ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1.
Available from http://www.i6doc.com/en/.

Impact of the initialisation of a blind unmixing method dealing with intra-class variability

Charlotte Revel[1,2] and Yannick Deville[1] and Veronique Achard[2] and Xavier Briottet[2] *

1-Universié de Toulouse, UPS-CNRS-OMP
IRAP (Institut de Recherche en Astrophysique et Planétologie)
14 Av. Edouard Belin, F-31400 Toulouse, France

2-ONERA "The French Aerospace Lab"
2 Av. Edouard Belin, 31000 Toulouse, France

Abstract. In hyperspectral imagery, unmixing methods are often used to analyse the composition of the pixels. Such methods usually suppose that a single spectral signature, called an endmember, can be associated with each pure material present in the scene. Such an assumption is no more valid for materials that exhibit spectral variability due to illumination conditions, weathering, slight variations of the composition, etc. In this paper, we investigate a new method based on the assumption of a linear mixing model, that deals with intra-class spectral variability. A new formulation of the linear mixing is provided. In our model a pure material cannot be described by a single spectrum in the image but it can in a pixel. An approach method is presented to handle this new model. It is based on pixel-by-pixel Nonnegative Matrix Factorization (NMF) methods. The methods are tested on a semi-synthetic data set built with spectra extracted from a real hyperspectral image and mixtures of these spectra. We particularly focused our tests to study the impact of the initialisation of our methods.

1 Problem statement

In the framework of remote sensing, unmixing is a common way to deal with mixed pixels in hyperspectral images. This technique extracts subpixel information from hyperspectral images. When no more information is available, this corresponds to a blind source separation problem. Blind unmixing aims at extracting the reflectance spectra of pure materials and the associated proportions in each pixel. An extensive review of unmixing is available in [1]. A common approach of unmixing problems assumes that the spectrum in a pixel is a linear mixture of pure reflectance spectra shared by all pixels. Under this assumption, each observed reflectance spectrum, $\mathbf{x_p} \in \mathbb{R}^{L\times 1}$, can be written as follows:

$$\mathbf{x_p} = \sum_{m=1}^{M} c_{pm}\mathbf{r_m} \ \forall p \in [\![1, P]\!]. \tag{1}$$

*This work was partly funded by the French ANR HYperspectral Imagery for Environmental urban Planning (HYEP) project (ANR 14-CE22-0016-01).

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1. Available from http://www.i6doc.com/en/.

where p is the pixel index and P the number of pixels, m the index of one of the M pure materials present "in the data", $\mathbf{r_m} \in \mathbb{R}^{L\times 1}$ is the reflectance spectrum of the m^{th} material and c_{pm} is the associated coefficient. Reflectance spectra and coefficients are assumed to be nonnegative. Besides, the coefficients corresponding to all materials m are most often assumed to sum to one in each pixel p [2]. This sum-to-one condition leads to:

$$\sum_{m=1}^{M} c_{pm} = 1, \ \forall p \in [\![1, P]\!]. \tag{2}$$

In the above model, each material is described by a single spectrum, but, at macroscopic scale, the spectrum of a same pure material extracted from two different locations of the image can vary. This can be due to illumination variations. However, from a physical point of view, each pixel spectrum can actually be decomposed in a set of component spectra. We provide a solution to this problem by introducing in Eq. (1) a dependency of $\mathbf{r_m}$ with respect to p, the associated pixel. This means that each pure material is no longer described by only one spectrum but by a set of spectra (one per pixel). So, the mixing model (1) becomes:

$$\mathbf{x_p} = \sum_{m=1}^{M} c_{pm}\mathbf{r_m}(p) \quad \forall p \in [\![1, P]\!] \tag{3}$$

where $\mathbf{r_m}(p)$ is the reflectance spectrum associated with the material m and the pixel p. The coefficients c_{pm} still verify the sum-to-one constraint, so Eq. (2) is still verified. This new model yields an ill-posed problem. To solve Eq. (3) under Eq. (2), Sec. 2 overviews two NMF unmixing methods. Sec. 3 presents experimental results and the conclusions are given in Sec. 4.

