



**HAL**  
open science

## A new hyperspectral unmixing method using co-registered hyperspectral and panchromatic images

Simon Rebeyrol, Yannick Deville, Véronique Achard, Xavier Briottet,  
Stéphane May

### ► To cite this version:

Simon Rebeyrol, Yannick Deville, Véronique Achard, Xavier Briottet, Stéphane May. A new hyperspectral unmixing method using co-registered hyperspectral and panchromatic images. Whispers 2019, Jan 2019, AMSTERDAM, Netherlands. hal-02409424

**HAL Id: hal-02409424**

**<https://hal.science/hal-02409424>**

Submitted on 13 Dec 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# A NEW HYPERSPECTRAL UNMIXING METHOD USING CO-REGISTERED HYPERSPECTRAL AND PANCHROMATIC IMAGES

*Simon Rebeyrol<sup>1,2</sup>, Yannick Deville<sup>1</sup>, Veronique Achard<sup>2</sup>, Xavier Briottet<sup>2</sup>, Stephane May<sup>3</sup>*

<sup>1</sup> IRAP, Université de Toulouse, UPS-CNRS-CNES, 14 Av. Edouard Belin, 31400 Toulouse, France

<sup>2</sup> ONERA The French Aerospace Lab, Département Optique et Techniques Associées (DOTA), 2 Av. Edouard Belin, 31055 Toulouse, France

<sup>3</sup> Centre national d'études spatiales (CNES), 18 Av. Edouard Belin, 31401 Toulouse Cedex 9, France

## ABSTRACT

In this paper, we propose a novel way of unmixing spaceborne hyperspectral images using a co-registered hyperspectral image and a high-spatial-resolution panchromatic image. The unmixing is achieved by a two-step approach. First, an extraction of pure pixels using a panchromatic homogeneity criterion is performed and a reconstruction error map is generated. Then, a local semi-supervised Non-negative Matrix Factorization (NMF) is performed on badly reconstructed areas in order to estimate material spectra that are not represented by pure pixels in the image. This approach is tested on real data for the first step, and on both synthetic and real data for the second step.

**Index Terms**— Hyperspectral, panchromatic, unmixing, spaceborne

## 1. INTRODUCTION

Earth remote sensing satellites, especially those designed for acquiring hyperspectral images (HS) [1], allow one to identify materials based on their specific spectral signatures. The main issue of spaceborne hyperspectral sensors is their ground sampling distance (GSD) which is typically in the range of 10 to 30 meters [2]. For such GSDs, material signatures, also known as endmembers, are therefore mixed into the pixels. The way of retrieving the pure material spectral signatures and their abundance fractions with very limited prior knowledge is called unsupervised hyperspectral unmixing and has been studied for several decades [3]. Most of the existing methods are based on the assumption of a linear mixing model. Additional constraints are typically assumed, such as the abundance sum-to-one constraint that induces interesting geometrical properties on the data and leads to a wide unmixing method family referred to as geometrical methods [4].

In such conditions, the observed data, minus the noise, are contained within a  $P - 1$  simplex formed by the endmembers, where  $P$  is the number of endmembers. Most of the geometrical methods use that property to find the endmembers in the observed data by, for example, making the

assumption that each material is represented by at least one pure pixel in the considered image. However, this assumption is often invalid in regard to the spaceborne sensor's spatial resolution, and the unmixing performance of such methods is therefore limited. Some well-known geometrical methods based on pure pixels are the Vertex Component Analysis (VCA), N-FINDR, the Fast Iterative Pixel Purity Index (FIPPI) and the Automated Target Generation Process (ATGP) algorithms. Other geometrical methods find the endmembers by minimizing the volume of the induced simplex. These methods mainly include the MVC-NMF and SISAL algorithms. Further information is available in [4]. Other families of unsupervised methods also exist which do not require the pure pixel assumption, *e.g.* the statistical methods [4] and those based on Non-negative Matrix Factorization (NMF) [5]. The NMF field offers a well-studied framework to estimate both the abundance fractions and the endmembers by iteratively minimizing a cost function. Several strategies exist, the well-known Lee and Seung approach [6] with multiplicative update rules or the projected gradient strategy with additive update rules [7]. The cost function can be penalized by adding constraint terms to force the abundance fractions or spectra to satisfy some hypotheses such as spectral smoothness [8] or abundance sparsity [9]. The NMF algorithms are however extremely sensitive to their initialization. In all these methods, only the hyperspectral data were used to estimate the endmembers and their abundance fractions in each pixel.

