

# DETECTION AND AREA ESTIMATION FOR PHOTOVOLTAIC PANELS IN URBAN HYPERSPSPECTRAL REMOTE SENSING DATA BY AN ORIGINAL NMF-BASED UNMIXING METHOD

Moussa Sofiane Karoui<sup>1,2,3</sup>, Fatima Zohra Benhalouche<sup>1,2,3</sup>, Yannick Deville<sup>2</sup>, Khelifa Djerriri<sup>1</sup>,  
Xavier Briottet<sup>4</sup> and Arnaud Le Bris<sup>5</sup>

<sup>1</sup> Centre des Techniques Spatiales, Arzew, Algeria

<sup>2</sup> IRAP, Université de Toulouse, UPS-OMP, CNRS, CNES, Toulouse, France

<sup>3</sup> LSI, Université des Sciences et de la Technologie d'Oran Mohamed Boudiaf, Oran, Algeria

<sup>4</sup> ONERA, The French Aerospace Lab, Toulouse, France

<sup>5</sup> Univ. Paris-Est, LASTIG MATIS, IGN, ENSG, Saint-Mandé, France

{Sofiane.Karoui, Fatima.Benhalouche, Yannick.Deville}@irap.omp.eu, kdjerriri@cts.asal.dz,  
Xavier.Briottet@onera.fr, Arnaud.Le-Bris@ign.fr

## ABSTRACT

Hyperspectral remote sensing data offer unique opportunities for the characterization of land surface in urban areas. However, no hyperspectral-unmixing based studies have been conducted to automatically detect photovoltaic panels, which represent one of the important components of energy systems in such areas.

In this paper, a hyperspectral-unmixing based method is proposed to detect photovoltaic panels and to estimate their areas. This approach is based on an original multiplicative nonnegative matrix factorization (NMF) algorithm with some known photovoltaic panel spectra. The proposed method can be considered as a partial/informed NMF approach. Experiments are conducted on realistic synthetic and real data to evaluate the performance of the proposed approach. In both cases, obtained results show that the proposed method yields much better overall performance than a method from the literature.

**Index Terms**— Hyperspectral imaging, hyperspectral unmixing, partial/informed nonnegative matrix factorization, detection and area estimation, photovoltaic panels

## 1. INTRODUCTION

The number of operating renewable energy systems has been increasing rapidly in recent years, particularly in urban areas of developed countries [1]. Photovoltaic panels constitute a significant portion of such systems. Indeed, these solar installations can generate greener energy from non-polluting resources, which has led government agencies, electricity grid operators and decision makers to encourage them by funding and tax reduction. Therefore, and to prevent malicious frauds with these alternative

energies, many organizations are increasingly interested in detailed information about these solar installations including their localization and energy production [1].

One of the methods for obtaining this detailed information, at least for the localization, on photovoltaic panels consists in making field surveys. However, this approach is costly and time consuming, which leads to the use of other faster and less expensive methods such as the integration of remote sensing data.

Recently, some approaches, using remote sensing data, were proposed for the automatic detection of photovoltaic installations and their localization [1], [2]. These methods use high spatial resolution airborne/spaceborne RGB images, therefore with a limited number of spectral bands. Nevertheless, such data do not allow an effective detection of photovoltaic panels mainly due to their material properties. Indeed, depending on the irradiation configurations, specular reflections can modify the visual properties of these panels, which makes their detection difficult.

To overcome the above limitations, high spectral resolution hyperspectral data may be considered. In fact, hyperspectral imaging sensors simultaneously gather images from hundreds of narrow and contiguous bands of the electromagnetic spectrum, which allows precise material recognition [3]. However, and principally due to the low spatial resolution of these sensors, mixed pixels, which are characterized by mixtures of spectra of more than one pure material (also called endmember), may occur in the acquired images and prevent direct detection of endmembers without advanced processing to unmix these spectra.

Linear Spectral Unmixing (LSU) is one of the most used techniques for processing hyperspectral data. It aims at linearly unmixing all spectra into a set of endmember spectra, and a collection of associated abundance fractions [4].

