Filtering-based endmember identification method for snapshot spectral images

Kinan Abbas, Matthieu Puigt, Gilles Delmaire, Gilles Roussel

FILTERING-BASED ENDMEMBER IDENTIFICATION METHOD FOR SNAPSHOT SPECTRAL IMAGES

Kinan Abbas, Matthieu Puigt, Gilles Delmaire, and Gilles Roussel

Univ. Littoral Côte d'Opale, LISIC – UR 4491, F-62219 Longuenesse, France

ABSTRACT

In this paper, we propose a new endmember estimation method for snapshot spectral imaging (SSI) systems using Fabry-Perot filters. Indeed, such systems only provide a part of the spectral content of a classical multi- or hyperspectral camera and restoring the full hyperspectral datacube from an SSI matrix is named "demosaicing". However, it was recently shown that a joint unmixing and demosaicing method allowed a much better unmixing performance than a two-stage approach consisting of a demosaicing step followed by an unmixing one. In this paper, we propose a new approach to estimate endmembers from the SSI image without requiring a demosaicing step. It inverts the Fabry-Perot filters and extends the "pure pixel" framework to the SSI sensor patch level, that we name the "pure patch" assumption. Our experiments show that our proposed scheme significantly outperforms state-of-the-art methods in terms of endmember estimation accuracy.

***Index Terms*—** Snapshot Spectral Imaging, Endmember Identification, Fabry-Perot Filter Inversion, Rank-one Approximation, Pure Patch Assumption

1. INTRODUCTION

An essential issue that hyperspectral imaging (HSI) sensors have to handle is how to collect the four-dimensional HSI data—i.e., two spatial, one spectral, and one temporal dimensions—utilizing a single detector, i.e., a 1D-array or 2D-plane detectors. Thus, different strategies in HSI acquisition designs have emerged from the disparity between detector demands and available dimensionality, leading to spatial, spectral, and frame scanning architectures [1]. Regardless of the approach, the main potential of all the methods is acquiring images with a high spectral, spatial, and temporal resolution. However, a shared characteristic among all scenarios is the need for repetitive scanning of the scene and the acquisition of many exposures (frames) to capture the complete spatio-spectral resolution data cube. Moreover, the need for miniaturization of the imaging systems implies that novel designs should seek to be without mechanical parts—such as moving mirrors—since they increase the complexity of the designs. Therefore, a new generation of HSI imaging architectures—known under the name Snapshot Spectral Imaging (SSI)—has been proposed to overcome the previously mentioned issues [2]. Among the recent strategies which emerged, compressive coded-aperture SSI (CASSI) systems [3] and SSI cameras using Fabry-Perot filters (FPfs) [4] only provide a 2-D image which is obtained from the 3-D HS data. With both cameras, a post-processing technique known as "demosaicing" must be applied to estimate the full HS data cube. However, the proposed methods significantly differ with respect to the chosen sensing technology. In this paper, we focus on the cameras using FPfs.

Traditional demosaicing methods use spatial and/or spectral correlation and are based on Weighted Bilinear interpolation (WB) [5], Binary Tree-Based Generic Demosaicing (BTES) [6], Iterative Spectral Difference (ItSD) [7], a pseudo-panchromatic image (PPID) [8], structural and adaptive nonlocal optimization (SaND) [9], and graph-regularized low-rank matrix completion (GRMR) [10]. More recently, deep learning methods have been proposed [11–14].

Once the 3-D HS image has been restored from the 2-D SSI one, we can apply any postprocessing technique of interest. In particular, extracting the spectral signatures of all the materials—aka endmembers—which are present in an observed scene is a very classical problem known as *unmixing*. Specifically, popular unmixing methods assume that for each endmember, there is at least one spatial pixel which contains the correspondent material only, such that the observed spectrum is equal to the endmember. The most widely used algorithms are the vertex component analysis (VCA) [15], N-FINDR [16], the pixel purity index (PPI), and the sequential maximum angle convex cone (SMACC) [1].

However, according to the authors in [10], performing classification after demosaicing SSI images provides a poor classification performance. Furthermore, in [17], we found that a two-stage strategy—which consists of first demosaicing an SSI image and then unmixing it—was much less accurate than jointly demosaicing and unmixing the same SSI data. In particular, we showed that our proposed approach—which was solving a low-rank matrix completion problem using a locally-rank-1 weighted nonnegative matrix factor-

This work was partly funded by the Région Hauts-de-France. Experiments presented in this paper were carried out using the CALCULCO computing platform, supported by SCoSI/ULCO.