Recently, a satellite project named HYPXIM [10] has been proposed by the French Centre National d'Études Spatiales (CNES) shipping two co-registered cameras. The first camera is a hyperspectral sensor with an 8 m GSD and an 11 nm spectral resolution. The second one is a panchromatic (PAN) camera with a 2 m GSD. The spatial resolution of the panchromatic camera allows one to sense small objects in the scene and therefore provides more spatial information than the hyperspectral sensor. This additional information should be of major interest for the challenging blind unmixing problem.

In this paper, we propose a novel method taking in account both PAN and HS images to estimate the endmembers.

This method is composed of two stages. The first stage, called Homogeneity-Based Endmember Extraction (HBEE), aims at finding all pure pixels in the HS image based on a homogeneity criterion applied to the co-registered PAN image in order to extract the pure pixel spectra. Then a clustering method is applied to these pure pixel spectra in order to gather spectra corresponding to the same material. The second stage aims at estimating the remaining material spectra that are not represented by pure pixels by applying local semi-supervised NMF (LS-NMF). Results are provided for both stages, including for real data. Finally, a conclusion is proposed.

## 2. MIXING MODEL

In this work, we use the linear mixing model described in (1):

$$\vec{y}_i = \sum_{p=1}^P x_{ip} \vec{s}_p + \vec{w} \quad x_{ip}, \vec{s}_p \geq 0 \quad (1)$$

where  $P$  is the number of endmembers in the scene,  $x_{ip}$  are the abundance fractions,  $\vec{s}_p$  are the pure material signatures,  $\vec{w}$  represents the noise, and  $\vec{y}_i$  an observed pixel spectrum. However, we do not use the sum-to-one constraint since, for real data, this assumption is not fulfilled, *e.g.* due to the spectral variability of each material [11].

## 3. PROPOSED APPROACH

### 3.1. Homogeneity-Based Endmember Extraction (HBEE)

A classical first stage in the unmixing process, before estimating the abundance fractions, is to find the pure pixels. We perform this stage in several steps, first, by applying a homogeneity-based criterion on the panchromatic pixels that cover the considered co-registered HS pixel. Let  $A_{ij}$  be the panchromatic sub-image corresponding to the  $y_{ij}$  hyperspectral pixel. It is reasonable to assume that if the sub-image  $A_{ij}$  has a homogeneous grey level then the corresponding  $y_{ij}$  is a pure pixel. The limitation of such a hypothesis is that different materials may have the same grey level, for example in shadowed areas. Nevertheless, this case is fairly unusual in the same sub-image  $A_{ij}$ . The homogeneity criterion we choose to apply is the max-min distance

$$h_{ij} = \max(A_{ij}) - \min(A_{ij}). \quad (2)$$

The criterion is strict since any panchromatic outlier will increase this distance. In order to detect pure pixels, one must choose a threshold value  $t_h$  under which  $A_{ij}$  is considered to be homogeneous and then, the corresponding pixel  $y_{ij}$  pure. This method is applied on all the sub-images  $A_{ij}$  to build a set  $l_h$  of spectra corresponding to the pure pixels.

This method is likely to extract several pure pixels that represent the same material. A subsequent clustering step, using the spectral angle (SA), is therefore required to derive

a single spectrum from all spectra corresponding to the same material. To this end, the collected spectra  $l_h$  are first sorted into a list  $L_h$  according to their PAN homogeneity criterion value  $h_{ij}$ . The user needs to choose a spectral angle threshold  $t_{sa}$ . The clustering operates as follows. During a first pass, *i.e.* for  $k = 1$ , it creates a new cluster  $C_k$  defined by the first spectrum  $\vec{s}_k$  of  $L_h$ . An SA value is computed between  $\vec{s}_k$  and each of the other remaining spectra in  $L_h$ . If this distance is below  $t_{sa}$ , the corresponding spectrum is then assigned to the current cluster  $C_k$  and removed from  $L_h$ . New passes, *i.e.* with increasing  $k$ , are then performed on the remaining list  $L_h$  in the same way, until that list is empty. The generated number of clusters gives us a strong indication about how many materials are represented by pure pixels in the image. The endmember matrix  $\hat{S}_h$  is constructed with the spectrum of each cluster having the best homogeneity criterion value.

