A method based on NMF dealing with intra-class variability for unsupervised hyperspectral unmixing

Charlotte Revel, Yannick Deville, Véronique Achard, Xavier Briottet

A METHOD BASED ON NONNEGATIVE MATRIX FACTORIZATION DEALING WITH INTRA-CLASS VARIABILITY FOR UNSUPERVISED HYPERSPECTRAL UNMIXING

Charlotte Revel[1,2], Yannick Deville[1]

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

Véronique Achard[2], Xavier Briottet[2]

[2]ONERA "The French Aerospace Lab"
2 Av. Édouard Belin, 31000 Toulouse, France

ABSTRACT

In hyperspectral imagery, unmixing methods are often used to analyse the composition of the pixels. Such methods usually supposed 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 proposed a new method based on the assumptions of a linear mixing model, that deals with within intra-class spectral variability. A new formulation of the linear mixing is proposed. It introduces not only a scaling factor but a complete representation of the spectral variability in the pure spectrum representation. In our model a pure material cannot be described by a single spectrum in the image but it can in a pixel. A method is presented to process this new model. It is based on a pixel-by-pixel Non-negative Matrix Factorization (NMF) method. The method is tested on a semi-synthetic set of data built with spectra extracted from a real hyperspectral image and mixtures of these spectra. Thus we demonstrate the interest of our method on realistic intra-class variabilities.

Index Terms— Hyperspectral unmixing, intra-class variability, pixel-by-pixel Non-negative Matrix Factorisation (NMF), real data set.

1. PROBLEM STATEMENT

In the framework of remote sensing the unmixing is a common way to deal with mixed pixels in hyperspectral images. This technique extracts subpixel information in hyperspectral images which are usually less spatially resolved than panchromatic or multispectral images. 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. Under this assumption, each observed reflectance spectrum, $\mathbf{x_p} \in \mathbb{R}^{L\times 1}$, can be written as follows:

$$\mathbf{x_p} = \sum_{m=1}^{M} c_{p,m}\mathbf{r_m} \quad \forall p \in \{1,\dots,P\} \tag{1}$$

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 $c_{p,m}$ is the associated coefficient. Reflectance spectra and coefficients are assumed to be nonnegative. Beside, 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:

$$\sum_{m=1}^{M} c_{p,m} = 1 \quad \forall p \in \{1,\dots,P\} \tag{2}$$

The above model is based on strong assumptions. Two of them can be singled out. First of all, the pure material spectra definition, indeed at macroscopic scale the spectrum of a same pure material extracted in two different locations of the image can vary. This can be due to illumination variations. Material can also be weathered, having mineral slight composition variations or being variously used [3]. This phenomenon, the so-called intra-class variability [4], then leads to reconsider the sum-to-one constraint as it is formulated in the linear mixing model. The unmixing as it is currently performed considers that all observed spectra can be reconstructed by the weighted sum of a single spectrum per class. And yet it was just pointed out that a single spectrum cannot describe all variations of a class of materials. 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 (1) a dependency of $\mathbf{r_m}$ 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) can be rewritten as:

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

where $\mathbf{r_m}(p)$ is the reflectance spectrum associated with the material m and the pixel p. The coefficients $c_{p,m}$ still verify the sum to one, so (2) is still verified.

This new model is an ill-posed problem. The features of the spectral variation inside the classes (sets of pure spectra associated with a same material) might bring additional information to further constrain the problem. If intra-class variability was only due to illumination variations, spectral variation could be modelled by a multiplying factor [5]. However in real data intra-class variability has a more complex structure.

To achieve an accurate unmixing in real data the intra-class variability has to be taken into account. The method proposed hereafter achieves unmixing by extracting M sets of endmember spectra. The Nonnegative Matrix Factorization has been extended and constrained to solve (3) under (2). Sec. 2 overviews three NMF unmix-

This work was partly supported by the French ANR Hyperspectral Imagery for Environmental Urban Planning (HYEP) project as from 2015.

