



**HAL**  
open science

## Spectral partitioning and fusion techniques for hyperspectral data classification and unmixing

Rita Ammanouil, Jean Abou Melhem, Joumana Farah, Paul Honeine

► **To cite this version:**

Rita Ammanouil, Jean Abou Melhem, Joumana Farah, Paul Honeine. Spectral partitioning and fusion techniques for hyperspectral data classification and unmixing. Proc. 6th International Symposium on Communications, Control, and Signal Processing (ISCCSP), 2014, Athens, Greece. pp.550-553, 10.1109/ISCCSP.2014.6877934 . hal-01965992

**HAL Id: hal-01965992**

**<https://hal.science/hal-01965992v1>**

Submitted on 27 Dec 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# SPECTRAL PARTITIONING AND FUSION TECHNIQUES FOR HYPERSPECTRAL DATA CLASSIFICATION AND UNMIXING

Rita Ammanouil<sup>1</sup>, Jean Abou Melhem<sup>1</sup>, Joumana Farah<sup>1</sup>, Paul Honeine<sup>2</sup>

<sup>1</sup>Telecommunications department, Faculty of Engineering, Holy-Spirit University of Kaslik, Lebanon

<sup>2</sup>Institut Charles Delaunay (CNRS), Université de technologie de Troyes, France

## ABSTRACT

Abstract — Hyperspectral images are characterized by their large contiguous set of wavelengths. Therefore, it is possible to benefit from this ‘hyper’ spectral information in order to reduce the classification and unmixing errors. For this reason, we propose new classification and unmixing techniques that take into account the correlation between successive spectral bands, by dividing the spectrum into non-overlapping subsets of correlated bands. Afterwards, classification and unmixing are performed on each subset separately, such as to yield several labels per pixel in the classification case, or abundances in the unmixing case. Then, several fusion techniques are proposed to obtain the final decision. Results show that spectral partitioning and appropriate fusion allow a significant gain in performance compared to previous classification and unmixing techniques.

*Index Terms*— Hyperspectral imaging, spectral pre-processing, classification, unmixing, fusion.

## 1. INTRODUCTION

Hyperspectral imaging is based on the acquisition of a scene in a large number of wavelengths, resulting in a data cube with two spatial dimensions and a spectral one. The large amount of information in these images makes them a powerful tool for surface materials’ detection in remote sensing. This detection can be performed either by classifying each pixel to a specific class, or by decomposing a pixel into a mixture of several pure classes [3][4]. One of the standard tools for classification is Support Vector Machines (SVMs), and a study of their effectiveness for hyperspectral data was proved in [1]. In this paper, we use the One Least Square Machine ‘One LSM’ [2], a variant of SVM with less computational complexity. On the other hand, several techniques were proposed for the identification of pure classes (also known as ‘endmembers’), like the well-known NFINDR [5] and SGA (Simplex Growing Algorithm) [6] techniques. After this identification, the abundances of endmembers are computed using unconstrained, equality constrained, or fully constrained approaches [7], [8]. Other techniques allow a joint estimation of endmembers and abundances, such as the Barycentric approach introduced in [9].

In both classification and unmixing contexts,

performance can be enhanced using spatial, spectral, or joint spatial-spectral pre-processing. Spatial pre-processing aims at exploiting the strong correlation between neighboring pixels through, for example, mathematical morphology operators, as was done in [10]. On the other hand, spectral pre-processing allows taking advantage of the correlation between neighboring spectral bands, as was done in [11] to perform clustering.

In this paper, we propose to divide the original wavelength interval into a set of correlated intervals. We refer to this process as sub-band partitioning. Afterwards, we perform the classification or the unmixing on each sub-interval and use this supplementary information in order to enhance the original result. However, the fusion of the obtained results is a challenging problem that we address through several fusion techniques. The rest of the paper is organized as follows: Section 2 discusses the characteristics of the spectral correlation matrix. In Section 3, we show how sub-band partitioning can be applied using the One LSM classifier, and we propose a distance-based fusion technique for this classifier. Then, sub-band partitioning is applied for spectral unmixing, along with several proposed metrics for the weighted fusion. Experimental results are provided and analyzed in Section 4. Finally, Section 5 concludes our work.

