A New Multiplicative Nonnegative Matrix Factorization Method for Unmixing Hyperspectral Images Combined with Multispectral Data

Yasmine Kheira Benkouider [1, 2], Moussa Sofiane Karoui [1, 2, 3], Yannick Deville [3] and Shahram Hosseini [3]

[1] Centre des Techniques Spatiales, Arzew, Algeria
[2] Laboratoire Signaux et Images, Université des Sciences et de la Technologie, Oran, Algeria
[3] Institut de Recherche en Astrophysique et Planétologie (IRAP), Université de Toulouse, UPS-OMP, CNRS, Toulouse, France
ybenkouider@cts.asal.dz, {Sofiane.Karoui, Yannick.Deville, Shahram.Hosseini}@irap.omp.eu

ABSTRACT

In these investigations, a novel algorithm is proposed for linearly unmixing hyperspectral images combined with multispectral data. This algorithm, which is used to unmix the considered hyperspectral image, is founded on nonnegative matrix factorization. It minimizes, with new multiplicative update rules, a novel cost function, which includes multispectral data and a spectral degradation model between these data and hyperspectral ones. The considered multispectral variables are also used to initialize the proposed algorithm. Tests, using synthetic data, are carried out to assess the performance of our algorithm and its robustness to spectral variability between the processed data. The obtained results are compared to those of state of the art methods. These tests prove that the proposed algorithm outperforms all other used approaches.

Index Terms— Hyper/multispectral imaging, linear unmixing, multiplicative update rule, nonnegative matrix factorization, spectral degradation model

1. INTRODUCTION

Hyperspectral sensors, with high spectral resolution, gather data in the visible and infrared wavelength domains, thus allowing precise classification of materials present in imaged areas. However, these sensors are characterized by a lower spatial resolution as compared to multispectral ones, which have a low spectral resolution. Therefore, each reflectance pixel-spectrum of a hyperspectral image is thus typically a combination of contributions from a number of reflectance endmember-spectra that are present in the acquired data. The combination is often assumed to be linear [1] and linear spectral unmixing (LSU) methods are used to linearly unmix all pixel-spectra into a set of endmember-spectra, and a collection of associated abundance fractions.

As the analyzed data are nonnegative, nonnegative matrix factorization (NMF) [2] methods, consisting in factorizing a nonnegative matrix into a product of two other nonnegative matrices, are an interesting tool to tackle the above problem.

In some applications, when unmixing a hyperspectral image, a high spatial resolution multispectral image of the same area or of another one containing the same endmembers as the considered hyperspectral image may be available. Many endmember-spectra extraction approaches applied to hyperspectral data do not use such information available in multispectral data. However, some NMF-based literature methods jointly use these two types of data for *improving the spatial resolution* of the considered hyperspectral image [3], [4]. Recently, in [5], the authors proposed an approach that combines these two types of images in order to unmix a hyperspectral one. This NMF-based approach not only uses the standard update rules but also a supplementary one. This additional rule, which is not derived from the considered criterion, constrains desired hyperspectral endmember-spectra to get the corresponding multispectral endmember-spectra values at some wavelengths. The latter spectra values can be precisely estimated from multispectral unmixed pixels (pure pixels), which are highly expected to be present because of the high spatial resolution of this image. However, this supplementary rule does not necessarily ensure the decrease of the considered criterion from an iteration to another during the optimization process.

This article reports on a new algorithm for linear endmember-spectra extraction from a hyperspectral image, also considering multispectral data. The proposed algorithm, which is used to unmix the considered hyperspectral image, is based on NMF. This algorithm minimizes, with new multiplicative update rules, a novel cost function, which combines multispectral data and a spectral degradation model between these data and hyperspectral ones. The considered multispectral variables are also used to initialize the proposed algorithm. In practice, with this algorithm, the new used cost function decreases from an iteration to another.

The rest of this article is organized in the following manner. The data model used in LSU techniques is described in Section 2. The designed algorithm is presented in Section 3. In Section 4, the experiments conducted with synthetic data are described. In this section, results obtained by the proposed method are compared with those provided by

ISBN 978-0-9928626-7-1 © EURASIP 2017

some classical methods. Eventually, a conclusion is given in Section 5.

