

COLLABORATIVE NONNEGATIVE MATRIX FACTORIZATION FOR REMOTELY SENSED HYPERSPECTRAL UNMIXING

Jun Li^{1,2}, José M. Bioucas-Dias², and Antonio Plaza¹

¹Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, University of Extremadura, E-10071 Caceres, Spain.

²Instituto de Telecomunicações, Instituto Superior Técnico, TULisbon, 1900-118, Lisbon, Portugal.

ABSTRACT

In this paper, we develop a new algorithm for hyperspectral unmixing which can provide suitable endmembers (and their corresponding abundances) in a single step. Hence, the algorithm does not require a previous subspace identification step to estimate the number of endmembers as it can cope with the two most likely scenarios in practice (i.e., the number of endmembers is correctly determined or overestimated *a priori*). The proposed approach, termed collaborative NMF (CoNMF), uses a collaborative regularization prior which forces the abundances corresponding to the overestimated endmembers to zero, such that it is guaranteed that only the true endmembers have fractional abundance contributions and the estimation of the number of endmembers is not required in advance. The obtained experimental results demonstrate that the proposed method exhibits very good performance in case the number of endmember is not available *a priori*.

Index Terms—Hyperspectral imaging, spectral unmixing, collaborativity.

1. INTRODUCTION

Spectral unmixing is an important task for remotely sensed hyperspectral data exploitation [1]. It amounts at estimating the abundance of spectrally pure components (called *endmembers* in hyperspectral imaging literature). In the last few years, several techniques have been developed for automatic endmember identification, with and without assuming the presence of pure spectral signatures in the input hyperspectral data [2–4]. However, the estimation of the number of endmembers (generally assumed to be known as prior information) remains a more elusive goal, despite the availability of subspace identification algorithms specifically designed for

this purpose [5, 6]. Let p denote the true number of endmembers in a hyperspectral image and let \hat{p} denote the number of endmembers estimated by a certain algorithm. In general, \hat{p} fluctuates around p but three situations are possible:

- If $\hat{p} = p$, we have an ideal situation and most existing endmember identification algorithms rely on this assumption [1].
- If $\hat{p} < p$, then the number of endmembers is underestimated. This situation is generally easy to identify as the reconstruction of the original hyperspectral image based, for instance, on the well-known linear mixture model [7] should provide a very high reconstruction error, thus allowing a trained analyst to identify that more endmembers are needed for the model to work properly.
- If $\hat{p} > p$, then the number of endmembers is overestimated. This is a difficult problem as compared with the underestimation situation, since the reconstruction error using more endmembers than needed would provide similar values than those obtained when $\hat{p} = p$. This is particularly critical for endmember identification algorithms designed without assuming that the true endmembers are present in the input hyperspectral data [8–14].

In this paper, we develop a new algorithm for hyperspectral unmixing which can provide suitable endmembers (and their corresponding abundances) when $\hat{p} = p$ and $\hat{p} > p$. Hence, the algorithm does not require a previous subspace identification step to estimate the number of endmembers as it can cope with the two most likely scenarios in practice (i.e., the number of endmembers is correctly determined or overestimated *a priori*). The algorithm addresses a limitation observed, for instance, in nonnegative matrix factorization (NMF) algorithms for endmember identification [9, 13], which tend to assign abundances to all obtained endmembers, such that overestimated endmembers (i.e., those resulting in the case that $\hat{p} > p$) will have abundance contributions which is not true in reality. On the other hand, minimum volume enclosing algorithms for endmember identification will fail

This work has been supported by the European Community's Marie Curie Research Training Networks Programme under contract MRTN-CT-2006-035927, Hyperspectral Imaging Network (HYPER-I-NET). Funding from the Portuguese Science and Technology Foundation, project PEst-OE/EEI/LA0008/2011, and from the Spanish Ministry of Science and Innovation (CEOS-SPAIN project, reference AYA2011-29334-C02-02) is also gratefully acknowledged.

in the case that $\hat{p} > p$ due to the determinant calculation involved in those algorithms. The proposed approach, termed collaborative NMF (CoNMF), uses a collaborative regularization prior [15] which forces the abundances corresponding to the overestimated endmembers to zero, such that it is guaranteed that only the true endmembers have fractional abundance contributions and the estimation of the number of endmembers is not required in advance. The obtained experimental results demonstrate that the proposed method exhibits very good performance in case the number of endmember is not available *a priori*.