---

This work was funded by the French ANR project on "HYperspectral imagery for Environmental urban Planning" (HYEP, no. ANR 14-CE22-0016-01).

In this paper, a hyperspectral-unmixing method is proposed to automatically detect photovoltaic panels and to estimate their areas. This approach is based on an original algorithm with known photovoltaic panel spectra, which has relationships with multiplicative Nonnegative Matrix Factorization (NMF) [5]. Therefore, the proposed method can be considered as a partial/informed NMF technique [6], [7].

Experiments based on realistic synthetic and real data are conducted to evaluate the performance of the proposed approach and of the standard multiplicative NMF one.

This paper is structured in the following manner. The mathematical data model used in the proposed approach is defined in Section 2. The designed algorithm is presented in Section 3. In Section 4, the conducted experiments are described. Section 5 concludes this paper.

## 2. MATHEMATICAL DATA MODEL

As mentioned above, in this section the mathematical data model, used in the LSU techniques for unmixing hyperspectral remote sensing images, is described. Each observed nonnegative reflectance pixel spectrum in considered data is supposed to be a linear mixture of the nonnegative reflectance endmember spectra contained in that pixel. As a result, the entire observed nonnegative reflectance hyperspectral image  $X \in R_+^{N \times K}$  can be modeled, in matrix form, as [8]

$$X = A S, \quad (1)$$

where each row of the matrix  $X$  represents one spectral band of the considered hyperspectral image. The  $K$  pixels of this image are restructured as a one-dimensional array.  $N$  represents the number of spectral bands of that image. Each column of the matrix  $A \in R_+^{N \times L}$  represents one endmember spectrum and each row of the matrix  $S \in R_+^{L \times K}$  represents all nonnegative abundance fractions, in all pixels, of one endmember. These coefficients must satisfy the well-known abundance sum-to-one constraint [4].  $L$  represents the number of endmembers.

As suggested above, the proposed approach assumes that some photovoltaic panel spectra are known and that they are similar to those involved in the observed data. Therefore, the matrix  $A$  can be expressed as

$$A = A_1 + A_2, \quad (2)$$

where the matrix  $A_1 \in R_+^{N \times L}$  contains, in its first  $L_1$  columns, the known photovoltaic panel spectra. The rest of this first matrix is composed of zeros. The matrix  $A_2 \in R_+^{N \times L}$  contains, in its last  $L_2$  columns, the other, unknown, endmember spectra and the rest of this second matrix is composed of zeros. Obviously,  $L = L_1 + L_2$ . Consequently, Equation (1) becomes

$$X = (A_1 + A_2) S = A_1 S + A_2 S. \quad (3)$$

In the latter equation, the unknown (to be estimated by the proposed algorithm, described hereafter) matrices are  $A_2$  and, in particular,  $S$ .

## 3. PROPOSED APPROACH

The method proposed to automatically detect photovoltaic panels in urban hyperspectral remote sensing data is based on the NMF-unmixing technique [5]. The variables involved in this method consist of two nonnegative matrices,  $\tilde{A}_2$  and  $\tilde{S}$  that respectively aim at estimating  $A_2$  and  $S$ , such that

$$X \approx A_1 \tilde{S} + \tilde{A}_2 \tilde{S}. \quad (4)$$

Thus, the proposed method can be considered as a partial/informed NMF one [6], [7]. The designed algorithm minimizes the following cost function

$$J = \frac{1}{2} \|X - A_1 \tilde{S} - \tilde{A}_2 \tilde{S}\|_F^2 \quad (5)$$

by using iterative and multiplicative gradient-based update rules.  $\|\cdot\|_F$  represents the Frobenius norm.