2. MATHEMATICAL DATA MODEL

In the following, the mathematical data model, used in LSU techniques for unmixing hyperspectral images, is described. Therefore, and as explained above, each pixel-spectrum in hyperspectral data is supposed to be a linear mixture of the endmember-spectra contained in the considered pixel. Thus, the whole observed reflectance hyperspectral image $X_h \in R_+^{N_h \mathrm{x} K_h}$ can be modeled, in matrix form, as [5]

$$X_h = A_h S_h, \tag{1}$$

where each row of the matrix X_h is one spectral band of the considered hyperspectral image. The K_h pixels of this image are here reorganized as a one-dimensional array. N_h is the number of spectral bands of this image. Every column of the matrix $A_h \in R_+^{N_h \mathrm{x} L}$ is one hyperspectral endmember-spectrum and the matrix $S_h \in R_+^{L \mathrm{x} K_h}$ contains abundance fractions: each row represents all abundance fractions of one endmember in all pixels. These abundance coefficients must satisfy the abundance sum-to-one constraint [5]. L is the (known or estimated) number of endmembers.

As mentioned above, the proposed algorithm is targeted at unmixing a hyperspectral image combined with multispectral data. This combination consists in using a relation linking these two types of data. We here use the relationship which links hyperspectral and multispectral endmember-spectra as follows

$$A_m \approx D_\lambda A_h, \tag{2}$$

where each column of the matrix $A_m \in R_+^{N_m \mathrm{x} L}$ represents one multispectral endmember-spectrum and N_m is the number of spectral bands of the multispectral data. $D_\lambda \in R_+^{N_m \mathrm{x} N_h}$ is a fixed linear operator performing the spectral degradation between the two considered data. Ideally, this sparse and nonnegative matrix models the transformation between the spectral responses of hyperspectral and multispectral sensors.

3. PROPOSED ALGORITHM

The designed algorithm is based on the NMF technique [2] to unmix a hyperspectral image, combined with multispectral data, by deriving an estimate $\tilde{A}_h$ of the collection of hyperspectral endmember-spectra and an estimate $\tilde{S}_h$ of the collection of abundance fractions, such that

$$X_h \approx \tilde{A}_h \tilde{S}_h. \tag{3}$$

Standard NMF algorithms are very easy to use and implement. The initial algorithm [6], [7] minimizes the cost function

$$J_1 = \frac{1}{2} \left\| X_h - \tilde{A}_h \tilde{S}_h \right\|_F^2 \tag{4}$$

by using iterative multiplicative update rules. $\|.\|_F$ represents the Frobenius norm.

To improve the performance of the NMF-unmixing process when a multispectral image is available in addition, the following cost function, which includes an additional term containing multispectral variables, is proposed here

$$J_2 = \frac{\alpha}{2} \left\| X_h - \tilde{A}_h \tilde{S}_h \right\|_F^2 + \frac{\beta}{2} \left\| \hat{A}_m - D_\lambda \tilde{A}_h \right\|_F^2, \tag{5}$$

where α and β are weighting coefficients that are set to the inverse of the number of elements in each considered norm. The matrix $\hat{A}_m$ corresponds to the set of estimated multispectral endmember-spectra that can be automatically derived by using multispectral unmixing processes with pure pixel assumption [8]-[10]. This matrix does not evolve during the updating stage of the proposed algorithm. Thus, the only optimized variables in the proposed algorithm are: $\tilde{A}_h$ and $\tilde{S}_h$.