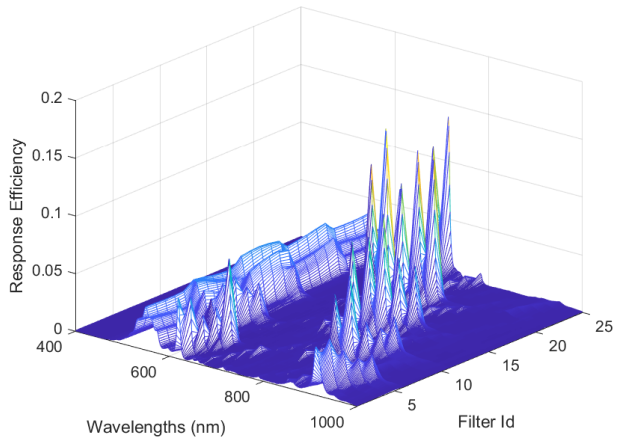

Fig. 1. Spectral response of the 25 spectral filters of the 5×5 mosaic Photon Focus SSI camera in the 400–1000 nm range.

ization framework and which is named K-means patch-based weighted non-negative matrix factorization (KPWNMF)—was providing much better estimates of the endmembers than the two-stage approaches while still providing a slightly better demosaicing performance.

In this work, we propose a novel approach to estimate endmembers from the SSI image. It significantly differs from [17] for the following reasons. Firstly, instead of a low-rank matrix completion framework, we here propose to invert the Fabry-Perot filters to recover the spectral content from "patches" of the SSI image. Then, we relax the abundance sparsity assumption needed in [17]: we here assume that the previously recovered spectra may be seen as linear mixtures of the endmembers and that some of the patches only contained one material, i.e., we extend the "pure pixel" idea to the patch level. The VCA algorithm is then used to derive the final endmembers, while the abundances can then be estimated using the method in [17][1].

The remainder of the paper is organized as follows. We introduce the SSI camera and the problem we aim to solve in Section 2. Then, section 3 presents our proposed method, whose performance is investigated in Section 4. We lastly conclude and introduce future work direction in Section 5.

2. PROBLEM STATEMENT

This section defines the SSI acquisition system and the problem we aim to solve. The SSI camera acquires a two-dimensional image of $m \times n$ pixels for each exposure, where m and n are the numbers of horizontal and vertical pixels, respectively. Moreover, the camera is assumed to observe k spectral bands. In practice, the considered SSI technology is based on a mosaic of FPfs which consist of $\sqrt{k} \times \sqrt{k}$ patches which are repeated along the sensor surface[2]. In the ideal scenario, an FPf allows the light from a minimal spectral range to propagate to a sensor and stops the light outside this range. However, these filters present additional harmonics around each wavelength of interest in the real implementation, as shown in Fig. 1. Fortunately, these filters are known and provided by the camera manufacturer [4].

In the remainder of this section, we focus on a single patch of FPfs. Denoting $y_i(\lambda_i)$ the i-th SSI pixel of the considered patch—assumed to theoretically observe the spectral information at λ_i nm—we get:

$$y_i(\lambda_i) = \sum_{j=1}^{k} h_i(\lambda_j) \cdot x_i(\lambda_j) + \omega_i, \tag{1}$$

where $h_i(\lambda)$ is the FPf associated with Pixel i, $x_i(\lambda)$ is the plain spectrum to be observed on Pixel i, and ω_i is some additive noise. Moreover, assuming a linear mixture model, the observed spectrum can be written as a mixture of endmembers, i.e.,

$$x_i(\lambda) = \sum_{l=1}^{p} g_{il} f_l(\lambda), \tag{2}$$

where p is the number of endmembers present in the observed scene, $f_l(\lambda)$ denotes the l-th endmember, and g_{il} its associated abundance proportion in Pixel i with

$$\forall l = 1, \ldots, p,\ 0 \leq g_{il} \leq 1 \quad \text{and} \quad \sum_{l=1}^{p} g_{il} = 1. \tag{3}$$

In this paper, we aim to estimate the p endmembers $f_l(\lambda)$ from the SSI image using Eqs. (1) and (2).

