(\text{SRE})$, where $\mathbf{x}$ is the true unmixing solution and $\widehat{\mathbf{x}}$ is the estimated one.

3. EXPERIMENTAL RESULTS

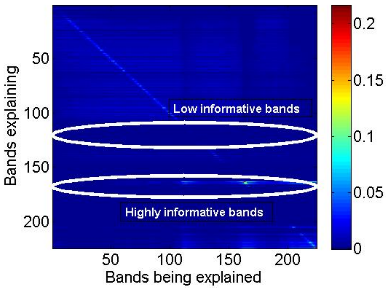
The proposed methodology was tested on various datacubes containing 4000 pixels, with the number of endmembers k_f being 4, 8 or 12. The data sets were generated using spectra from a library $\mathbf{A}$ containing 240 spectral signatures with 224 bands from the USGS spectral library available online[1]. The endmembers were randomly selected and all datacubes were contaminated with i.i.d. Gaussian noise for three signal-to-noise (SNR) levels: 20, 30 and 40dB. It should be noted that most hyperspectral sensors acquire data with SNR larger than 30dB, which makes the last two cases more appropriate for a real case analysis.

The band selection is performed using CLSUnSAL with $\mathbf{Y}^T$ as input matrix (as explained in Subsection 2.1). One example of coefficients matrix is shown in Fig. 1.a). In this particular case, the parameter λ was set to a high value ($\lambda = 30$). Such a matrix should give a response to the question: what are the most significant spectral bands? or, in other words: what are the spectral bands which best explain the observed data? The informative bands should have large coefficients, while the non-informative ones should have low coefficients. The bands with the lowest contribution to the matrix of coefficients (thus, to the observed data) can be dropped. In this work, the number of bands retained in the experiments varies between 50 and 200, with an increment of 30. Experiments with the full data (224 bands) were also performed for comparison purposes.

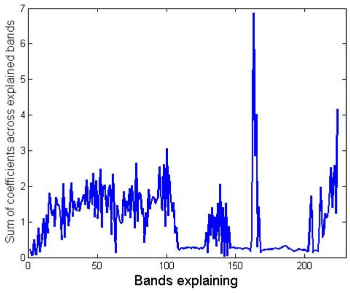
In Fig. 1.a), note the *striped* pattern of the coefficients. Two sets of bands are highlighted to make a distinction between informative and non-informative bands. While the most discriminative bands spread their non-zero coefficients horizontally (they explain other bands following a linear model), some of the others are vanishing (their cofficients ar almost zero). In Fig. 1.b), we plot the sum of coefficients for all bands. These coefficients are then arranged in increasing order and the bands corresponding to the highest values are selected.

The first issue to analyze is the order of the operations. In Fig. 2, we plot the SRE(dB) values for the same dataset ($k = 8$, SNR=30dB) in two situations – the band selection is performed before and after library pruning, respectively. Note that the SRE(dB) values in the latter case are clearly higher and more stable, although the differences are minor if the number of retained bands is large

[1]See: http://speclab.cr.usgs.gov/spectral.lib06

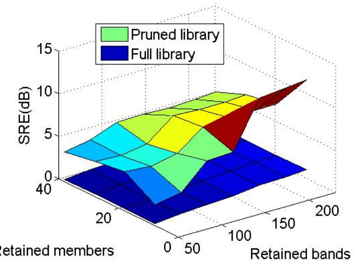


a) Matrix of coefficients

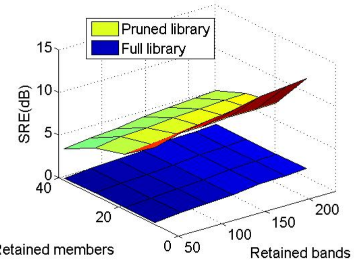


b) Contribution of individual bands to the data

Fig. 1. The use of matrix of coefficients for band selection.