Then we estimate the abundance fraction matrix  $\hat{X}_h$  with the Non-Negative constrained Least Square algorithm (NNLS) [12] in order to calculate the normalized reconstruction error map  $R$  between the original hyperspectral image  $Y$  and the reconstructed one  $\hat{Y}$ , whose entries read

$$R_{ij} = \frac{\|y_{ij} - \hat{y}_{ij}\|}{\|\hat{y}_{ij}\|}. \quad (3)$$

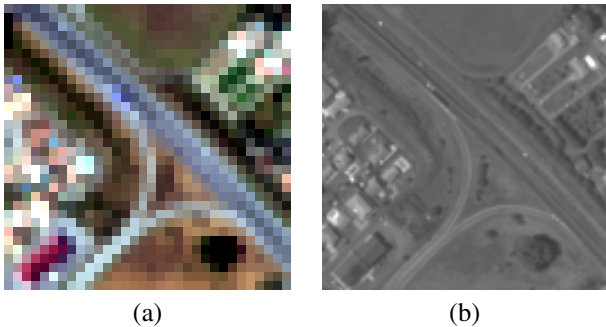
### 3.2. Local Semi-supervised Non-negative Matrix Factorization (LS-NMF)

Since there is no guarantee that all the endmembers are represented by pure pixels and then are extracted by the HBEE method, the reconstruction error map has high values in areas where the non-extracted material spectra are located. Such spatial information allows us to locally unmix the remaining endmembers by making the assumption that there is only one material to estimate in each poorly reconstructed area. This unmixing stage only aims at estimating the local endmember. A classical NMF [6] with multiplicative update rules is successively applied to each area which yielded a high reconstruction error. The endmember matrix is initialized by concatenating the endmembers  $\hat{S}_h$  extracted with the HBEE stage, the estimated spectra from previous areas treated during the LS-NMF stage and the mean spectrum over the considered local area. For the initial abundance fraction matrix, we make the assumption that the unknown material is far more abundant than the others in the area. We therefore build this initial abundance fraction matrix  $X_{init}$  as follows. We use the mean abundance fraction, previously estimated to reconstruct the image, of each material in a strip of pixels situated around the considered area to take into account the two most abundant nearby materials. We set these two abundances in the initial matrix  $X_{init}$  and divide them by an arbitrary  $a$  value. The other material abundances are set to a small value. We then force the abundance fraction of the unknown material to conform, only here, to the sum-to-one constraint. After the NMF run, the new estimated endmember is added to the matrix  $\hat{S}_h$

and this stage is repeated on all the local badly reconstructed areas of the updated reconstruction error map. Once all the poorly reconstructed areas have been treated, we estimate the abundance fractions on the entire image using the NNLS algorithm and the endmembers estimated from both the HBEE and LS-NMF stages in order to complete the unmixing process.

#### 4. SIMULATIONS OF HYPXIM-P DATA

The image used to test the proposed unmixing method was acquired by the AisaFenix<sup>1</sup> HS sensor over the Mauzac village (France) in August 2017. The GSD of this acquisition is equal to 55 cm. Since our investigation is focused on spaceborne hyperspectral data, a multi-step simulation has been performed to simulate the HYPXIM-P image from the airborne radiance image. First, an atmospheric correction is applied on the AisaFenix image using COCHISE [13] to get a reflectance image with a 55 cm GSD. A direct radiative transfer model COMANCHE [13] is then applied to obtain a radiance image at the satellite altitude. We then produce the 8 m and 2 m GSD hyperspectral radiance images by means of spatial integration knowing the expected point-spread function and noise parameters. A last COCHISE step is applied to simulate the observed reflectance image. A co-registered PAN image is generated by summing over the  $[0.4 - 0.8] \mu m$  domain the satellite simulated 2 m GSD radiance image (see Fig. 1).