3. PROPOSED METHOD

We now introduce our proposed method. First of all, let us emphasize the fact that inverting the k FPfs $h_i(\lambda)$ in Eq. (1) is ill-posed. In order to estimate the endmembers, we need to add additional assumptions. We first assume that in some patches, the matrix X which is defined over the considered patch as

$$X \triangleq \begin{bmatrix} x_1(\lambda_1) & \ldots & x_1(\lambda_k) \\ \vdots & & \vdots \\ x_k(\lambda_1) & \ldots & x_k(\lambda_k) \end{bmatrix} \tag{4}$$

is approximately rank-1, i.e., $\forall j = 1, \ldots, k$, for any given indices i_0 and i_1, $x_{i_0}(\lambda_j) \approx x_{i_1}(\lambda_j)$. Such an assumption is valid for miniaturized SSI systems which can be quite close to the area they tend to observe. For the sake of readability, we drop the indices and reduce the matrix X to a single vector $\underline{x} = [x(\lambda_1), \ldots, x(\lambda_k)]$. Denoting ^{T} the transposition,

$$\underline{y} \triangleq [y_1(\lambda_1), \ldots, y_k(\lambda_k)]^T, \tag{5}$$

[1] This is the reason why we only focus on the endmember estimation in this paper.

[2] Typical values of k are 16 or 25, such that the patch is of size 4×4 or 5×5, respectively. Moreover, m and n are both proportional to $\sqrt{k}$.

the vector of the k observed pixels in a patch, $\underline{\omega} \triangleq [\omega_1, \ldots, \omega_k]^T$ the associated vector of additive noise, and

$$H \triangleq \begin{bmatrix} h_1(\lambda_1) & \ldots & h_1(\lambda_k) \\ \vdots & & \vdots \\ h_k(\lambda_1) & \ldots & h_k(\lambda_k) \end{bmatrix} \quad (6)$$

the matrix of FPfs, we get

$$\underline{y} = H \cdot \underline{x}^T + \underline{\omega}. \quad (7)$$

Our approach thus reads as follows. For each patch, we first aim to recover a tentative spectrum $\underline{x}$ from Eq. (7). In practice, as the matrix H can be ill-conditionned[3], we aim to solve a penalized optimization problem, i.e.,

$$\min_{\underline{x} \geq 0} \frac{1}{2} \|\underline{y} - H \cdot \underline{x}^T\|_2^2 + \frac{\alpha}{2} \|D \cdot \underline{x}^T\|_2^2, \quad (8)$$

where D is the square matrix accounting for the discrete derivative of the spectrum $\underline{x}$, and α stands for the penalization term. Eq (8) is a quadratic problem which can be rewritten as [18]

$$\min_{\underline{x} \geq 0} \frac{1}{2} \left\| \begin{pmatrix} \underline{y} \\ 0 \end{pmatrix} - \begin{pmatrix} H \\ \sqrt{\alpha}\, D \end{pmatrix} \cdot \underline{x}^T \right\|_2^2. \quad (9)$$

In practice, the error value $\|\underline{y} - H \cdot \underline{\hat{x}}\|_2$—where $\underline{\hat{x}}$ is the estimated spectrum obtained from Eqs. (8) or (9)—provides a measure of rank-one approximation. Indeed, if in Eq. (7) the content of X cannot be approximated by a single vector, then $\|\underline{y} - H \cdot \underline{\hat{x}}\|_2$ will be high. On the contrary, it will be low if the rank-one approximation is a valid assumption. By only keeping the estimated vectors linked with the lowest errors—say q vectors—we get a $q \times k$ data matrix $\mathbb{X}$ where the FPf effects are removed. Let us recall that these q vectors are associated with q patches which are approximately rank-1. In practice, such a situation may be met if (i) only one endmember is active over the patch, or (ii) the abundance proportions are constant over the patch. The first scenario may be seen as a patch extension of the "pure pixel assumption" [1]—that we call "pure patch assumption"—while the second means that several endmembers are present in similar proportions over the patch.

Denoting G and F the $q \times k$ and $k \times k$ matrices of abundances and endmembers associated with the q rank-1 patches, we get a simple linear mixing relationship

$$\mathbb{X} \approx G \cdot F. \quad (10)$$

In this paper, we further assume that at least one pure patch exists for each endmember to estimate. This means that Eq. (10) can be solved by any unmixing method using the pure patch assumption. In order to show the proof of concept, we use VCA [15]. Algorithm 1 shows the whole structure of the approach, named FPVCA for Filter Patch-based VCA.