2. PROPOSED APPROACH

Let $\mathbf{Y} \equiv [\mathbf{y}_1, \dots, \mathbf{y}_n] \in \mathbb{R}^{d \times n}$ be a hyperspectral image with n spectral vectors and d spectral bands. The hyperspectral image is represented as a matrix, holding in its columns the spectral vectors $\mathbf{y}_i \in \mathbb{R}^d$, for $i = 1, \dots, n$. Under the linear mixing model [7], we have:

$$\begin{aligned} \mathbf{Y} &= \mathbf{M}\mathbf{S} + \mathbf{n} \\ \text{s.t. : } & \mathbf{S} \geq 0, \mathbf{1}_p^T \mathbf{S} = \mathbf{1}_n^T, \end{aligned} \quad (1)$$

where $\mathbf{M} \equiv [\mathbf{m}_1, \dots, \mathbf{m}_p] \in \mathbb{R}^{d \times p}$ is a so-called mixing matrix containing p endmembers, \mathbf{m}_i denotes the i -th endmember signature, and $\mathbf{S} = [\mathbf{s}_1, \dots, \mathbf{s}_p]^T \in \mathbb{R}^{p \times n}$ is the abundance matrix containing the endmember fractions for every pixel in the hyperspectral scene. Finally, \mathbf{n} collects the errors that affect the measurement process (e.g., noise). For each pixel, the abundance fractions should be no less than zero and sum to one, which are known as the non-negativity and sum-to-one constraints. These are represented in Eq. (1), where $\mathbf{1}_p$ is a column vector containing p ones. In this work, we solve the optimization problem in Eq. (1) using NMF as follows:

$$\begin{aligned} (\widehat{\mathbf{M}}, \widehat{\mathbf{S}}) &= \arg \min_{\mathbf{M}, \mathbf{S}} \frac{1}{2} \|\mathbf{Y} - \mathbf{M}\mathbf{S}\|_2^2 + \alpha \sum_{i=1}^{\hat{p}} \|\mathbf{s}_i\|_2^q \\ &+ \frac{\beta}{2} \sum_{i=1}^{\hat{p}} \|\mathbf{m}_i - \bar{\mathbf{y}}\|_2^2 \\ \text{s.t. : } & \mathbf{S} \geq 0, \mathbf{1}_p^T \mathbf{S} = \mathbf{1}_n^T, \end{aligned} \quad (2)$$

where α and β are regularization parameters and $\bar{\mathbf{y}}$ is the mean value of the data. The regularization term on the abundance matrix is a collaborative prior which promotes complete lines of zeros, where $0 < q \leq 1$ is a parameter controlling the degree of collaborativity. The regularization term on the right hand side, similar to those introduced in [16, 17], pushes the endmembers towards the data set mass center, thus promoting minimum size of the convex hull defined by the endmembers. Contrarily to the widely used determinant regularizer, the proposed minimum size regularizer is convex and well defined for any number of endmembers, even when p is overestimated.

The optimization problem in Eq. (2) is nonconvex and,

hence, very difficult to solve. In this work, we adopt a greedy iterative scheme to obtain local minima. At the $t + 1$ -th iteration, we solve the following optimizations with respect to \mathbf{M} and \mathbf{S} :

$$\widehat{\mathbf{M}}_{(t+1)} = \arg \min_{\mathbf{M}} \underbrace{\frac{1}{2} \|\mathbf{Y} - \mathbf{M}\widehat{\mathbf{S}}_{(t)}\|_2^2 + \frac{\beta}{2} \sum_{i=1}^{\hat{p}} \|\mathbf{m}_i - \bar{\mathbf{y}}\|_2^2}_{f_{\mathbf{M}|\mathbf{S}}(\mathbf{M}|\mathbf{S}_{(t)})}} \quad (3)$$

and

$$\begin{aligned} \widehat{\mathbf{S}}_{(t+1)} &= \arg \min_{\mathbf{S}} \underbrace{\frac{1}{2} \|\mathbf{Y} - \widehat{\mathbf{M}}_{(t)}\mathbf{S}\|_2^2 + \frac{\alpha \times q}{2} \sum_{i=1}^{\hat{p}} \frac{\|\mathbf{s}_i\|_2^2}{\|\mathbf{s}_{i(t)}\|_2^{2-q}}}_{f_{\mathbf{S}|\mathbf{M}}(\mathbf{S}|\mathbf{M}_{(t+1)})}} \\ \text{s.t. : } & \mathbf{S} \geq 0, \mathbf{1}_p^T \mathbf{S} = \mathbf{1}_n^T. \end{aligned} \quad (4)$$

The right hand side term in (4) is a quadratic lower bound for the sparsity inducing regularizer $\sum_{i=1}^{\hat{p}} \|\mathbf{s}_i\|_2^q$. Thus the minimization (4) decreases the objective function (2).