To easily obtain the gradient expressions, the cost function  $J$  is rewritten as follows

$$J = \frac{1}{2} \text{Tr}(X X^T) - \text{Tr}(A_1 \tilde{S} X^T) - \text{Tr}(\tilde{A}_2 \tilde{S} X^T) + \text{Tr}(\tilde{A}_2 \tilde{S} \tilde{S}^T A_1^T) + \frac{1}{2} \text{Tr}(A_1 \tilde{S} \tilde{S}^T A_1^T) + \frac{1}{2} \text{Tr}(\tilde{A}_2 \tilde{S} \tilde{S}^T \tilde{A}_2^T), \quad (6)$$

where  $\text{Tr}(\cdot)$  and  $(\cdot)^T$ , respectively, denote the matrix trace and the matrix transpose. Considering the properties provided in [9], the gradient expressions of  $J$ , with respect to the two matrix variables, are in matrix form

$$\frac{\partial J}{\partial \tilde{A}_2} = -X \tilde{S}^T + A_1 \tilde{S} \tilde{S}^T + \tilde{A}_2 \tilde{S} \tilde{S}^T, \quad (7)$$

$$\frac{\partial J}{\partial \tilde{S}} = -A_1^T X - \tilde{A}_2^T X + \tilde{A}_2^T A_1 \tilde{S} + A_1^T \tilde{A}_2 \tilde{S} + A_1^T A_1 \tilde{S} + \tilde{A}_2^T \tilde{A}_2 \tilde{S}. \quad (8)$$

Since the proposed minimization algorithm is a gradient-based one, the following iterative update rule is considered

$$\theta \leftarrow \theta - \varphi_\theta \odot \frac{\partial J}{\partial \theta}, \quad (9)$$

where  $\odot$  represents the element-wise multiplication.  $\theta$  denotes one of the two matrix variables ( $\tilde{A}_2$  or  $\tilde{S}$ ), and  $\varphi_\theta$  is a learning rate in matrix form. To guarantee the nonnegativity constraint, an iterative and multiplicative rule is derived, from the above gradient-based one, in the following manner: as it appears in (7) and (8), the gradient expressions of  $J$  with respect to  $\theta$  can be written as the difference of two nonnegative functions such that  $\frac{\partial J}{\partial \theta} =$

$\frac{\partial J^+}{\partial \theta} - \frac{\partial J^-}{\partial \theta}$ . The nonnegative function  $\frac{\partial J^+}{\partial \theta}$  corresponds to

the terms of (7) or (8) preceded by a plus sign, whereas  $\frac{\partial J^-}{\partial \theta}$  corresponds to the terms preceded by a minus sign. The nonnegativity constraint can be satisfied by initializing  $\theta$  with a nonnegative value and choosing the value of the learning rate matrix  $\varphi_\theta$  according to

$$\varphi_\theta = \theta \oslash \frac{\partial J^+}{\partial \theta}, \quad (10)$$

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

$$\theta \leftarrow \theta \odot \frac{\partial J^-}{\partial \theta} \oslash \frac{\partial J^+}{\partial \theta}. \quad (11)$$

Accordingly, the final proposed iterative and multiplicative update rules for the  $\tilde{A}_2$  and  $\tilde{S}$  matrix variables read

$$\tilde{A}_2 \leftarrow \tilde{A}_2 \odot X \tilde{S}^T \oslash (A_1 \tilde{S} \tilde{S}^T + \tilde{A}_2 \tilde{S} \tilde{S}^T + \epsilon), \quad (12)$$

$$\tilde{S} \leftarrow \tilde{S} \odot (A_1^T X + \tilde{A}_2^T X) \oslash (\tilde{A}_2^T A_1 \tilde{S} + A_1^T \tilde{A}_2 \tilde{S} + A_1^T A_1 \tilde{S} + \tilde{A}_2^T \tilde{A}_2 \tilde{S} + \epsilon). \quad (13)$$

In the latter expressions,  $\epsilon$  (a very small and positive value) is added to the denominator of each update rule to avoid division by zero. It should here be noted that the above update rules guarantee the nonnegativity constraint (when the considered matrix variables are nonnegatively initialized) and experimentally yield a descent algorithm, which is such that  $J$  decreases throughout its optimization.