At this stage, let us stress again that due to the fact that a miniaturized SSI camera may be placed very close to an area to observe, the rank-1 and the pure patch assumptions are not very constraining. Moreover, we would like to emphasize the differences between this method and the one we proposed in [17]: our previous work was using a weighted matrix factorization framework, which significantly differs from the filter inversion strategy used in this paper. Moreover, contrary to [17], we do not assume all the rank-one patches to be pure.

Algorithm 1 Filter Patch-based Vertex Component Analysis

Input:
the SSI matrix
p the number of endmembers
H the harmonics matrix
α the regularization parameter
Output:
F the estimated endmember matrix
Processing:

```
for r = 1 to nb_patches do
    Let y_r the SSI vector linked to Patch r
    Estimate x using Eq (9)
    if ||y_r − H x||_2 is "low-enough" then
        Add x as a new line of X
    end if
end for
F = VCA(X, p)
```

4. EXPERIMENTS AND RESULTS

Experiments are conducted using SSI simulations on the spheric image, which is generated using the HYDRA toolbox [19]. It has 128×128 pixels with three endmembers, i.e., water, concrete, and metal. The image satisfies abundance sum-to-one and abundance non-negativity constraints. Moreover, for each endmember, we generate one spatial area where only this one is present, so that the pure patch assumption is satisfied. We consider 4×4 and 5×5 spectral filter patterns for ideal filters—where Eq. (1) reduces to $y_i(\lambda_i) = x(\lambda_i)$ [4]—and real FPfs—whose responses are provided by IMEC and used in [10]—and we randomly select either $k = 16$ or 25 spectral bands out of the available ones. Lastly, we only keep for generating $\mathbb{X}$ the vectors $\underline{x}$ such that the rank-one approximation error $\|\underline{y} - H\underline{x}\|_2$ is below or equal to the average of these errors computed over all the patches.

We compare the performance achieved by our proposed FPVCA method with our previously proposed KPWNMF approach [17] and five 2-step demosaicing-then-unmixing

[3] For example, the 5×5 filter matrix H using real FPfs in [10] has two rows which are almost null.

[4] In that case, it is not possible to solve Eq. (8). We just generate $\underline{x}$ by replicating the SSI values and we set $\|\underline{y} - H\underline{x}\|_2 = 0$.

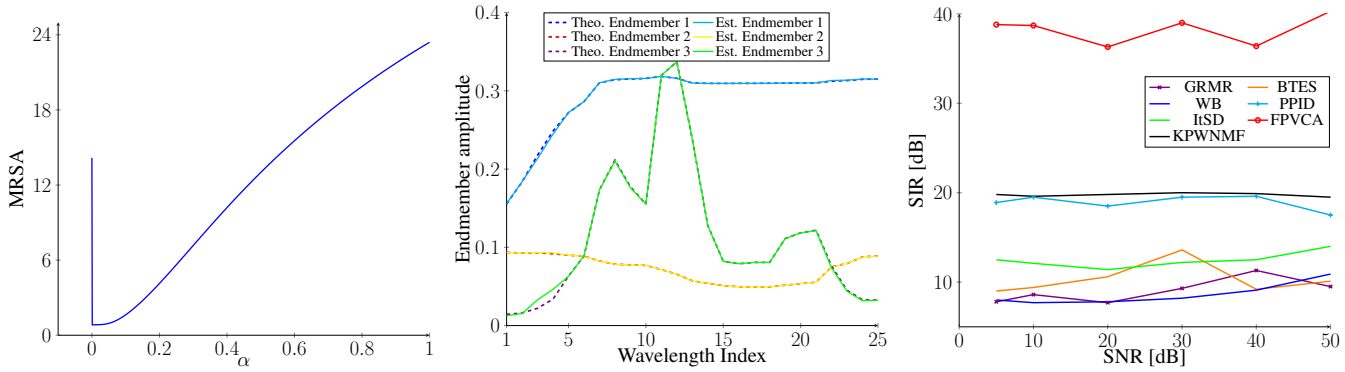

Fig. 2. Left: Influence of the regularization parameter α on the reached MRSA value. Center: Estimated spectra in case of real filter of size 5×5. Right: SIRs obtained with the 5×5 real filter with respect to additive noise.

methods. For the latter, we consider five state-of-the-art demosaicing methods—i.e., GRMR [10], BTES [6], WB [5], PPID [8], and ItSD [7]—while in the second step, we extract the endmembers in the restored datacube using the VCA algorithm. As a consequence, all the considered methods should take advantage of the presence of pure pixels.