ing methods from the classical to the above-mentioned one. 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 spectral reflectance matrix if only one spectrum per class is considered and $\mathbf{C} = [\mathbf{c_1}, \ldots, \mathbf{c_P}]^T$ the coefficient matrix. For all pixels p, $\mathbf{c_p} = [\mathbf{c_{p,1}}, \ldots, \mathbf{c_{p,M}}]^T$ 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 of 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, $\mathbf{R_P} = \begin{bmatrix} \mathbf{R}(1) \\ \vdots \\ \mathbf{R}(P) \end{bmatrix}$ the matrix containing all the pure spectra. Then Eq. (3) may be rewritten as follows:

$$\mathbf{x_p}^T = \mathbf{c_p}^T \mathbf{R}(p). \tag{5}$$

Let $\mathbf{C_P} \in \mathbb{R}^{P \times PM}$ be a block diagonal matrix, denoting the new coefficient matrix:

$$\mathbf{C_P} = \begin{bmatrix} \mathbf{c_1}^T & 0 \ldots 0 & \ldots & 0 \ldots 0 \\ 0 \ldots 0 & \mathbf{c_2}^T & \ldots & 0 \ldots 0 \\ & & \ddots & \\ 0 \ldots 0 & 0 \ldots 0 & \ldots & \mathbf{c_P}^T \end{bmatrix} \tag{6}$$

So Eq. (4) yields matrix writing of the mixing model (3):

$$\mathbf{X} = \mathbf{C_P R_P}. \tag{7}$$

The sum-to-one constraint, (2), is kept, so is the nonnegativity constraint. In each method hereafter develops the normalisation of $\mathbf{c_p}$ or $\mathbf{C_p}$ is performed after each update. Even with these constraints the equations (5) and (7) remain non-convex problems.

2.1. NMF

Nonnegative Matrix Factorization (NMF) has been adapted to solve remote sensing unmixing problems [6]. NMF aims at decomposing a matrix in a product of two nonnegative matrices, in our case the coefficient matrix, $\mathbf{C_P}$ and the reflectance matrix, $\mathbf{R_P}$. The necessary assumption of this method is the nonnegativity of the two searched matrices. The sum-to-one constraint can be added to NMF. The method then can be applied to the observation matrix $\mathbf{X}$. To perform NMF a cost function, the reconstruction error (RE), is minimised, however the obtained minima may be local ones. This point can be overcome with a good initialisation. This will be discussed in a later part (Sec. 3).

2.2. Unconstrained Pixel-by-pixel NMF (UP-NMF)

The classical way to carry out NMF to unmix data does not allow to extract one set of pure reflectance spectra per pixel. We therefore developed an extended version of Lin's standard NMF algorithm [7], so as to decompose each observed spectrum into pure spectra which are specific to the considered pixel. This extended version aims at minimising $J_1(p) = \frac{1}{2}\|\mathbf{x_p}^T - \mathbf{c_p}^T\mathbf{R}(p)\|_2^2$. We will show elsewhere that the resulting update of the estimated $\mathbf{c_p}$ and $\mathbf{R}(p)$ can be expressed as follows before their projection onto $\mathbb{R}^+$:

$$\begin{aligned} \mathbf{c_p}^T &\longleftarrow \mathbf{c_p}^T + \alpha_{c_p}(\mathbf{x_p}^T - \mathbf{c_p}^T\mathbf{R}(p))\mathbf{R}(p)^T \\ \mathbf{R}(p) &\longleftarrow \mathbf{R}(p) + \alpha_{R(p)}\mathbf{c_p}(\mathbf{x_p}^T - \mathbf{c_p}^T\mathbf{R}(p)) \end{aligned}$$

where α_{c_p} and $\alpha_{R(p)}$ are respectively the coefficient update step and the reflectance spectra one for the pixel p. From now, let us call this method Unconstrained Pixel-by-pixel NMF (UP-NMF). This update before the projection can also be written by using the estimated matrices $\mathbf{R_p}$ and $\mathbf{C_p}$. It yields to a similar writing to classical NMF:

$$\begin{aligned} \mathbf{C_P} &\longleftarrow \mathbf{C_P} + \alpha_{C_P}(\mathbf{X} - \mathbf{C_P R_P})\mathbf{R_P}^T \\ \mathbf{R_P} &\longleftarrow \mathbf{R_P} + \alpha_{R_P}\mathbf{C_P}^T(\mathbf{X} - \mathbf{C_P R_P}) \end{aligned}$$

except that the required entries of $\mathbf{C_p}$ are fixed to zero according to Eq. (6).