**Fig. 1.** (a) 8 m GSD hyperspectral reflectance image, (b) 2 m GSD panchromatic radiance image.

A seventeen-class spectral ground truth (GT) is generated from the 55 cm GSD hyperspectral reflectance image by means of photo-interpretation. Each class of materials contains several dozens of spectra. Some classes may depict the same material, as the intraclass spectral variability led us to split materials into two classes. The presence of some objects, such as cars and pools, can cause issues in the unmixing process due to their sparse distribution, their small size and/or their low reflectance. In order to limit their influence in any

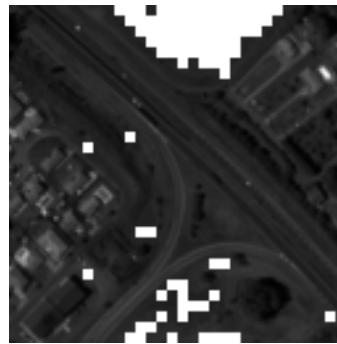
unmixing method, we chose to add representative spectra of cars and pools in the endmember set.

## 5. RESULTS

This section presents results for the HBEE and LS-NMF stages. The performance criteria used to assess the performance of the algorithm are the mean, over all estimated endmembers, of their spectral angle (SA) with respect to the closest spectra in the ground truth library, and the spatial mean of the Normalized Root Mean Square Error (NRMSE) between the original and re-built image pixel spectra.

### 5.1. HBEE stage results

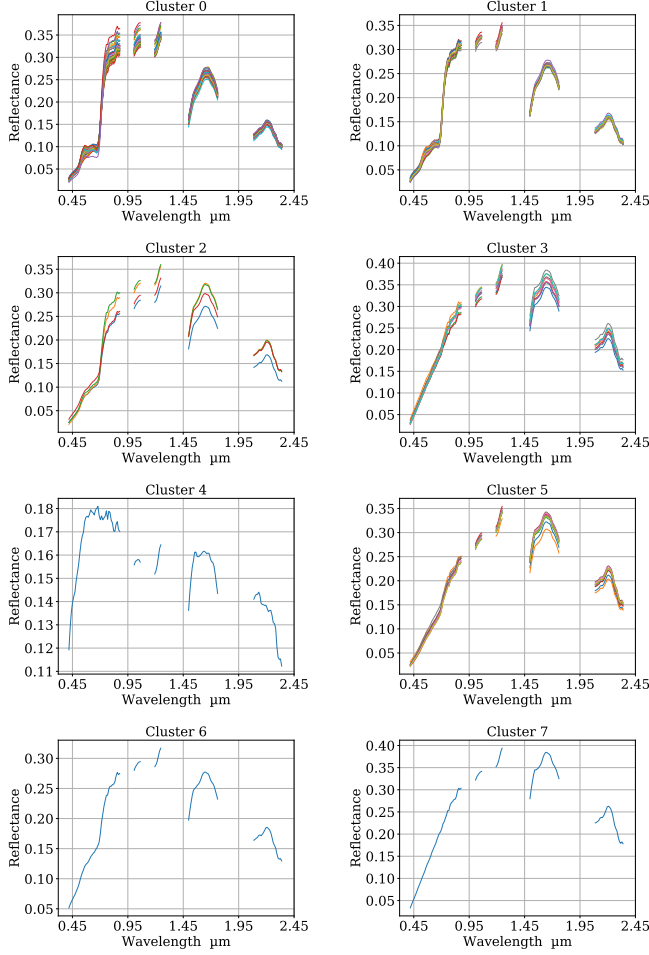
The HBEE stage has been performed with a set of parameters detailed hereafter. The homogeneity threshold  $t_h$  has been set to 0.045 by computing the average max-min distance of visually homogeneous areas. As it is commonly stated that two spectra represent the same material for an SA lower than 0.1 rad, the SA threshold  $t_{sa}$  has been set to 0.05 rad ( $2.86^\circ$ ) for strictness purpose. The pure 8 m x 8 m PAN areas then extracted are shown in Fig. 2. Large homogeneous area agglomerates exist as well as single isolated homogeneous areas.