2 Unmixing methods

Let $\mathbf{X} = [\mathbf{x_1}, \ldots, \mathbf{x_P}]^T$ denote the hyperspectral data matrix, $\mathbf{R} = [\mathbf{r_1}, \ldots, \mathbf{r_M}]^T$ the pure reflectance spectra matrix if only one spectrum per class is considered and $\mathbf{C} = [\mathbf{c_1}, \ldots, \mathbf{c_P}]^T$ the coefficient matrix. For each pixel p, $\mathbf{c_p} = [c_{p1}, \ldots, c_{pM}]^T$ is an M-dimensional vector, containing the set of coefficients associated with the pixel p. The number of pure spectra, M, is assumed to be known in the rest of the paper. The linear mixing model (1) can then be written as follows:

$$\mathbf{X} = \mathbf{CR}. \tag{4}$$

To obtain a similar expression for the model (3), let $\mathbf{R}(p) = [\mathbf{r_1}(p), \ldots, \mathbf{r_M}(p)]^T$ be the set of M constituent material spectra associated with the pixel p, $\tilde{\mathbf{R}} = \begin{bmatrix} \mathbf{R}(1) \\ \ldots \\ \mathbf{R}(P) \end{bmatrix} \in \mathbb{R}^{PM\times L}$ the matrix containing all the pure spectra. Then Eq. (3) may be rewritten as follows:

$$\mathbf{X} = \tilde{\mathbf{C}}\tilde{\mathbf{R}}. \tag{5}$$

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1.
Available from http://www.i6doc.com/en/.

with $\tilde{\mathbf{C}} = \begin{bmatrix} \mathbf{c_1}^T & 0\dots0 & \dots & 0\dots0 \\ 0\dots0 & \mathbf{c_2}^T & \dots & 0\dots0 \\ & & \ddots & \\ 0\dots0 & 0\dots0 & \dots & \mathbf{c_P}^T \end{bmatrix} \in \mathbb{R}^{P\times PM}$ the new coefficient matrix. The sum-to-one constraint, Eq. (2), is kept, so is the nonnegativity constraint. In each method hereafter developed, the normalisation of $\mathbf{c_p}$ or $\tilde{\mathbf{C}}$ is performed after each update to verify the sum-to-one constraint.

2.1 Unconstrained Pixel-by-pixel NMF (UP-NMF)

The classical way to carry out NMF to unmix data does not allow one to extract one set of pure reflectance spectra per pixel. We therefore developed an extended version of Lin's standard NMF algorithm [3], so as to decompose each observed spectrum into pure spectra which are specific to the considered pixel. This extended version aims at minimising $J_1(p) = \frac{1}{2}\|\mathbf{x_p}^T - \mathbf{c_p}^T\mathbf{R}(p)\|_2^2$. From now, let us call this method Unconstrained Pixel-by-pixel NMF (UP-NMF). In [4] we developed the calculation of the update rules of this algorithm. These rules before the $\mathbb{R}^+$ projection can be written in a similar way as in classical NMF:

$$\tilde{\mathbf{C}} \longleftarrow \tilde{\mathbf{C}} + \alpha_{\tilde{C}}(\mathbf{X} - \tilde{\mathbf{C}}\tilde{\mathbf{R}})\tilde{\mathbf{R}}^T$$
$$\tilde{\mathbf{R}} \longleftarrow \tilde{\mathbf{R}} + \alpha_{\tilde{R}}\tilde{\mathbf{C}}^T(\mathbf{X} - \tilde{\mathbf{C}}\tilde{\mathbf{R}})$$

except that the required entries of $\tilde{\mathbf{C}}$ are fixed to zero : see its expression above.

Matrices are initialized by the same set of spectra for all pixels: $\mathbf{R}(p) = \mathbf{R}^{(0)}$. On the contrary, the initialisation of $\mathbf{c_p}$ can be different for each pixel, by performing a fully constrained least square (FCLS) unmixing on the previously found $\mathbf{R}^{(0)}$ for instance.

Due to the high under-determinacy of the optimisation problem, the behavior of UP-NMF is not accurate enough. Spectra $\mathbf{r_m(p)}$ from a same class m may evolve so differently that they tend to define several classes of materials. To limit this spreading of spectra from a same class, constraints are required. Such a constraint is proposed hereafter.