As standard NMF methods, the designed algorithm is not assured to provide a unique result and its convergence point depends on its initialization. Thus, and to avoid random initialization (from the viewpoint of the proposed algorithm) of the considered variables, initial estimates of unknown endmember spectra (stacked in the last  $L_2$  columns of matrix  $\tilde{A}_2$ ) are calculated by the Vertex Component Analysis (VCA) method [4]. Also, each initial value of the matrix  $\tilde{S}$  is set to  $1/L$ .

An additional constraint to be considered is the abundance sum-to-one property. To this end, the technique described in [10] is used.

The adaptation of the considered matrix variables is stopped when the number of iterations reaches a predefined maximum value.

#### 4. EXPERIMENTAL RESULTS

Tests using realistic synthetic and real data are conducted to evaluate the performance of the designed algorithm. Also, the obtained results are compared to those of the standard Lee and Seung multiplicative NMF method [5].

##### 4.1. Tested data

Four spectra of urban environment components (photovoltaic panel, tiles, grass and tree) are used to create a

realistic synthetic hyperspectral image according to the considered linear mixing model (1), except that some “random noise” is first added to the photovoltaic panel spectrum in each image pixel, in order to validate the robustness of the tested methods to spectral variability. The above four pure spectra contain 214 samples, ranging from 0.4 to 2.5  $\mu\text{m}$ , and are obtained by ground measurements using a spectrometer. The abundance fractions are created from a land cover classification of a real urban image, with the above-mentioned four materials, by averaging pixel classification values on a non-overlapping sliding window.

Two real sub-images from a full hyperspectral scene are also used in the conducted tests. The airborne hyperspectral image of 1.6m spatial resolution and 214 spectral bands covers a part of Toulouse (France) urban area.

In the conducted tests, the above ground-measured photovoltaic panel spectrum is considered as a known spectrum (i.e.  $L_1=1$ ) for both realistic synthetic and real data.

##### 4.2. Performance evaluation criteria

For the realistic synthetic data, the Normalized Mean Square Error (NMSE), between the original and estimated photovoltaic panel abundance fractions, is used to evaluate the performance of the tested methods. A smaller value of this criterion indicates a better detection /area estimation for photovoltaic panels.

For the real data, and in addition to a visual inspection of the obtained results, areas of photovoltaic panels, computed from the obtained thresholded abundance fractions (to avoid false detections), are compared with those manually digitized on very high spatial resolution (0.1m) ortho-images of the same zones.

##### 4.3. Results and discussion

The proposed approach, called “Multi-Part-NMF” for “Multiplicative Partial NMF”, and the standard multiplicative NMF (Multi-NMF) technique [5] are applied to the above-defined data. These two tested methods are initialized with the same matrix values.

The next table shows the values of the performance criterion considered for the realistic synthetic data.

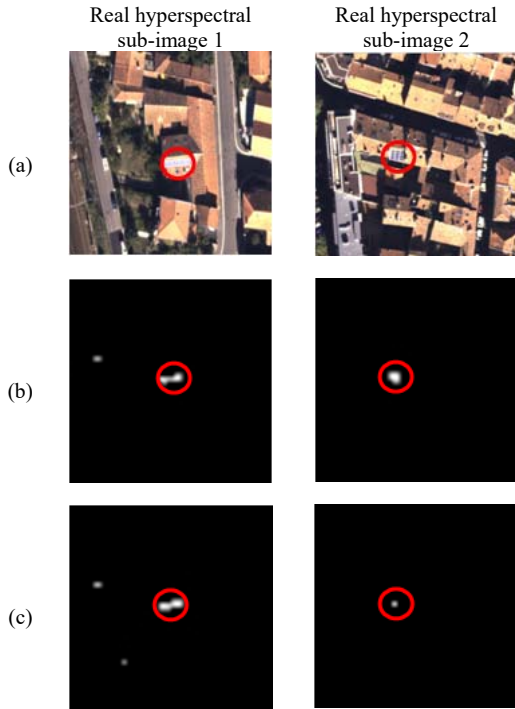
**Table I.** Photovoltaic panel NMSE values (%).