Problem (3) is quadratic. Its solution is given by

$$\widehat{\mathbf{M}}_{(t+1)} = (\mathbf{Y}\mathbf{S}_{(t)}^T + \beta\bar{\mathbf{Y}}\bar{\mathbf{Y}})(\mathbf{S}_{(t)}\mathbf{S}_{(t)}^T + \beta\mathbf{I}_{\hat{p}})^{-1}, \quad (5)$$

where $\mathbf{I}_{\hat{p}} \in \mathbb{R}^{\hat{p} \times \hat{p}}$ is the identity matrix and $\bar{\mathbf{Y}}_{\hat{p}} = [\bar{\mathbf{y}}, \dots, \bar{\mathbf{y}}]_{\hat{p}} \in \mathbb{R}^{d \times \hat{p}}$.

Problem (4) is quadratic with convex constraints. We solve it using the projected gradient method. That is, at iteration t , we run the iterative procedure

$$\widehat{\mathbf{S}}^{(i+1)} = P\left(\widehat{\mathbf{S}}^{(i)} - \lambda^{(i)} \nabla_{\mathbf{S}} f(\mathbf{S}^{(i)} | \mathbf{M}_{(t+1)})\right), \quad (6)$$

where P denotes successive projection onto the nonnegative orhant and onto the unit simplex. From expression (4) we compute $\nabla_{\mathbf{S}|\mathbf{M}}$:

$$\begin{aligned} \nabla_{\mathbf{S}|\mathbf{M}} f(\mathbf{S}|\mathbf{M}_{(t+1)}) &= \widehat{\mathbf{M}}_{(t+1)}^T (\widehat{\mathbf{M}}_{(t+1)}\mathbf{S} - \mathbf{Y}) \\ &+ \alpha \times q \left[\frac{s_{ij}}{\|\mathbf{s}_{i(t)}\|_2^{2-q}} \right], \end{aligned} \quad (7)$$

where s_{ij} is denotes the abundance fraction of material \mathbf{m}_i at the j -th pixel, and $[c]$ is a matrix with the i -th line and the j -th column element equal to c . If $\|\mathbf{s}_{i(t)}\|_2 = 0$, we take its gradient to zero. Parameters $\lambda^{(i)}$ is chosen such that $f(\mathbf{S}^{(i+1)} | \mathbf{M}_{(t+1)}) \leq f(\mathbf{S}^{(i)} | \mathbf{M}_{(t+1)})$.

To lighten the computational complexity of the unmixing algorithm, we project the data onto an overestimated subspace. This is a simple procedure because real hyperspectral data belongs, very often, to a low dimensional subspace.

3. EXPERIMENTAL RESULTS

In this section, we evaluate the proposed CoNMF method using synthetic hyperspectral data in comparison with the minimum volume constrained NMF (MVC-NMF) [9]. The simulated images are generated according to the linear mixture model in Eq. (1), following the procedure described in the work of MVC-NMF [9]. The simulated images comprise 58×58 pixels and $p = 4$ endmembers, and the spectral signatures are selected from the United States Geological Survey (USGS) library¹. To ensure that no pure pixel is present in the simulated images, we discard all pixels with abundance fractions larger than 0.8. Zero-mean Gaussian noise is added to the simulated images with signal-to-noise ratio (SNR) given by $\text{SNR} = 10 \log_{10}(E[(\mathbf{MS})^T(\mathbf{MS})]/[\mathbf{n}^T \mathbf{n}])$. In this work, we set the SNR in the simulated images to 20 dB. In all experiments, we projected the data into a \hat{p} -dimensional subspace and set $q = 1/4$.