**Fig. 2.** Locations of the homogenous 8 m panchromatic areas (in white).

As seen in Fig. 3, some clusters provided by this stage are quite similar. This is due to the low spectral angle threshold value  $t_{sa}$ . This allows us to address the spectral variability by considering several endmembers for a unique material. This result shows that very few materials are represented by pure pixels in the image and therefore, the reconstruction error map in Fig. 4 obtained after the method detailed in Section 3.1 has high values where unknown materials are indeed present. In our case, the materials within the most poorly reconstructed areas are the metal sheet roof tile of a large building and a large tree with some sparse shadow. In this paper, we focus on these major badly reconstructed areas by thresholding the reconstruction error image with a threshold set to  $t_{RE} = 0.12$ .

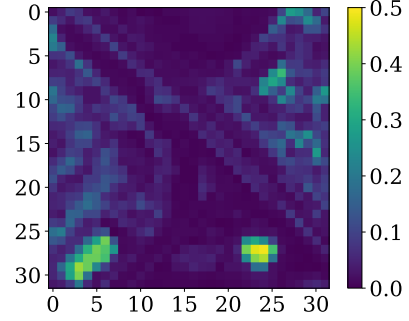
<sup>1</sup><http://www.specim.fi/products/aisafenix-hyperspectral-sensor/>



**Fig. 3.** The eight clusters of spectra, resulting from the clustering step.

## 5.2. LS-NMF stage results

In this section, we present a first test on synthetic data and then propose a proof of concept on real data. The synthetic data are generated as follows: three real spectra have been selected to be the known endmembers  $S_i$ . To fulfill the assumption of an unknown local predominant endmember, the abundance fractions  $X_{ij}$  are randomly set and the coefficients corresponding to the known endmembers are divided by three. The abundance fraction of the endmember to be estimated is then calculated to conform to the sum-to-one constraint. The observed  $N$  synthetic spectra are then generated. The unknown endmember is a metal sheet spectrum. The initialisation matrices are set as mentioned in Section 3 but the abundance fractions of the known material are randomly set and divided by  $a$  for this experiment on synthetic data. The results for this experiment are shown in Table 1 for ten iterations of the NMF with  $a$ , a coefficient dividing the initial random abundance fractions of the known endmembers. As can be seen, after ten iterations and for any value of  $a$ , the NMF out-



**Fig. 4.** Reconstruction error map (NRMSE) after the HBEE stage.

N	$a$							
	1	2	3	4	5	6	7	10
5	1.90	0.69	1.68	2.23	2.57	2.80	2.98	3.33
10	3.19	0.38	0.77	1.38	1.76	2.04	2.23	2.61
18	3.68	0.97	0.20	0.87	1.30	1.60	1.83	2.26

**Table 1.** Spectral angle value (in degrees) between the estimated endmembers and the true endmember after 10 NMF iterations on the synthetic data. The spectral angles of the initial guess (iteration 0) are  $5.08^\circ$ ,  $4.26^\circ$ ,  $4.08^\circ$  respectively for  $N=5, 10, 18$ .

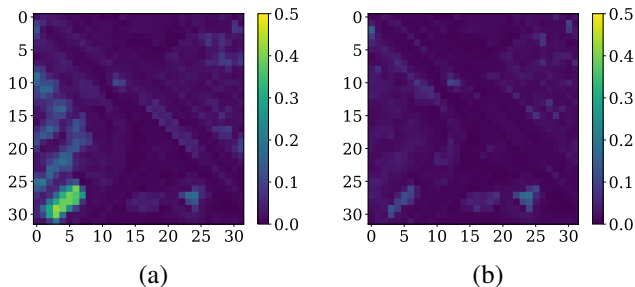
put has estimated a better endmember than the initial guess (the mean of the observed data) in a spectral angle sense. The optimal value of  $a$  (2 or 3) is coherent with the true abundance values of the processed data.

For real data, tests using the LS-NMF algorithm presented in Section 3.2 were done after the HBEE stage shown in Section 5.1. The dividing coefficient value  $a$  was set arbitrarily to three. This stage treated two poorly reconstructed areas. The spectral angles between the estimated spectrum and the corresponding 55 cm GT are shown in Table 2.