## 2. SPECTRAL CORRELATION AND SUB-BAND PARTITIONING

The hyperspectral image is modeled by a matrix  $X = [x_1 \dots x_n]$  of size  $p \times n$ , a set of  $n$  observations (one for each pixel) with  $p$  spectral components each. The inter-band correlation matrix  $C_x = XX^T$  estimates the pairwise normalized correlation coefficients between each pair of columns in the matrix  $X^T$ .  $C_x$  is a square symmetric matrix, with size  $(p \times p)$  and diagonal elements equal to 1. The off-diagonal elements  $c_{ij}$  express the correlation between the spectral pair  $(i, j)$ . Large (resp. small) elements of  $C_x$  denote a high (resp. low) correlation between the two bands. Fig. 1 shows the inter-band correlation matrix obtained with the Pavia University, with size  $(103 \times 103)$ . Light pixels, close to 1, denote a high correlation, whereas dark ones indicate little correlation between the two corresponding bands. From Fig. 1, we see that the inter-band correlation matrix generally exhibits clusters. In this example, it underlies two white

squares, which implies that the bands from 1 to 70 are highly correlated, and the same goes for the bands from 71 to 103. Similar results were observed with the Indian Pines', and Cuprite's correlation matrices.

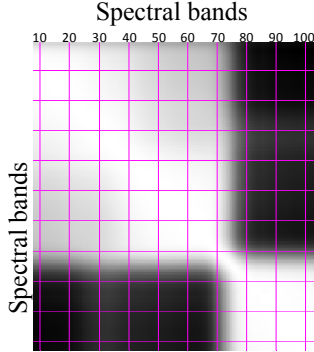


Figure 1: Inter-band correlation matrix of the Pavia University

### 3. FUSION TECHNIQUES FOR CLASSIFICATION AND UNMIXING

Given that the spectral bands can be divided into a set of correlated intervals, we propose to add new stages for the classification and unmixing procedures. Instead of only performing the classification, respectively the unmixing, on the whole spectral interval, we suggest to additionally perform the classification on each sub-band interval. These so called sub-classifications can be used to correct the potential mistakes of the original classifier.

The proposed scheme (Fig. 2) consists of four phases. The first step is the estimation of the inter-band correlation. Afterwards, we divide the bands into correlated subsets, each subset containing the same pixels but with a different band margin. Then, classification is performed using the original cube, as well as each subset separately. Finally, via an adequate fusion function, we fuse all obtained results.

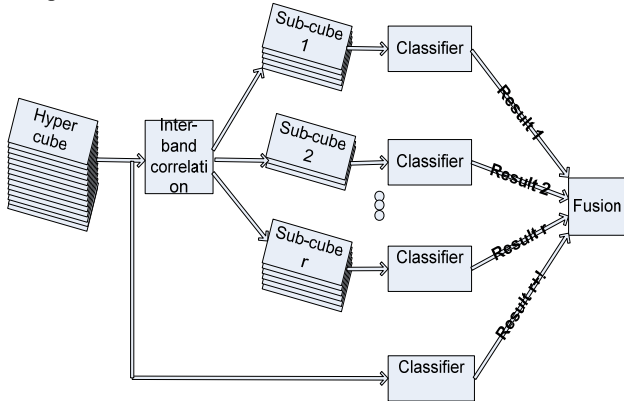


Figure 1: Classification with sub-band partitioning

Note that this scheme can be applied for both classification and unmixing, i.e., the input of the fusion function can be either the class of a pixel in the case of classification, or the abundances in the case of unmixing. However, the main challenge in this scheme is to find the optimal fusion function with regards to the material detection

approach in use. In Section 3.1, we will provide an application of this scheme using the One LSM classifier, and, in Section 3.2, we investigate four unmixing algorithms. In each section, we will propose adequate fusion techniques to deduce the final decision.

#### 3.1. The classification case

The One LSM algorithm, recently proposed in [2], is used to solve a supervised learning classification problem. Most importantly, the complexity of the algorithm in the multiclass mode is lower than other algorithms reported in the literature such as SVM. The basic setting of a binary supervised learning classification problem is as follows. Given a training set  $S = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$  constituted by  $n$  observations  $x_i$  with their respective classes  $y_i = \pm 1$ , the algorithm should be able to find a decision function  $f$  such that:

$$\begin{cases} f(x) > 0, y = 1 \\ f(x) < 0, y = -1 \end{cases}$$

Once the decision function  $f$  has been determined based on the training set, it is used to estimate the class membership of any new observation. Equivalently, for any observation  $x$ , we can write:  $y = \arg \max_{y=\pm 1} yf(x)$  (1)

Note that (1) is given here in its binary context, with a number of classes  $m=2$ , but could be easily extended to the multiclass case where 'y' takes 'm' possible values. On the other hand, when using the One LSM classifier in the proposed scheme, as shown in Fig.2, we will have  $r+1$  classes for every pixel prior to the fusion step:  $r$  decisions from the  $r$  sub-cubes' classification and 1 decision from the original cube's classification. For the fusion of these results, we tested two voting techniques.