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 full constraint least square unmixing on the previously found $\mathbf{R}^{(0)}$ for instance. Both $\mathbf{R}(p)$ and $\mathbf{c_p}$ initializations will be discussed below (Sec. 3.3).

We can note that the coefficient and reflectance matrices are obtained with a scaling factor. Indeed $\mathbf{x_p}$ can be rewritten as :

$$\mathbf{x_p} = \sum_{m=1}^{M} \frac{1}{k_m(p)} c_{p,m} \times k_m(p)\mathbf{r_m}(p) \quad \forall p \in \{1, \ldots, P\} \tag{8}$$

with $k_m(p)$ a scaling factor. This point will be discussed in Sec. 3. Indeed it has to be taken into account during the result evaluation.

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.

2.3. 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:

$$J_2 = \frac{1}{2}\|\mathbf{X} - \mathbf{C_p R_p}\|_F^2 + \mu \sum_{m=1}^{M} Tr(Cov(\mathbf{R}_{\mathcal{C}_m})) \tag{9}$$

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 show that the resulting update of UP-NMF under this added constraint and before the $\mathbb{R}^+$ projection and zero forcing in $\mathbf{C_p}$ is:

$$\begin{aligned} \mathbf{C_P} &\longleftarrow \mathbf{C_P} + \alpha_{C_P}.(\mathbf{X} - \mathbf{C_P R_P})\mathbf{R_P}^T \\ \mathbf{R_P} &\longleftarrow \mathbf{R_P} + \alpha_{R_P}.(\mathbf{C_P}^T(\mathbf{X} - \mathbf{C_P R_P}) \\ &\qquad - \frac{2\mu}{P}(\mathbf{Id}_{PM} - \frac{1}{P}\mathbf{U})\mathbf{R_P}) \end{aligned}$$

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

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

3. TEST 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 well describe the scene: tile, vegetation (mainly plane trees) and asphalt. Due to its 1.8 m resolution, this image contains many pure spectra. Hence we were able to extract a large number of spectra to take into account various phenomena responsible for intra-class variability: illumination variation, various material weatherings... Some of these spectra are shown in Fig. 1, where atmospheric absorption bands were removed.

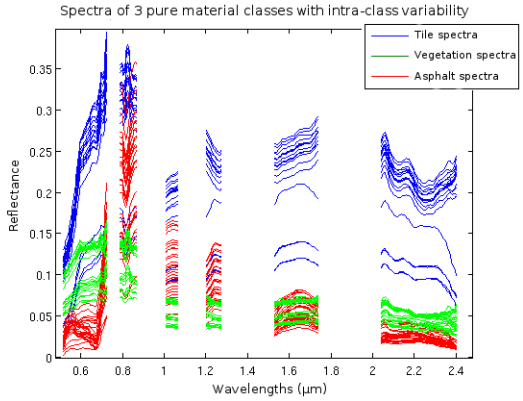

Fig. 1. Reflectance spectra of three pure material classes: Tile (blue), Vegetation (red) and Asphalt (green) illustrating the intra-class variability.

The tested data are semi-synthetic. The pure spectra used to create the mixed data (hereafter call reference spectra) are extracted from the above urban area image. Thus they describe realistic intra-class variabilities. Coefficients are randomly chosen while respecting the sum-to-one constraint. The mixing model is Eq. (3)'s one.

Tests were performed to evaluate the benefits of the developed method. They also allowed us to analyse the impact of initialisation on results. Various initial spectral matrices, $\mathbf{R}^{(0)}$ (cf. Sec. 2.3) were tested: (i) M spectra are randomly selected from observed data, (ii) the M purest spectra are extracted with a standard method [1] (N-FINDR, VCA, ...). 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 with a full constraint least square (FCLS) method.

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 angles between these two sets of spectra. Two other criteria were also computed. The first one is the reconstruction error (RE), that shows the effects of our method on the global reconstruction of the image. The second one is the mean square error computed on the coefficients. Computing these errors is possible for semi-synthetic data for which all the mixing parameters are known. Other criteria must be used if real data are considered, particularly if no accurate ground truth is available.