Material	SA(init, GT)	SA(final, GT)
Metal sheet	6.64	5.65
Vegetation	3.53	3.34

**Table 2.** Spectral angle between the estimated spectrum and the groundtruth before the first NMF iteration and after the end of the NMF on the HYPXIM-P image.

These preliminary results show a significant enhancement of the reconstruction error, especially in the treated areas but also where the estimated endmembers are in the scene (*e.g.* the grass field located on the image top right). Although some of the GT spectra have not been estimated, *e.g.* tiles and awning spectra, this LS-NMF stage estimates satisfactory endmembers that are not represented by pure pixels in the image by using prior knowledge of the panchromatic image. Further investigations on the combined reconstruction error map and panchromatic image will be conducted in order



**Fig. 5.** Evolution of the reconstruction error map after processing the first area containing the vegetation (a) and, in addition, the second area containing a metal sheet (b).

to estimate the remaining endmembers. As compared to the standard methods in Table 3, the proposed approach shows similar or better performance by estimating two endmembers in addition to those extracted from pure pixels, with a total of 10 estimated endmembers. The other methods have been applied, assuming the same number of endmembers to be estimated.

Criterion	HBee-LSNMF	VCA	NFINDR	ATGP	SISAL	MVCNMF
$RE(\%)$	1.72	1.69	1.88	2.22	2.11	1.41
$SAM(^{\circ})$	5.44	6.60	5.19	5.75	8.85	5.9

**Table 3.** Compared performances for 10 estimated endmembers.

## 6. CONCLUSION

A new blind unmixing method using both panchromatic and hyperspectral images was developed to extract the endmembers. A first stage using a homogeneity criterion followed by a spectral clustering is used to build a first set of endmembers that are present in pure pixels. The remaining endmembers are then estimated locally using both previously extracted spectra and the reconstruction error map. The results are promising since the method yields results similar to or better than those of the standard methods. Current works aim at improving the different stages, especially the LS-NMF stage by taking in account the panchromatic image.

## 7. REFERENCES

- [1] M. Eismann, *Hyperspectral Remote Sensing*, SPIE, 2012.
- [2] K. Staenz, A. Mueller, and U. Heiden, "Overview of terrestrial imaging spectroscopy missions," in *2013 IEEE International Geoscience and Remote Sensing Symposium - IGARSS*, July 2013, pp. 3502–3505.
- [3] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 44–57, Jan 2002.
- [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, pp. 354–379, 2012.
- [5] V. Paul Pauca, J. Piper, and R. J. Plemmons, "Non-negative matrix factorization for spectral data analysis," *Linear Algebra and its Applications*, vol. 416, no. 1, pp. 29 – 47, 2006, Special Issue devoted to the Haifa 2005 conference on matrix theory.
- [6] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.
- [7] C. Lin, "Projected gradient methods for nonnegative matrix factorization," *Neural Computation*, vol. 19, no. 10, pp. 2756–2779, 2007.
- [8] X. Liu, W. Xia, B. Wang, and L. Zhang, "An approach based on constrained nonnegative matrix factorization to unmix hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 2, pp. 757–772, Feb 2011.
- [9] Z. Yang, G. Zhou, S. Xie, S. Ding, J. Yang, and J. Zhang, "Blind spectral unmixing based on sparse nonnegative matrix factorization," *IEEE Transactions on Image Processing*, vol. 20, no. 4, pp. 1112–1125, April 2011.
- [10] V. Carrere, X. Briottet, S. Jacquemoud, R. Marion, A. Bourguignon, M. Chami, M. Dumont, A. Minghelli-Roman, C. Weber, M. Lefevre-Fonollosa, and M. Mandea, "Hypxim: A second generation high spatial resolution hyperspectral satellite for dual applications," in *2013 5th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, June 2013, pp. 1–4.
- [11] 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.
- [12] R. Bro and S. D. Jong, "A fast nonnegativityconstrained least squares algorithm," *Journal of Chemometrics*, vol. 11, no. 5, pp. 393–401, 10 1997.
- [13] C. Miesch, L. Poutier, V. Achard, X. Briottet, X. Lenot, and Y. Boucher, "Direct and inverse radiative transfer solutions for visible and near-infrared hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 7, pp. 1552–1562, July 2005.