(a) Unmixing with band selection before library pruning



(b) Unmixing with band selection after library pruning

Fig. 2. Influence of the order of operations on the unmixing results.

and the number of retained members is small. This behavior is common to all the tested datasets, thus we can conclude that the band selection should always be performed after the library pruning step. All the following results are reported folowing this rule.

Another issue to be analyzed is the unmixing performance w.r.t. the data noise. In Fig. 3, we compare the unmixing performances for the three noise levels for the dataset containing $k = 4$ endmembers. In this plot, one horizontal axis shows the number of retained library members, while the other marks the number of spectral bands retained.

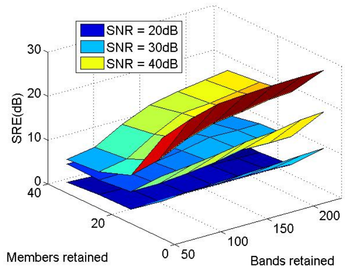
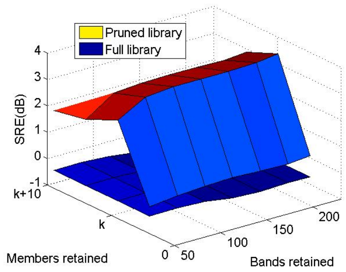


Fig. 3. Noise influence on the unmixing with band selection ($k = 4$).

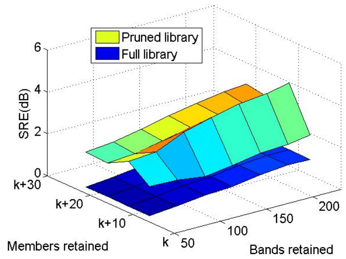
It is easy to notice, from the plots displayed so far, that higher performances can be achieved when a minimal number of library members are retained. Ideally, the k library members with lowest projection errors should be the endmembers themselves. However, this is not always the case. In high noise, or when the number of endmembers is high, retaining form the library a number of members equal to the number of endmembers can be a risky approach. This might be due to several reasons: nonlinearity of the data, incorrect estimation of the number of endmembers, incorrect inference of the data subspace etc. In Fig. 4, we illustrate this drawback for two representative cases: high data noise (SNR=20dB and $k = 4$; see Fig. 4.a) and high number of endmembers (SNR=30dB and $k = 12$; see Fig. 4.b). For illustration purposes, the SRE(dB) obtained with the full library is also displayed, thus the y-axis should not be taken into account in this case. Note that when $k_f = k$, even if a large number of bands is retained, the unmixing performance degrades, compared to the cases when $k_f > k$. This indicates that at least one endmember was not correctly identified. However, a slight improvement over the performance obtained with the full library is still visible and, fortunately, the case illustrated in Fig. 4.a (very high noise) is unlikely to happen in real scenarios.

The last open question is how low can be the number of selected bands such that the unmixing performance stays reasonable. We take the performance level SRE(dB)=5 as a reference to indicate if a solution is useful or not [4]. It should be noted that the unmixing using the complete full library reached this performance only in very advantageous conditions: low number of endmembers ($k = 4$) and low noise (SNR=40dB). For the very high noise scenario (SNR=20dB), none of the tested combinations reached the targeted performance, except for the situaton in which the number of endmembers is $k = 4$, the number of retained library spectra is $k_f = k = 4$ and the number of retained bands is 50. It means that, even in this scenario, useful solutions can be obtained when a low number of spectral bands is selected. However, even if the unmixing with band selection and library pruning improved the unmixing compared to the experiments with full library, this situation should still be treated with caution, due to the drawbacks mentioned before (the probability to select an incorrect set of endmembers is higher).

Table 3 shows the minimum number of bands needed to reach useful unmixing performance for all tested unmixing settings when SNR=30dB and SNR=40dB. The sign † indicates that the unmixing did not reach the targeted performance, no matter the number of retained bands. Note that, under mild assumptions, high unmixing performances can be achieved by band selection and library pruning while keeping a very low number of bands (50 bands represent less



(a) High data noise (SNR=20dB)



(b) Large number of constituent endmembers (SNR=30dB)

Fig. 4. Risks affecting the performance of sparse unmixing with band selection.

than 23% of the original datacube).