2.2 Inertia-constrained Pixel-by-pixel NMF (IP-NMF)

Our extended method is based on limiting class inertia to reduce the risk for an estimated pure spectrum to go out of its own class. This limitation is introduced in the optimisation problem by adding a penalty term in the cost function. The function J_2 to be minimised becomes:

$$J_2 = \frac{1}{2}\|\mathbf{X} - \tilde{\mathbf{C}}\tilde{\mathbf{R}}\|_F^2 + \mu\sum_{m=1}^{M} Tr(Cov(\mathbf{R}_{\mathcal{C}_m})) \tag{6}$$

where $\mathbf{R}_{\mathcal{C}_m} \in \mathbb{R}^{P\times L}$ denotes the set of reflectance spectra of the m^{th} pure material extracted in the P pixels, μ the constraint parameter and $Cov(\mathbf{R}_{\mathcal{C}_m})$ the $\mathbf{R}_{\mathcal{C}_m}$ covariance matrix. Our calculations [4] show that the resulting update

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1. Available from http://www.i6doc.com/en/.

of IP-NMF under this added constraint and before the $\mathbb{R}^+$ projection and zero forcing in $\tilde{\mathbf{C}}$ is:

$$\tilde{\mathbf{C}} \longleftarrow \tilde{\mathbf{C}} + \alpha_{\tilde{C}}.(\mathbf{X} - \tilde{\mathbf{C}}\tilde{\mathbf{R}})\tilde{\mathbf{R}}^T$$
$$\tilde{\mathbf{R}} \longleftarrow \tilde{\mathbf{R}} + \alpha_{\tilde{R}}.(\tilde{\mathbf{C}}^T(\mathbf{X} - \tilde{\mathbf{C}}\tilde{\mathbf{R}}) - \frac{2\mu}{P}(\mathbf{Id}_{PM} - \frac{1}{P}\mathbf{U})\tilde{\mathbf{R}})$$

with $\mathbf{Id}_{PM}$ the identity matrix of size PM and $\mathbf{U} \in \mathbb{R}^{PM \times PM}$ the "spectrum selection" matrix,

$$\mathbf{U} = \begin{bmatrix} \mathbf{Id}_M & \cdots & \mathbf{Id}_M \\ \vdots & \ddots & \vdots \\ \mathbf{Id}_M & \cdots & \mathbf{Id}_M \end{bmatrix}.$$

2.3 Initialisation methods

Initialisation of NMF is a well known problem. Our methods might also be very sensitive to the initial value of the matrix. We here propose various methods to evaluate the impact of initialisation on results. Various initial spectral matrices, $\mathbf{R}^{(0)}$ (cf. Sec. 2.1) were tested: (i) M spectra are randomly selected from observed data, (ii) the M purest spectra are extracted with a standard method (VCA) [5], (iii) the M initial spectra correspond to the computed averages of the M classes. The initial coefficient matrix is obtained in two different ways: (a) by giving the same constant, $\frac{1}{M}$ to all coefficients, (b) by extracting coefficients associated with initialised spectra initialised with the FCLS method.

3 Tests results

3.1 Test description

The image used to extract pure spectra was taken above Toulouse, France. The city center architecture led us to choose three pure material classes which correctly describe the scene: tile, vegetation (mainly plane trees) and asphalt. Due to its 1.8 m resolution, this image contains many pure spectra. These pure real spectra (hereafter called reference spectra) describe realistic intra-class variabilities. So they were used to build semi-synthetic mixed data to performed the tests. Coefficients were randomly chosen while respecting the nonnegativity and the sum-to-one constraints. The mixing model is Eq.(5)'s one.

Tests were performed on these data with the various initialisation described in Section 2.3.

3.2 Evaluation criteria

Evaluation criteria were chosen to assess the benefits of our method. A major point to be evaluated is the correspondence between estimated pure material reflectance spectra and spectra really present in the pixel. To this end we computed, in each pixel, the spectral angle (SAM) errors between these two sets

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1.
Available from http://www.i6doc.com/en/.

of spectra (Table 1). Another criterion was also computed, the reconstruction error (RE), that shows the effects of our method on the global reconstruction of the image (Table 2). Computing these errors is possible for semi-synthetic data for which all the mixing parameters are known.

3.3 Results

It appears that the random initialisation of $\tilde{\mathbf{R}}$ gives poor results regarding the SAM criterion (higher error that with standard VCA method) for both proposed methods, regardless of the $\tilde{\mathbf{C}}$ initialisation. Indeed the initial spectra are mixed, which leads the iterative methods to converge toward local minima.

A second point to be noted is the very low reconstruction error for every initialisation of the UP-NMF method. Both UP-NMF and IP-NMF are built to minimise the reconstruction error, which leads to good performances regarding this criterion. UP-NMF aim at separately minimising this error in each pixel, which explains the close to 0 reconstruction error (RE). RE is not the best criterion to analyse UP-NMF results.