Table 1 shows an evaluation of the performance of CoNMF and MVC-NMF in the unmixing of simulated hyperspectral data with SNR=20dB. Both algorithms shared the same initialization condition, given by the endmembers produced by the vertex component analysis (VCA) algorithm in [18]. As this algorithm starts with a random initialization condition, each value in Table 1 is obtained by averaging 10 random tests. In the table, several performance discriminators are reported, namely the root mean reconstruction error (RMSE), the spectral angle distance (SAD) [7] and two error metrics focused on the quality of the estimated endmembers, $\|\hat{\mathbf{M}} - \mathbf{M}\|_F$, and on the quality of the estimated abundances, $\frac{1}{n \times p} \|\hat{\mathbf{S}} - \mathbf{S}\|_F$, where F denotes the Frobenius norm of a given matrix \mathbf{X} , such that $\|\mathbf{X}\|_F \equiv \sqrt{\text{trace}\{\mathbf{X}\mathbf{X}^T\}}$. Finally, $\xi_c = \frac{1}{\hat{p}-p} \sum_{i=1}^{\hat{p}-p} n_0(i)/n$ is a new metric introduced to measure the effectiveness of the collaborativity, which measures the number of zeros in s_i –denoted by $n_0(i)$ –. Several conclusions can be obtained from Table 1. First and foremost, it can be seen that the CoNMF algorithm is quite robust in the case that $\hat{p} \geq p$. Second, when the number of endmembers is underestimated ($\hat{p} < p$) the RMSE metric provides a very high value and should allow a trained analyst in detecting that the value of \hat{p} should be increased. Finally, it is important to note that the collaborative term introduced in the proposed CoNMF enforces that the abundance values associated to the overestimated endmembers to zero in the case that $\hat{p} > p$, as indicated by the values of ξ_c reported in Table 1.

To conclude this paper, Fig. 1 illustrates graphically the performance of the two compared unmixing methods by projecting the simulated data on a two-dimensional subspace. In Fig. 1, we show the true endmembers in black color, the initial condition in cyan color (this corresponds to the endmembers provided by VCA which are used to initialize both algorithms: CoNMF and MVC-NMF, and the iterative process for

CoNMF (in red color) and MVC-NMF (in blue color) from the initial to the final condition. As shown by Fig. 1, the proposed CoNMF always provides endmembers which are very close to the true ones when $\hat{p} \geq p$. Although these results are quite encouraging, further experimentation with real hyperspectral data sets is highly desirable. These will be provided in the final version of this paper.

4. REFERENCES

- [1] A. Plaza, Q. Du, J. Bioucas-Dias, X. Jia, and F. Kruse, "Foreword to the special issue on spectral unmixing of remotely sensed data," *IEEE Trans. Geo. Rem. Sens.*, vol. 49, no. 11, pp. 4103–4110, 2011.
- [2] Q. Du, N. Raksuntorn, N. H. Younan, and R. L. King, "End-member extraction for hyperspectral image analysis," *Applied Optics*, vol. 47, pp. 77–84, 2008.
- [3] M. Parente and A. Plaza, "Survey of geometric and statistical unmixing algorithms for hyperspectral images," *Proceedings of the 2nd IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pp. 1–4, Reykjavik, Iceland, Jun. 14–16, 2010.
- [4] J. M. Bioucas-Dias and A. Plaza, "Hyperspectral unmixing: geometrical, statistical, and sparse regression-based approaches," *Proceedings of SPIE, Image and Signal Processing for Remote Sensing XVI*, vol. 7830, pp. 1–15, Toulouse, France, Sept. 20–23, 2010.
- [5] C.-I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Trans. Geo. Rem. Sens.*, vol. 42, no. 3, pp. 608–619, 2004.
- [6] J. M. Bioucas-Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geo. Rem. Sens.*, vol. 46, no. 8, pp. 2435–2445, 2008.
- [7] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 44–57, 2002.
- [8] M. D. Craig, "Minimum-volume transforms for remotely sensed data," *IEEE Trans. Geo. Rem. Sens.*, vol. 32, pp. 542–552, 1994.
- [9] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Trans. Geo. Rem. Sens.*, vol. 45, no. 3, pp. 765–777, 2007.
- [10] J. Li and J. Bioucas-Dias, "Minimum volume simplex analysis: a fast algorithm to unmix hyperspectral data," *Proc. IEEE International Geoscience and Remote sensing Symposium*, vol. 3, pp. 250–253, 2008.
- [11] J. Bioucas-Dias, "A variable splitting augmented lagrangian approach to linear spectral unmixing," *Proceedings of the 1st IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pp. 1–4, Grenoble, France, Aug. 26–28, 2009.
- [12] Tsung-Han Chan, Chong-Yung Chi, Yu-Min Huang, and Wing-Kin Ma, "A convex analysis-based minimum-volume enclosing simplex algorithm for hyperspectral unmixing," *IEEE Transactions on Signal Processing*, vol. 57, pp. 4418–4432, 2009.
- [13] A. Huck, M. Guillaume, and J. Blanc-Talon, "Minimum dispersion constrained nonnegative matrix factorization to unmix hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 6, pp. 2590–2602, june 2010.
- [14] J. M. Bioucas-Dias and J. Nascimento, "Hyperspectral unmixing based on mixtures of Dirichlet components," *IEEE Trans. Geo. Rem. Sens.*, 2011, in press.
- [15] J. Bioucas-Dias and M. Figueiredo, "An alternating direction algorithm for collaborative hierarchical sparse regularization," *Technical Report, IST*, 2011.
- [16] M. Berman, H. Kiiveri, R. Lagerstrom, A. Ernst, R. Dunne, and J. F. Huntington, "ICE: a statistical approach to identifying endmembers in hyperspectral images," *IEEE Trans. Geo. Rem. Sens.*, vol. 42, no. 10, pp. 2085–2095, 2004.
- [17] M. Schmidt M. Arngren and J. Larsen, "Bayesian nonnegative matrix factorization with volume prior for unmixing of hyperspectral images," *Journal of Signal Processing Systems*, vol. 10, pp. 1–18, 2009.
- [18] J. M. P. Nascimento and J. M. Bioucas-Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geo. Rem. Sens.*, vol. 43, no. 4, pp. 898–910, 2005.