Multi-Part-NMF	Multi-NMF
<b>16.62</b>	81.76

The above table shows that the proposed approach achieves much better detection of photovoltaic panels than the standard multiplicative NMF algorithm. Indeed, this table reveals that the proposed Multi-Part-NMF algorithm is able to reach 16.62% NMSE for photovoltaic panels, whereas the literature Multi-NMF technique attains 81.76%.

The next figure shows the ortho-image and estimated (by the two tested algorithms) photovoltaic panel abundance fraction maps for the two considered real hyperspectral sub-

images. This figure confirms, by visual inspection, that the proposed algorithm is capable to better detect photovoltaic panels than the tested literature technique.



**Figure 1.** (a) Ortho-image. Estimated photovoltaic panel abundance fraction maps: (b) Multi-Part-NMF, (c) Multi-NMF. Red circles surround photovoltaic panels.

The table below shows the photovoltaic panel areas obtained from the two real hyperspectral sub-images.

**Table II.** Photovoltaic panel areas (m<sup>2</sup>).

	Manual digitization	Multi-Part-NMF	Multi-NMF
Sub-image 1	15	<b>15.27</b>	21.26
Sub-image 2	12	<b>13.06</b>	1.10

The above table shows that the proposed algorithm is able to automatically estimate photovoltaic panel areas much more accurately than the tested literature algorithm.

## 5. CONCLUSION

In this investigation, an original unmixing-based method was proposed to automatically detect photovoltaic panels in urban hyperspectral remote sensing data and to estimate their areas. The designed algorithm, which uses iterative and multiplicative update rules, is based on nonnegative matrix factorization (NMF) with known photovoltaic panel spectra. Compared with the standard multiplicative NMF technique, and according to the obtained results, the proposed algorithm proves to be very attractive in the framework of

the detection of photovoltaic panels in urban hyperspectral remote sensing data, and can be of great help to electricity grid operators and decision makers.

Future extensions of this investigation will particularly consist in improving the initialization stage of the proposed algorithm by using advanced LSU techniques to set initial unknown endmember spectra. Other extensions of this work will consist in performing further tests, by applying the proposed approach to other synthetic/real data.

## 6. REFERENCES

- [1] J.M. Malof, K. Bradbury, L.M. Collins and R.G. Newell, "Automatic detection of solar photovoltaic arrays in high resolution aerial imagery," *Applied Energy*, vol. 183, pp. 229-240, 2016.
- [2] S. Puttemans, W.V. Rans and T. Goedemé, "Detection of photovoltaic installations in RGB aerial imaging: A comparative study," *Proc. GEOBIA 2016*, Enschede, The Netherlands, 2016.
- [3] C.-I Chang, *Hyperspectral data processing: Algorithm design and analysis*, Wiley, 2013.
- [4] 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.
- [5] 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.
- [6] A. Limem, G. Delmaire, M. Puigt, G. Roussel and D. Courcot, "Non-negative matrix factorization using weighted beta divergence and equality constraints for industrial source apportionment," *Proc. IEEE International Workshop on Machine Learning for Signal Processing* (IEEE MLSP 2013), Southampton, UK, 2013.
- [7] L. Tong, J. Zhou, Y. Qian, X. Bai and Y. Gao, "Nonnegative-Matrix-Factorization-Based Hyperspectral Unmixing with Partially Known Endmembers," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54(11), pp. 6531-6544, 2016.
- [8] 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," *Proc. IEEE International Conference on Geoscience and Remote Sensing Symposium* (IEEE IGARSS 2012), pp. 3074-3077, Munich, Germany, 2012.
- [9] K.B. Petersen and M.S. Pedersen, *The matrix Cookbook*, Technical Univ. Denmark, 2008 [Online]. Available: URL:<http://matrixcook-book.com/>
- [10] 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.