**1- Absolute Sub-Vote (ASV):** In this voting scheme, when we have an absolute majority among the sub-votes (from the sub-classifiers), the winning class is the one agreed on by the sub-classifiers. Otherwise, the class proposed by the original classifier remains unchanged.

**2- Distance-Based Vote (DBV):** In this voting strategy, a soft decision is taken, i.e., the value of the maximum product in (1) is taken into account, rather than only considering its argument. While collecting the votes, the sub-classifiers' and main classifier's votes are multiplied by their corresponding weight  $d_{res}$  that allows taking into account the confidence of the classifier's decision:

$$d_{res} = y_{res} \times f(x), \text{ with } y_{res} = \arg \max_{y=\pm 1} yf(x).$$

#### 3.2. The unmixing case

In the unmixing context, the pixel's spectrum is a mixture of pure material's spectra, known as endmembers. The first step in unmixing is to identify the endmembers present in the scene, followed by the abundances' calculation, i.e. determining the contribution of each endmember in a pixel. The abundances should be positive, with a sum equal to one. Let  $x$  be a pixel's spectrum,  $x_e$  an endmember's spectrum,

and  $\alpha$  its corresponding abundance in the pixel's spectrum. Assuming that the data has a linear mixture model, and considering  $q$  endmembers in the scene, we can write a

pixel's spectrum as:  $x = \sum_{i=1}^q \alpha_i x_{e,i}$ , with  $\sum_{i=1}^q \alpha_i = 1$  and

$$\alpha_i \geq 0, i = 1, \dots, q.$$

The problem of abundances' calculation has been addressed using different algorithms in the literature [7], [8], [9], like the unconstrained, equality constrained, fully constrained, and barycentric approaches, that mainly differ by the constraints they consider for this calculation. The endmembers' abundances constitute the inputs to the fusion function of Fig.2, in the unmixing case. We propose the following fusion function consisting of two steps:

1. Fuse the  $r$  results of the  $r$  sub-cubes.
2. Choose between the abundances obtained from the original data (input  $r+1$ ) and those obtained from step 1.

In step 1, the fusion consists of a weighted average of the estimated sub-abundances, with a weight proportional or inversely proportional to a certain metric. Let

$a_i = [\alpha_i^1 \ \dots \ \alpha_i^q]^T$ ,  $i = 1 \ \dots \ r$ , be the vectors of estimated sub-abundances, and  $d_i$ ,  $i = 1, \dots, r$ , the metric values. Let  $x_i = [x_i^1 \ \dots \ x_i^p]^T$  be the original pixel's spectrum and  $\hat{x}_i$  its reconstructed version using the estimated abundances. The final abundances from step 1 are taken as:

$$a = \sum_{i=1}^r \left( \frac{1}{d_i} a_i \right) / \sum_{i=1}^r \frac{1}{d_i} \quad (2)$$

We proposed and studied four different distance metrics for this fusion:

**1- Simple averaging (AVG):** In this case, all sub-results are given the same weight  $d_i = 1, \forall i$ .

**2- Averaging based on the reconstruction angle (ANG),** i.e., the angle difference between the reconstructed and original pixels' spectra:  $d_i = \cos^{-1}(\langle x_i, \hat{x}_i \rangle / \|x_i\| \cdot \|\hat{x}_i\|)$ .