3.3. Results

In the evaluation of the results, we focused on two points: (i) the results of IP-NMF compared with standard and UP-NMF, (ii) the impact of the initialisation on the results.

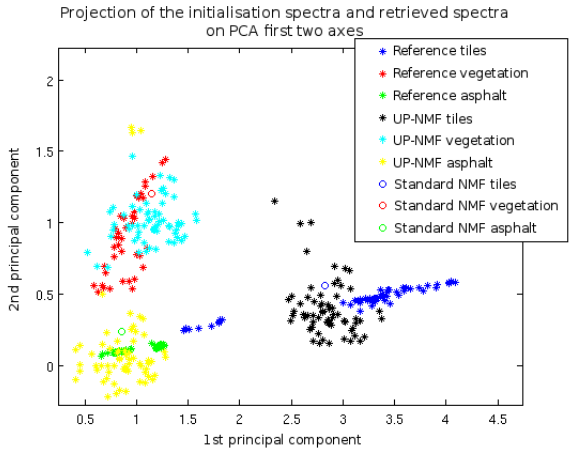

Fig. 2. Projection onto the first two PCA axes of initialisation spectra (red, green, blue stars), standard NMF spectra (red, green, blue circles) and UP-NMF spectra (cyan, yellow and black stars).

To analyse the general contribution of the IP-NMF, the initialisation parameters are fixed. $\mathbf{R}^{(0)}$ is built with the reflectance spectra extracted from the mixed data with the N-FINDR algorithm. As explained in Sec. 2.2, all the pixels are initialized with the same set of spectra. Initialisation of coefficients is carried out by setting all of them to the same value, $\frac{1}{M}$. The constraint parameter, μ, varied from 0 (spectra spreading) to 200 (convergence toward a single spectrum). The corresponding results are illustrated in Fig. 3 and Fig. 4. Fig. 2 shows the result of UP-NMF. As mentioned in Section 2.2 the main risk is the spectra spreading. It partly occur in Fig.2: the variabilities of the retrieved spectra are a bit high regarding the variability of the spectra used in the mixing. It reinforces our idea to limit the classes inertia. Results in Table. 1 confirms this observation. The average SAM of this method is the highest due to the poor reconstruction of some spectra. In this case (as for standard NMF) the close to 0 reconstruction error (RE) signification is relative. Indeed reconstruction error minimisation is what UP-NMF and standard NMF look for. The IP-NMF leads to Fig. 3 and Fig. 4. In the first one, the constraint parameter μ was fixed to 30, and to 100 in the second one. Dots localisation evolution shows the impact of the inertia constraint. If μ is too high, the method gets closer to standard NMF since it yields very compact sets of spectra for each class. In the case when μ is suitable, the widths of the scatter plots formed by the extracted spectra are comparable to those of the scatter plots formed by the reference spectra. It has to be noted that the spectra resulting from constrained method are somewhat shifted

with respect to the initialisation spectra scatter plots. Indeed as explained in Sec. 2.2, by Eq. (8) scaling factors are contained in the spectra. And hence dots can move onto the axes passing through the origin. Thus the dot position onto its class principal axis is contained in the scale factor. That is why we are $\mu = 30$ is a suitable parameter value: it allows to retrieve the variation orthogonal to the principal axis of the class. Since scaling factor does not affect SAM, computing these SAM allow to confirm this point. Table 1 shows improvements brought by IP-NMF.

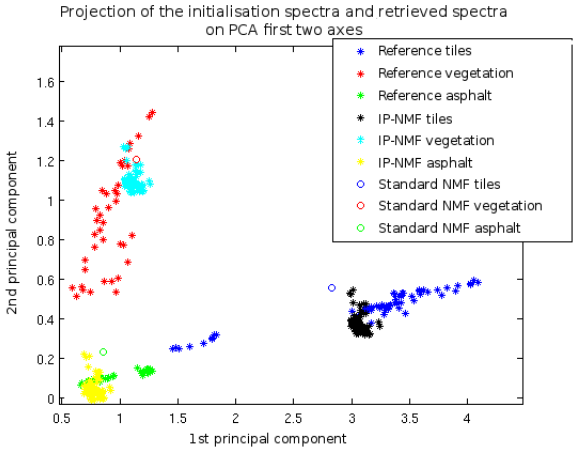

Fig. 3. Projection onto the first two PCA axes of initialisation spectra (red, green, blue stars), standard NMF spectra (red, green, blue circles) and IP-NMF spectra (cyan, yellow and black stars) with $\mu = 30$.