We propose to optimize the cost function J_2 with an iterative gradient-based algorithm, which practically yields a monotonic decrease, from an iteration to another, of J_2. In order to easily permit one to obtain the gradient expressions, the cost function J_2 can be expressed as

$$\begin{aligned} J_2 = {} & \frac{\alpha}{2} \mathrm{Tr}\left(X_h X_h^{\mathrm{T}}\right) - \alpha \mathrm{Tr}\left(X_h \tilde{S}_h^{\mathrm{T}} \tilde{A}_h^{\mathrm{T}}\right) + \\ & \frac{\alpha}{2} \mathrm{Tr}\left(\tilde{A}_h \tilde{S}_h \tilde{S}_h^{\mathrm{T}} \tilde{A}_h^{\mathrm{T}}\right) + \frac{\beta}{2} \mathrm{Tr}\left(\hat{A}_m \hat{A}_m^{\mathrm{T}}\right) - \\ & \beta \mathrm{Tr}\left(D_\lambda \tilde{A}_h \hat{A}_m^{\mathrm{T}}\right) + \frac{\beta}{2} \mathrm{Tr}\left(D_\lambda \tilde{A}_h \tilde{A}_h^{\mathrm{T}} D_\lambda^{\mathrm{T}}\right). \end{aligned} \tag{6}$$

In the latter expression, Tr(.) and (.)$^{\mathrm{T}}$, respectively, represent the matrix trace and the matrix transpose. Using the properties provided in [11], the gradient expressions of the considered cost function are as follows (in matrix form)

$$\frac{\partial J_2}{\partial \tilde{A}_h} = \alpha\left(\tilde{A}_h \tilde{S}_h \tilde{S}_h^{\mathrm{T}} - X_h \tilde{S}_h^{\mathrm{T}}\right) + \beta\left(D_\lambda^{\mathrm{T}} D_\lambda \tilde{A}_h - D_\lambda^{\mathrm{T}} \hat{A}_m\right), \tag{7}$$

$$\frac{\partial J_2}{\partial \tilde{S}_h} = \alpha\left(\tilde{A}_h^{\mathrm{T}} \tilde{A}_h \tilde{S}_h - \tilde{A}_h^{\mathrm{T}} X_h\right). \tag{8}$$

Considering that the proposed algorithm is a gradient-based one, the following iterative update rule is used

$$\theta \leftarrow \theta - \varphi_\theta \odot \frac{\partial J_2}{\partial \theta}, \tag{9}$$

where $\odot$ denotes the element-wise multiplication. θ corresponds to the matrix $\tilde{A}_h$ or $\tilde{S}_h$, and φ_θ is a learning rate in matrix form. Applying this rule is not sufficient since it

ISBN 978-0-9928626-7-1 © EURASIP 2017

does not ensure nonnegativity. To guarantee this constraint, an iterative and multiplicative rule is obtained from the above gradient-based one as follows. As it appears in (7) and (8), the derivative $\frac{\partial J_2}{\partial \theta}$ of the considered criterion with respect to θ can always be expressed as the difference of two nonnegative functions such that $\frac{\partial J_2}{\partial \theta} = \frac{\partial J_2}{\partial \theta}^{+} - \frac{\partial J_2}{\partial \theta}^{-}$. In the gradient expressions (7) and (8), the nonnegative function $\frac{\partial J_2}{\partial \theta}^{+}$ corresponds to the terms preceded by a plus sign, while $\frac{\partial J_2}{\partial \theta}^{-}$ corresponds to the terms preceded by a minus sign. The nonnegativity constraint can be fulfilled by initializing θ with a nonnegative value and then choosing the value of the learning rate matrix φ_θ according to

$$\varphi_\theta = \theta \oslash \frac{\partial J_2}{\partial \theta}^{+}, \tag{10}$$

where $\oslash$ stands for element-wise division. Thus, the update rule (9) becomes

$$\theta \leftarrow \theta \odot \frac{\partial J_2}{\partial \theta}^{-} \oslash \frac{\partial J_2}{\partial \theta}^{+}. \tag{11}$$

As a result, the final proposed iterative and multiplicative rules for the $\tilde{A}_h$ and $\tilde{S}_h$ matrices read

$$\tilde{A}_h \leftarrow \tilde{A}_h \odot (\alpha X_h \tilde{S}_h^{\mathrm{T}} + \beta D_\lambda^{\mathrm{T}} \hat{A}_m) \oslash (\alpha \tilde{A}_h \tilde{S}_h \tilde{S}_h^{\mathrm{T}} + \beta D_\lambda^{\mathrm{T}} D_\lambda \tilde{A}_h + \epsilon), \tag{12}$$

$$\tilde{S}_h \leftarrow \tilde{S}_h \odot (\tilde{A}_h^{\mathrm{T}} X_h) \oslash (\tilde{A}_h^{\mathrm{T}} \tilde{A}_h \tilde{S}_h + \epsilon), \tag{13}$$

where ϵ is a very small and positive value, which is added to the denominator of each update rule to avoid division by zero.