Table 1. Minimum number of retained bands to achieve unmixing performance higher than SRE(dB)=5

	SNR=30dB			SNR=40dB		
k	4	8	12	4	8	12
$k_f = k$	50	†	†	50	†	50
$k_f = k + 10$	50	50	†	50	50	110
$k_f = k + 20$	50	50	†	50	50	110
$k_f = k + 30$	50	50	†	50	50	170
$k_f = m$	†	†	†	110	†	†

4. CONCLUSIONS

In this paper, we investigated the performance of sparse hyperspectral unmixing when band selection and library pruning are performed simultaneously as part of the unmixing procedure. Our experiments show that, no matter the characteristics of the data, improvements are achieved compared to the results obtained using the complete full spectral library as input. It was demonstrated that band selection should be performed after library pruning for stable results. As expected, the data noise influences the unmixing quality: higher noise implies lower performance. On another hand, lower noise leads to larger improvements when the anlyzed strategy is employed. A very interesting observation is that caution should be taken during the library pruning when the data is affected by high level of noise and/or the number of endmembers is high. In these cases, the number of retained library spectra should be larger than the number of endmembers in the image, as several factors might influence the correct identification of the endmembers: data nonlinearities, errors in data subspace estimation, errors in estimating the number of endmembers. It was shown that useful unmixing results can be obtained in regular datasets (*e.g.*, the noise is not very high and the number of endmembers not very large, which is a situation often encountered in hyperspectral data) even for a very low number of spectral bands retained (a number of bands lower than than 23% of the number of original bands was tested in our experiments).

5. REFERENCES

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

[2] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical and sparse regression-based approaches," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 354–379, 2011.

[3] J. Bioucas-Dias and M. Figueiredo, "Alternating direction algorithms for constrained sparse regression: application to hyperspectral unmixing," *2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, vol. 1, pp. 1–4, Reykjavik, Iceland, 2010.

[4] M.-D. Iordache, J. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 6, pp. 2014–2039, 2011.

[5] M.-D. Iordache, J. M. Bioucas-Dias, A. Plaza, and B. Somers, "Music-csr: Hyperspectral unmixing via multiple signal classification and collaborative sparse regression," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 7, pp. 4364–4382, July 2011.

[6] J. M. Kim, O. K. Lee, and J. C. Ye, "Compressive music with optimized partial support for joint sparse recovery," *arXiv:1102.3288v2*, June 2011.

[7] J. M. Kim, O. K. Lee, and J. C. Ye, "Compressive music: A missing link between compressive sensing and array signal processing," *arXiv:1004.4398v5*, June 2011.

[8] Y. Eldar and H. Rauhut, "Average case analysis of multichannel sparse recovery using convex relaxation," *IEEE Transactions on Information Theory*, vol. 56, no. 1, pp. 505–519, 2010.

[9] M.-D. Iordache, A. Okujeni, S. van der Linden, J. M. Bioucas-Dias, A. Plaza, and B. Somers, "On the use of collaborative sparse regression in hyperspectral unmixing chains," *Proceedings of 6^{t}h IEEE GRSS Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pp. 1–4, Lausanne, Switzerland, 2014.

[10] M.-D. Iordache, J. Bioucas-Dias, and A. Plaza, "Collaborative sparse regression for hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 1, pp. 341–354, January 2014.

[11] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, 1986.

[12] G. Bienvenu and L. Kopp, "Adaptivity to background noise spatial coherence for high resolution passive methods," in *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP'80.* IEEE, 1980, vol. 5, pp. 307–310.

[13] J. Bioucas-Dias and J. Nascimento, "Hyperspectral subspace identification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 46, no. 8, pp. 2435–2445, 2008.