**3- Averaging based on the Mean Square Error (MSE):**

$$d_i = \frac{1}{p} \sum_{j=1}^p (\hat{x}_i^j - x_i^j)^2.$$

**4- Averaging based on the Mean Absolute Difference**

$$\text{(MAD): } d_i = \frac{1}{p} \sum_{j=1}^p |\hat{x}_i^j - x_i^j|.$$

In step 2, we need to evaluate the accuracy of the abundances obtained from step 1 and those given by the original hyperspectral cube, in order to choose the best candidate. For this purpose, we compute the spectral angle of the reconstructed pixel, for each one of the two results, and choose the positive abundances with the smallest

$$\text{spectral angle: } \theta(x, \hat{x}) = \cos^{-1}(\langle x, \hat{x} \rangle / \|x\| \cdot \|\hat{x}\|). \quad (3)$$

## 4. EXPERIMENTAL RESULTS

### 4.1. Classification

The proposed spectral pre-processing techniques are applied to the Pavia University data set, which has a total of 103 bands and 9 classes of materials. We compare the performance of the proposed scheme with respect to the simple classification in terms of the overall error percentage. In the aim of generalizing the results, we use the 5-fold cross validation approach [12]. According to the study in Section 2, the spectral band is partitioned into 2 intervals: [1 70] and [71 103], which accounts for two sub-classifiers ( $r=2$ ). Moreover, at the classification level, in Fig. 2, we apply two dimensionality reduction techniques, PCA (Principal Component Analysis) and Gaussian KPCA (Kernel PCA), for which we respectively consider 150 and 50 pixels per class. The number of retained bands  $L$  tested with each technique is taken in the set  $\{3, 10, 30\}$ . Table 1 shows the overall error percentage obtained with the classical classification (with no partitioning), and with the proposed scheme, using the two proposed fusion techniques ASV and DBV. Since our scheme introduces additional stages compared to the classical classification, we also show the execution time of each method, i.e., the time consumption necessary to evaluate the classes for the whole data set (for the case without partitioning), or for the  $r+1$  evaluations as well as their fusion (with partitioning). In general, sub-band partitioning and fusion allow reducing the error percentage, especially with applied with PCA and the DBV technique, with a slight increase in the execution time. With Gaussian KPCA, ASV has better performance. However, in both cases, when the number of retained bands is relatively high ( $L=30$ ), the performance gain is minor.

Table 1: Overall error percentage (%) and time consumption (s) for the One LSM classifier, obtained with the Pavia University.

		No partitioning	ASV	DBV
PCA	$L=3$	48.370 % (0.333 s)	42.148 % (0.887 s)	41.556 % (0.890 s)
	$L=10$	2.593 % (0.360 s)	2.519 % (1.013 s)	2.074 % (1.022 s)
	$L=30$	0.741 % (0.360 s)	0.519 % (1.045 s)	0.519 % (1.040 s)
Gaussian KPCA	$L=3$	18.889 % (0.430 s)	17.778 % (0.678 s)	23.333 % (0.681 s)
	$L=10$	1.556 % (0.511 s)	1.111 % (0.787 s)	1.333 % (0.787 s)
	$L=30$	0.667 % (0.535 s)	0.222 % (0.802 s)	0.222 % (0.805 s)

### 4.2. Unmixing

For the unmixing experiments, we use the Cuprite scene 1 which, after removing the water absorption bands, has 188 bands and a total of 47750 pixels. The proposed fusion schemes are tested with the following spectral interval

partitions: [1 101], [102 135], and [132 188]. In addition to the time consumption, the following metrics are used to evaluate the performance of the proposed unmixing techniques:

- Negativity Percentage (**NP**): The percentage of pixels having at least one negative abundance value.
- No Equality Percentage (**NEP**): The percentage of pixels whose abundances' sum is different than one.
- Average Spectral Angle (**ASA**): Average angle difference between the reconstructed and original images, obtained by averaging the pixels' reconstruction angles, estimated by (3).

Table 2 compares the performance of the Basic unmixing scheme (without partitioning) towards the three proposed fusion techniques, using four algorithms for the abundances' estimation: unconstrained, equality constrained, fully constrained, and barycentric. In general, the proposed scheme yields better results compared to the basic unmixing. In particular, when applied with the unconstrained algorithm, our fusion techniques were able to reduce the Negativity Percentage (NP) from 90% to as low as 3.89% with the AVG fusion. Similarly, with the equality constrained approach, the NP was reduced from 29% down to 0.791% with AVG. On the other hand, no important improvement was achieved with the fully constrained approach since it already has an NP of 0%.

Table 2: NP, NEP, ASA, and T for unmixing, obtained with the Cuprite scene 1.