To assess the performance of the tested methods, we compare the quality of the restored spectra using the Signal-to-Interference Ratio (SIR, reported in dB) and Mean-Removed Spectral Angle (MRSA) computed over the rows of F. The MRSA measures how close two endmembers are (neglecting scaling and translation) [20]. We first investigate the influence of α on the estimation of the endmembers. The left plot of Fig. 2 shows the MRSA with respect to the value of α in the case of a 5×5 filter. We see that a small value of α (typically, $0.001 \leq \alpha \leq 0.02$) allows a much lower MRSA than no regularization ($\alpha = 0$). Then, when α gets higher, the error is increased. In the remainder of the tests, we set this parameter to $\alpha = 0.005$.

Table 1 provides the average SIR and MRSA of the estimated spectra achieved by the various methods in noiseless mixtures. This table shows that the unmixing performance obtained with our proposed FPVCA method is always much better than the other methods. This is probably because all the tested 2-stages approaches introduce some datacube estimation errors that affect the quality of the estimated spectra. Moreover, the performance of our previously proposed KPWNMF approach is slightly better than the SotA other methods but still much lower than FPVCA. This may be due to the fact that the abundance sparsity assumption in [17]—i.e., all the rank-one patches are pure patches—is not necessarily satisfied in these simulations. Still, KPWNMF assumes all the vectors in $\mathbb{X}$ to be close to the true endmembers and estimates the latter using K-means. Such a clustering algorithm may be affected by the spectra obtained from non-pure patches, which lowers the KPWNMF performance.

Please also note the drop of SIR in the case of 5×5 real filters. While our proposed FPVCA approach still outperforms the other methods, its performance decrease is due to the fact that H is severely ill-conditioned. The proposed Tikhonov regularization serves to correct the effect of the matrix H. Furthermore, the angle between the estimated spectra and the real one is small, and they tend to be almost identical, as shown in the central plot of Fig. 2.

Table 1. Average SIR and MRSA (into brackets) values for the considered experiment. In bold: the best values

	4 x 4 patch		5 x 5 patch	
	Ideal Filter	**Real Filter**	**Ideal Filter**	**Real Filter**
GMRM	61.6 (4.0)	56.7 (4.8)	10.2 (4.0)	11.5 (3.6)
BTES	58.2 (4.4)	55.5 (4.9)	9.5 (4.4)	9.7 (4.5)
WB	61.4 (3.8)	57.7 (4.9)	11.7 (3.8)	12.2 (3.9)
PPID	73.1 (2.3)	63 (3.7)	28 (2.3)	16.3 (4.2)
ItSD	61.2 (4.8)	58.6 (4.6)	14.6 (4.8)	14.8 (4.8)
KPWNMF	77.8 (2.2)	64.3 (4.8)	31.8 (2.2)	19.2 (4.0)
FPVCA	**253.0 (0)**	**253.0 (0.0)**	**246.0 (0.0)**	**67.4 (0.8)**

Lastly, the influence of additive noise is shown on the right plot of Fig. 2, for the 5×5 real filter. All the tested methods are rather stable to the additive noise but the reached performance is lower than in the noiseless case. Still, our proposed FPVCA approach significantly outperforms all the tested methods, which shows the relevance of our work.

5. CONCLUSION AND DISCUSSION

In this paper, we proposed a new method for estimating endmembers from SSI images. It works on an SSI patch level and it mainly consists of finding rank-one patches and of restoring the spectral content from the latter. Then, any unmixing method assuming pure pixels—e.g., VCA—can be applied to the collected data. The proposed method significantly outperforms all the tested 2-stage approaches—which consist of applying any demosaicing method to restore the original datacube, and then of applying the same unmixing method as our proposed approach—but also a specific joint unmixing and

demosaicing method that we previously proposed. In future work, we aim to investigate the use of our proposed method on real SSI data. We will also investigate its demosaicing performance. We further aim to extend it to the case when the FPfs are not constant over the sensor surface [18] and when some endmember spectral variability is met in the acquisition process.