By providing an initial nonnegative value of the updated matrices $\tilde{A}_h$ and $\tilde{S}_h$, the above iterative and multiplicative rules ensure the nonnegativity constraint and yield a descent algorithm, which is such that the proposed cost function practically decreases at each step of its optimization stage.

The proposed algorithm, as usual NMF methods, is not guaranteed to give a unique solution and its convergence point possibly depends on its initialization. To avoid random initialization from the viewpoint of the designed algorithm, and as the initialization stage, the initial estimated hyperspectral endmember-spectra $\tilde{A}_h$ are calculated from the estimated multispectral endmember-spectra $\hat{A}_m$ as follows: starting from each N_m-point multispectral endmember-spectrum, an inter/extrapolation is performed by using cubic spline approximation [12], in order to obtain a first rough approximation of each hyperspectral endmember-spectrum (with N_h samples). In order to satisfy the nonnegativity constraint of the proposed NMF algorithm, negative inter/extrapolated values are set to a very small positive value ε. Besides, each initial value of the matrix $\tilde{S}_h$ is set to $1/L$.

Another constraint to be taken into account is the abundance sum-to-one property. To this end, we use the technique described in [13].

A predefined maximum number of iterations is used as the stopping criterion of the proposed algorithm.

4. EXPERIMENTAL RESULTS

As stated in Section 1, tests using synthetic data are carried out to assess the performance of the proposed algorithm and its robustness to spectral variability between the considered data. Also, the obtained results are compared to those of literature methods.

4.1. Performance evaluation criteria

We first compute the spectral angle mapper (SAM) [9] involving the original and estimated hyperspectral endmember-spectra. A smaller angle corresponds to better spectra extraction.

We then compute the Normalized Mean Square Error (NMSE) [9]. This criterion is also calculated by using the original and estimated hyperspectral endmember-spectra. Again, a smaller NMSE value corresponds to better spectra extraction.

4.2. Tested data and experiments design

The tested synthetic data are generated as follows. Hyperspectral endmember-spectra are randomly chosen from a spectral library measured from 0.4 to 2.5 µm with 420 wavelengths and compiled by the United States Geological Survey (USGS) [14]. The number of endmembers is varied from 2 to 10 in our tests. For a fixed number of endmembers, the selected endmember-spectra are used to generate a 30×30 pixel synthetic hyperspectral image according to the considered linear mixing model (1). The abundance coefficients used to generate these mixtures are randomly selected. These abundance fractions maps are pre-processed so as to satisfy the abundance sum-to-one constraint. Also, for each selected hyperspectral endmember-spectrum, a corresponding multispectral endmember-spectrum is derived therefrom by just averaging the samples of the latter hyperspectral spectrum over the wavelength regions corresponding to the Landsat Enhanced Thematic Mapper Plus (ETM+) bands 1-5 and 7 (instead of deriving these multispectral spectra by means of the multispectral unmixing methods that were mentioned above). For each selected hyperspectral endmember-spectrum, this yields a multispectral endmember-spectrum with six wavelengths, corresponding to the centers of these bands. These multispectral endmember-spectra are used as mentioned in the proposed algorithm.

ISBN 978-0-9928626-7-1 © EURASIP 2017

Also, for a fixed number of endmembers, 10 runs are performed to reduce the random effect of the results. The next section contains the means of the results obtained in these runs.

4.3. Results and discussion

We hereafter compare the performance of the designed and three NMF-based literature methods for the generated data. These literature methods are: modified NMF [5], standard multiplicative NMF [2], [6], [7], and minimum volume constrained NMF (MVC-NMF) [15]. The latter two NMF-based literature methods do not make use of multispectral data and are only applied to generated hyperspectral images, while the proposed and the literature modified NMF methods take multispectral data into account.
All tested methods are initialized with the same matrix values.
The maximum number of iterations used in each tested algorithm is set to 1000.
The next figures show the mean values of the considered performance criteria for each algorithm.