¹<http://speclab.cr.usgs.gov/spectral-lib.html>

Table 1. Evaluation of the performance of CoNMF and MVC-NMF in the unmixing of simulated hyperspectral data with SNR=20dB for different \hat{p} . The true number of endmembers in this experiment is $p = 4$.

Number of estimated endmembers (\hat{p})	CoNMF					MVC-NMF				
	RMSE	SAD	$\ \hat{\mathbf{M}} - \mathbf{M}\ _F$	$\frac{1}{n \times p} \ \hat{\mathbf{S}} - \mathbf{S}\ _F$	ξ_c	RMSE	SAD	$\ \hat{\mathbf{M}} - \mathbf{M}\ _F$	$\frac{1}{n \times p} \ \hat{\mathbf{S}} - \mathbf{S}\ _F$	
3	0.0346	-	-	-	-	0.0315	-	-	-	-
4	0.0012	1.2770	0.2880	0.0009	-	0.0012	0.6444	0.1446	0.0005	
5	0.0020	1.8887	0.4005	0.0026	0.3287	0.0015	7.0552	1.4857	0.0215	
6	0.0026	2.2472	0.4391	0.0065	0.3899	0.0024	6.5750	1.3948	0.0376	
7	0.0026	2.9030	0.6006	0.0087	0.3657	0.0031	6.5163	1.7264	0.0486	

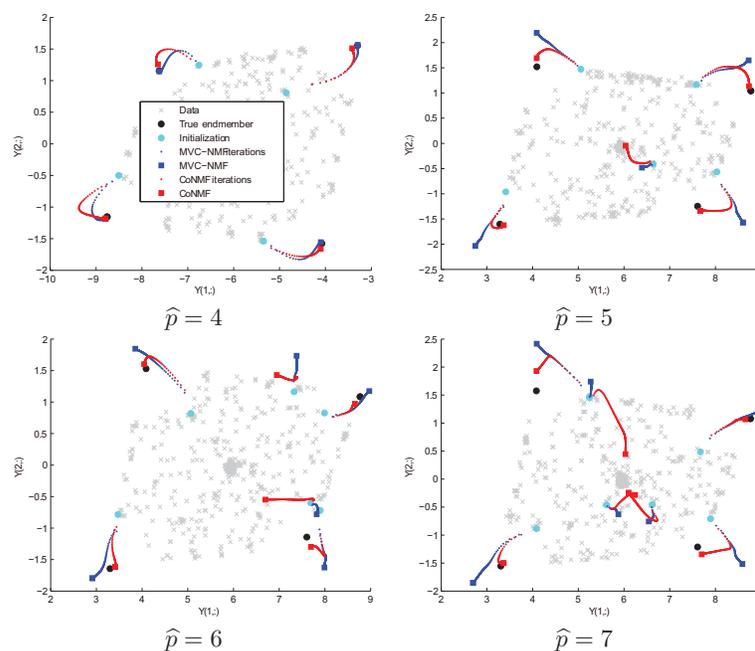


Fig. 1. Graphical assessment of the performance of CoNMF and MVC-NMF in the unmixing of simulated hyperspectral data with SNR=20dB for cases in which $\hat{p} \geq p$. The $p = 4$ true endmembers are displayed in black color, the initial condition used for both algorithms (endmembers produced by the VCA algorithm) is displayed in cyan color, and the iterative process for CoNMF (in red color) and MVC-NMF (in blue color), from the initial to the final condition, is also represented in graphical form.