	N-FINDR	Standard NMF	UP-NMF	IP-NMF ($\mu = 30$)
RE	17.5	0.60	2×10^{-5}	2.94
SAM	7.72	7.73	9.44	5.47
Coefficient error (in %)	4.0	4.7	3.8	3.8
Times (in sec)	0.03	0.02	4.40	4.44

Table 1. Results of various unmixing methods.

The second point consists in analysing the impact of initialisation. We fixed μ parameter at 30 for IP-NMF. Scenarios described in Sec. 3.1 were carried out. Results of UP-NMF and IP-NMF were compared to standard NMF and a classical unmixing method (N-FINDR + FCLSU). It appears that a poor initialization leads to poor results for both the standard NMF UP-NMF and IP-NMF. Yet, compared with standard NMF, IP-NMF improves the average spectral angle error in every initialisation case. We also noted that if both the spectra and the coefficient are initialised too close to a local minimum the results of our method are close to this initialisation.

4. CONCLUSION

We have developed two new hyperspectral unmixing methods to address intra-class variability. This investigation is a first step in the development of a more robust method. Indeed the dependency on the initialisation is too strong. Besides the constraint can also be improved to better take data spatial variability into account. The

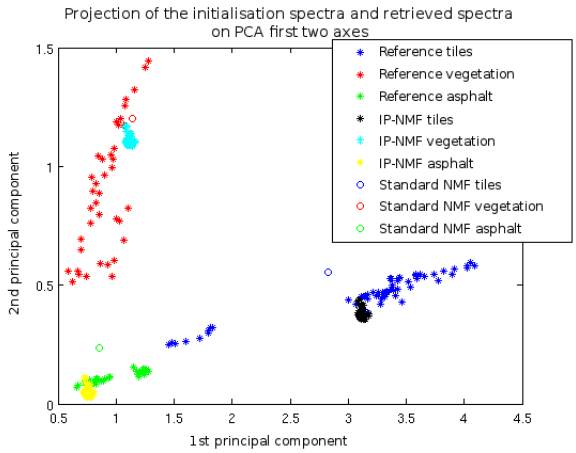

Fig. 4. Projection onto the first two PCA axes of initialisation spectra (red, green, blue stars), standard NMF spectra (red, green, blue circles) and IP-NMF spectra (cyan, yellow and black stars) with $\mu = 100$.

inertia limitation constraint can be discussed. Current work aims at developing an extended version of this approach both in terms of initialisation and constraint applied to the cost function. We also plan to work on ground truth to obtain good qualities ones. Then we will be able to carry out IP-NMF on real images and pertinently analyse results.

5. REFERENCES

[1] J.M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Qian 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, no. 2, pp. 354–379, Apr. 2012.

[2] Nirmal Keshava, "A survey of spectral unmixing algorithms," *Lincoln Laboratory Journal*, vol. 14, no. 1, pp. 55 – 78, 2003.

[3] S. Lacherade, C. Miesch, X. Briottet, and H. Le Men, "Spectral variability and bidirectional reflectance behaviour of urban materials at a 20 cm spatial resolution in the visible and near-infrared wavelengths. a case study over toulouse (france)," *International journal of remote sensing*, vol. 26, no. 17, pp. 3859–3866, 2005.

[4] A. Zare and K.C. Ho, "Endmember variability in hyperspectral analysis: Addressing spectral variability during spectral unmixing," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 95–104, Jan. 2014.

[5] J.M.P. Nascimento and J.M. Bioucas Dias, "Does independent component analysis play a role in unmixing hyperspectral data?," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 1, pp. 175–187, Jan. 2005.

[6] Yuntao Qian, Sen Jia, Jun Zhou, and A. Robles-Kelly, "Hyperspectral unmixing via sparsity-constrained nonnegative matrix factorization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4282–4297, Nov. 2011.

[7] Chih-Jen Lin, "Projected gradient methods for nonnegative matrix factorization," *Neural Computation*, vol. 19, no. 10, pp. 2756–2779, 2007.