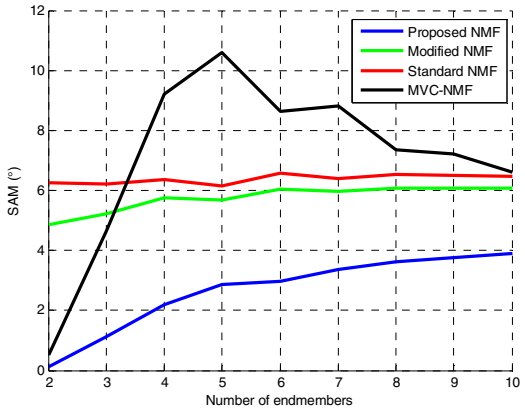

Figure 1. Mean SAM values (°).

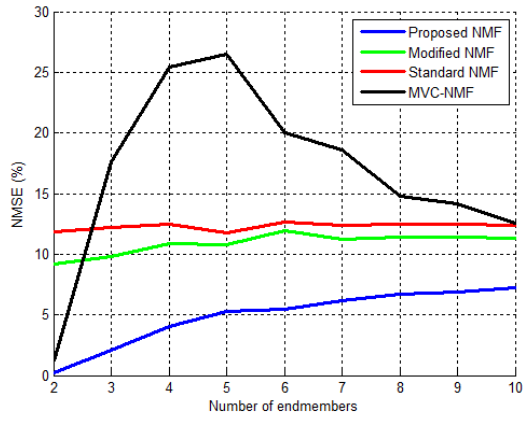

Figure 2. Mean NMSE values (%).

Globally, the above figures prove that the proposed algorithm yields much better performance than all other used literature methods. For each number of endmembers, the mean of SAM is less than 4° for the proposed approach, while this mean of SAM is, globally, close to or greater than 6° for the tested literature methods. Also, for each number of endmembers, the used literature methods achieve, globally, a mean of the NMSE values greater than 10%, while the proposed algorithm provides a mean of the NMSE values around 5%.
As stated throughout this paper, the proposed algorithm is designed to unmix hyperspectral images combined with multispectral data. However, some spectral variability between these two types of images may occur. In order to evaluate the robustness of the proposed algorithm to this spectral variability, another test is performed with the above generated synthetic data. In this test, and in order to add spectral variability between the hyperspectral and multispectral data, a random reflectance uniformly distributed between -5% and 5% of the original value is added to each sample of each created multispectral endmember-spectrum whereas the hyperspectral image is not modified. This addition is carried out while ensuring the nonnegativity constraint. The proposed and the literature modified NMF methods are applied to these modified synthetic data, and the mean values of the considered performance criteria are reported in the next tables.

Table I. Mean SAM values (°) obtained with spectral variability.

Number of endmembers	Proposed NMF	Modified NMF
2	**1.14**	4.24
3	**1.77**	5.31
4	**3.14**	6.12
5	**3.71**	6.51
6	**4.24**	6.61
7	**4.56**	6.73
8	**4.94**	6.80
9	**5.25**	7.05
10	**5.54**	7.21

Table II. Mean NMSE values (%) obtained with spectral variability.

Number of endmembers	Proposed NMF	Modified NMF
2	**5.49**	10.67
3	**6.39**	11.73
4	**8.44**	13.47
5	**9.11**	13.75
6	**9.99**	14.35
7	**10.41**	14.82
8	**11.07**	14.99
9	**11.67**	15.19
10	**12.15**	15.48

ISBN 978-0-9928626-7-1 © EURASIP 2017

The above tables show that the proposed algorithm is robust to the spectral variability between the hyperspectral and multispectral data, with a limited loss of performance as compared to the performance obtained by the same method without spectral variability. These tables also show that the proposed approach is more robust, to this spectral variability, than the tested literature modified NMF method.

5. CONCLUSION

In this work, a new algorithm, based on nonnegative matrix factorization, was proposed for linear endmember spectra extraction from remote sensing hyperspectral images combined with multispectral data.