		NP (%)	NEP (%)	ASA (%)	T(s)
Un-constrained	Basic	90.18	99.99	0.0497	0.0780
	AVG	3.89	99.69	0.0962	1.2010
	ANGLE	25.60	99.88	0.0777	2.6360
	MSE	39.44	99.90	0.0745	2.2620
	MAD	31.86	99.91	0.0762	2.2150
Equality constrained	Basic	29.31	0	0.0841	0.0780
	AVG	0.791	0	0.0975	1.9340
	ANGLE	0.896	0	0.0969	2.8860
	MSE	1.038	0	0.0958	2.4800
	MAD	0.902	0	0.0965	2.4810
Fully constrained	Basic	0	0	0.0867	6.2090
	AVG	0	0	0.0867	29.593
	ANGLE	0	0	0.0867	31.294
	MSE	0	0	0.0866	31.310
	MAD	0	0	0.0867	30.826
Barycentric	Basic	30.54	0	0.0849	1.3120
	AVG	1.327	0	0.1000	5.9800
	ANGLE	0.938	0	0.0994	7.0630
	MSE	1.149	0	0.0978	6.4960
	MAD	0.963	0	0.0989	6.5800

However, what is important to note is that the proposed schemes, used with the unconstrained and equality constrained techniques, were able to nearly reach the fully constrained performance, with less execution time. For

instance, the equality constrained algorithm tested with AVG only requires 1.93 s to reach a similar result to the basic fully constrained, which needs 6.2 s. Finally, with the barycentric approach, the proposed fusion techniques, in particular MAD and ANGLE, were able to reduce the NP from 30.54% to less than 1%.

## 5. CONCLUSION

In this paper, we proposed a new scheme for spectral pre-processing in which we take advantage of the correlation between adjacent bands and divide the spectral interval into correlated sub-intervals. We showed how classification and unmixing can be performed on each sub-interval in addition to the original spectrum. Using adequate fusion techniques, we were able to combine the results such as to correct the errors of the original classifier and improve performance, at the expense of a slight increase in the system complexity.

## 6. REFERENCES

- [1] J. A. Gualtieri and S. Chettri, "Support vector machines for classification of hyperspectral data," Proc. of IEEE-IGARSS 2000, Hawaii, pp. 813-815, July 2000.
- [2] P. Honeine, Z. Noumir, and C. Richard, "Multiclass classification machines with the complexity of a single binary classifier," Signal Processing, vol. 93, pp. 1013-1026, 2013.
- [3] Nirmal Keshava and John F. Mustard, "Spectral unmixing," IEEE signal proc. magazine, January 2002.
- [4] G. Martin, A. Plaza, "Spatial-Spectral pre-processing prior to endmember identification and unmixing of remotely sensed hyperspectral data," IEEE journal of selected topics in applied earth observations and remote sensing, April 2012.
- [5] M. E. Winter, "N-FINDR: An algorithm for fast autonomous spectral endmember determination in hyperspectral data," Proc. SPIE, 1999, vol. 3753, pp. 266-277.
- [6] C. Chang, C.-C. Wuand, W.-M. Liu, and Y.-C. Quyang, "A new growing method for simplex-based endmember extraction algorithm," IEEE Trans. Geosci. Remote Sens., vol. 44, no. 10, pp. 2804-2819, Oct. 2006.
- [7] D Heinz and C. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," IEEE Trans. Geosci. Remote Sens., vol. 39, no. 3, pp. 529-545, Mar. 2001.
- [8] J. S. Cardenas and L. Wang, "Fully constrained linear spectral unmixing: Analytic solution using fuzzy sets," IEEE Trans. Geosci. Remote Sens., vol. 48, no. 11, pp. 3992-4002, Nov. 2010.
- [9] P. Honeine and C. Richard, "Geometric Unmixing of Large Hyperspectral Images: A Barycentric Coordinate Approach," IEEE Trans. Geosci. Remote Sens., vol. 50, no. 6, June 2012.
- [10] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Advances in Spectral-Spatial Classification of Hyperspectral Images," Proc. of the IEEE, vol.101, no.3, pp.652-675, March 2013.
- [11] Y. Yan, Y. Zhao, H. Xue, X. Kou, and Y.Liu, "Integration of Spatial-Spectral Information for Hyperspectral Image Classification," Int. Conf. on Geosci. and Remote Sensing, 2010.
- [12] P. Refaeilzadeh, L. Tang, and H. Liu, "Cross-Validation," Arizona State University, November 2008.