We particularly focus on the differences between the second (ii) and the third (iii) initializations of $\tilde{\mathbf{R}}$. Indeed it seems that the mean spectra are the best initial spectra we can expect since they describe the average behavior of each class. The obtained SAM and RE confirm this expectation. However these mean spectra are not available in blind unmixing. That is why we tested an initialisation with the results of a standard blind unmixing method. VCA aims at extracting spectra that correspond to the vertices of the data simplex. So these spectra may be expected to be close to the average spectra. Table 1 shows that the IP-NMF method over performed the VCA method both for (a) and (b) initialisations of $\tilde{\mathbf{C}}$. However the improvement is more significant when $\tilde{\mathbf{C}}$ is initialised by a regression of the spectra on the data (FCLS). The improvement of the results is true for UP-NMF and IP-NMF. So, the proposed initialisations in case of blind unmixing allow to increase the performance of unmixing compared to the VCA + FCLS method. Computing the coefficient error (root mean square error) also shows an improvement of the result accuracy: errors of 4% and 6% are respectively computed with IP-NMF and VCA methods.

4 Conclusion

Two blind unmixing methods intended to address intra-class variability are described in the paper. The proposed methods are based on NMF approaches. The tests performed on semi-synthetic data have shown that the results of UP-NMF and IP-NMF over-performed the standard methods, but only for some initialisation conditions. For a non-blind situation, one of the best initialisations of $\tilde{\mathbf{R}}$ seems to be the average spectrum of each class, combined with FCLS initialisation for $\tilde{\mathbf{C}}$. In the blind case, this investigation shows that spectra resulting from VCA are good candidates for $\tilde{\mathbf{R}}^{(0)}$, combined with FCLS coefficient for $\tilde{\mathbf{C}}$ initialisation. For the tested initialisation sets, our new methods are more accurate than VCA+FCLS both spectrally and for the mixing coefficients. Current

ESANN 2017 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 26-28 April 2017, i6doc.com publ., ISBN 978-287587039-1. Available from http://www.i6doc.com/en/.

$\tilde{\mathbf{C}}$ initialisation	$\tilde{\mathbf{R}}$ initialisation	NMF	UP-NMF	IP-NMF ($\mu = 30$)	VCA + FCLS
$\frac{1}{M}$ (a)	Random (i)	15.98	16.32	16.07	7.7
	VCA (ii)	16.40	9.21	7.43	
	Mean (iii)	16.55	7.63	4.99	
FCLS (b)	Random (i)	12.9	15.03	12.53	
	VCA (ii)	7.17	7.56	7.05	
	Mean (iii)	5.11	4.60	4.99	

Table 1: Spectral errors (SAM) for various methods (NMF, UP-NMF, IP-NMF, and VCA) and for various $\tilde{\mathbf{C}}$ and $\tilde{\mathbf{R}}$ initialisations.

$\tilde{\mathbf{C}}$ initialisation	$\tilde{\mathbf{R}}$ initialisation	NMF	UP-NMF	IP-NMF ($\mu = 30$)	VCA + FCLS
$\frac{1}{M}$ (a)	Random (i)	1.22	$< 10^{-3}$	4.07	6.68
	VCA (ii)	1.22	$< 10^{-3}$	4.12	
	Mean (iii)	1.22	$< 10^{-3}$	2.81	
FCLS (b)	Random (i)	6.34	$< 10^{-3}$	0.59	
	VCA (ii)	4.72	$< 10^{-3}$	0.44	
	Mean (iii)	1.93	$< 10^{-3}$	0.38	

Table 2: Reconstruction errors (RE) for various methods (NMF, UP-NMF, IP-NMF, and VCA) and for various $\tilde{\mathbf{C}}$ and $\tilde{\mathbf{R}}$ initialisations.

work aims at developing an extended version of this approach both in terms of initialisation and constraint applied to the cost function.

References

[1] 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*, 5(2):354–379, April 2012.

[2] N. Keshava. A survey of spectral unmixing algorithms. *Lincoln Laboratory Journal*, 14(1):55 – 78, 2003.

[3] C-J. Lin. Projected gradient methods for nonnegative matrix factorization. *Neural Computation*, 19(10):2756–2779, 2007.

[4] C. Revel, Y. Deville, V. Achard, and X. Briottet. A method based on nonnegative matrix factorization dealing with intra-class variability for unsupervised hyperspectral unmixing. In *7th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, June 2015.

[5] J.M.P. Nascimento and J.M. Bioucas Dias. Vertex component analysis: a fast algorithm to unmix hyperspectral data. *IEEE Transactions on Geoscience and Remote Sensing*, 43(4):898–910, April 2005.