The proposed algorithm was used to unmix a hyperspectral image. Using new multiplicative update rules, it minimizes a new cost function that combines multispectral data and a spectral degradation model between these data and hyperspectral ones.

Our tests show that the proposed method significantly outperforms some literature ones, which do not take into account multispectral data. Also, this new approach is robust to spectral variability that may exist between the two types of considered remote sensing data.

The combination of hyperspectral and multispectral data is therefore very attractive for linear hyperspectral endmember-spectra extraction.

Future extensions of this work may consist in applying the proposed algorithm to other synthetic data and also to real hyperspectral and multispectral images.

6. 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*, vol. 5(2), pp. 354-379, 2012.

[2] A. Cichocki, R. Zdunek, A.H. Phan, and S.-I. Amari, *Nonnegative matrix and tensor factorizations. Applications to exploratory multi-way data analysis and blind source separation*, Wiley, Chichester, UK, 2009.

[3] N. Yokoya, T. Yairi, and A. Iwasaki, "Coupled Nonnegative Matrix Factorization Unmixing for Hyperspectral and Multispectral Data Fusion," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50(2), pp. 528-537, 2012.

[4] M.S. Karoui, Y. Deville, F.Z. Benhalouche, and I. Boukerch, "Hypersharpening by Joint-Criterion Nonnegative Matrix Factorization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55(3), pp. 1660-1670, 2017.

[5] M.S. Karoui, S. Hosseini, Y. Deville, and A. Ouamri, "Modified Nonnegative Matrix Factorization for Endmember Spectra Extraction from Highly Mixed Hyperspectral Images Combined with Multispectral Data," *in Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing* (IEEE ICASSP 2017), New Orleans, USA, 2017.

[6] D.D. Lee and H.S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, pp. 788-791, 1999.

[7] D.D. Lee and H. S. Seung, "Algorithms for Non-Negative Matrix Factorization," *Advances in Neural Information Processing Systems*, vol. 13, pp. 556-562, MIT Press, 2001.

[8] M.S. Karoui, Y. Deville, S. Hosseini, and A. Ouamri, "Blind spatial unmixing of multispectral images: New methods combining sparse component analysis, clustering and non-negativity constraints," *Pattern Recognition*, vol. 45(12), pp. 4263-4278, 2012.

[9] M.S. Karoui, Y. Deville, S. Hosseini, and A. Ouamri, "A New Spatial Sparsity-Based Method for Extracting Endmember Spectra from Hyperspectral Data with Some Pure Pixels," *In Proc. of the IEEE International Conference on Geoscience and Remote Sensing Symposium* (IEEE IGARSS 2012), pp. 3074-3077, Munich, Germany, 2012.

[10] M.S. Karoui, Y. Deville, S. Hosseini, and A. Ouamri, "Blind unmixing of remote sensing data with some pure pixels: extension and comparison of spatial methods exploiting sparsity and nonnegativity properties," *In Proc. of the 8th IEEE International Workshop on Systems, Signal Processing and their Applications* (IEEE WoSSPA 2013), pp. 42-49, Algiers, Algeria, 2013.

[11] K.B. Petersen and M.S. Pedersen, *The matrix Cookbook*, Technical Univ. Denmark, 2008 [Online]. Available: URL:http://matrixcook-book.com/

[12] C. De Boor, *A Practical Guide to Splines*, Springer-Verlag, NY, USA, 2001.

[13] D.C. Heinz, C.I. Chang, "Fully Constrained Least Squares Linear Spectral Mixture Analysis Method for Material Quantification in Hyperspectral Imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39(3), pp. 529-545, 2001.

[14] R.N. Clark, G.A. Swayze, R. Wise, E. Livo, T. Hoefen, R. Kokaly, and S.J. Sutley, *USGS digital spectral library splib06a*. U.S. Geological Survey, Digital Data Series, 231, 2007.

[15] L. Miao and H. Qi, "Endmember Extraction From Highly Mixed Data Using Minimum Volume Constrained Nonnegative Matrix Factorization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45(3), pp. 765-777, 2007.

ISBN 978-0-9928626-7-1 © EURASIP 2017