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 J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, pp. 354–379, Apr. 2012.

[2] N. A. Hagen and M. W. Kudenov, "Review of snapshot spectral imaging technologies," *Optical Engineering*, vol. 52, no. 9, p. 090901, 2013.

[3] G. R. Arce, D. J. Brady, L. Carin, H. Arguello, and D. S. Kittle, "Compressive coded aperture spectral imaging: An introduction," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 105–115, 2013.

[4] B. Geelen, C. Blanch, P. Gonzalez, N. Tack, and A. Lambrechts, "A tiny VIS-NIR snapshot multispectral camera," in *Advanced Fabrication Technologies for Micro/Nano Optics and Photonics VIII* (G. von Freymann, W. V. Schoenfeld, R. C. Rumpf, and H. Helvajian, eds.), SPIE, Mar. 2015.

[5] J. Brauers and T. Aach, "A color filter array based multispectral camera," in *12. Workshop Farbbildverarbeitung* (G. C. Group, ed.), (Ilmenau), October 5-6 2006.

[6] L. Miao, H. Qi, R. Ramanath, and W. Snyder, "Binary tree-based generic demosaicking algorithm for multispectral filter arrays," *IEEE Trans. Image Process.*, vol. 15, pp. 3550–3558, Nov. 2006.

[7] J. Mizutani, S. S. Ogawa, K. Shinoda, M. Hasegawa, and S. Kato, "Multispectral demosaicking algorithm based on inter-channel correlation," in *Proc. IEEE VCIP'14*, pp. 474–477, 2014.

[8] S. Mihoubi, O. Losson, B. Mathon, and L. Macaire, "Multispectral demosaicing using pseudo-panchromatic image," *IEEE Trans. Comput. Imaging*, vol. 3, pp. 982–995, Dec. 2017.

[9] L. Bian, Y. Wang, and J. Zhang, "Generalized MSFA engineering with structural and adaptive nonlocal demosaicing," *IEEE Trans. Image Process.*, vol. 30, pp. 7867–7877, 2021.

[10] G. Tsagkatakis, M. Bloemen, B. Geelen, M. Jayapala, and P. Tsakalides, "Graph and rank regularized matrix recovery for snapshot spectral image demosaicing," *IEEE Trans. Comput. Imaging*, vol. 5, pp. 301–316, June 2019.

[11] K. Dijkstra, J. van de Loosdrecht, L. R. B. Schomaker, and M. A. Wiering, "Hyperspectral demosaicking and crosstalk correction using deep learning," *Machine Vision and Applications*, vol. 30, pp. 1–21, July 2018.

[12] T. A. Habtegebrial, G. Reis, and D. Stricker, "Deep convolutional networks for snapshot hypercpectral demosaicking," in *Proc. IEEE WHISPERS'19*, Sept. 2019.

[13] Z. Pan, B. Li, H. Cheng, and Y. Bao, "Joint demosaicking and denoising for CFA and MSFA images using a mosaic-adaptive dense residual network," in *Computer Vision – ECCV 2020 Workshops*, pp. 647–664, Springer International Publishing, 2020.

[14] K. Feng, Y. Zhao, J. C.-W. Chan, S. Kong, X. Zhang, and B. Wang, "Mosaic convolution-attention network for demosaicing multispectral filter array images," *IEEE Trans. Comput. Imaging*, vol. 7, pp. 864–878, 2021.

[15] J. M. P. Nascimento and J. M. Bioucas-Dias, "Vertex component analysis: a fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, 2005.

[16] M. E. Winter, "N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data," in *Imaging Spectrometry V*, vol. 3753, pp. 266–275, International Society for Optics and Photonics, 1999.

[17] K. Abbas, M. Puigt, G. Delmaire, and G. Roussel, "Méthode de démélange et dématriçage conjoints fondée sur la complétion de rang un pour les images multispectrales "snapshot"," in *Proc. of GRETSI*, 2022.

[18] P. Chatelain, G. Delmaire, M. Puigt, and G. Roussel, "Inversion de réseaux de filtres de Fabry-Perot pour la restauration de cubes hyperspectraux," in *Proc. of GRETSI*, Sept. 2022.

[19] Grupo de Inteligencia Computacional, UPV/EHU, "Hyperspectral imagery synthesis (EIAs) toolbox." `http://www.ehu.es/ccwintco/index.php/Hyperspectral_Imagery_Synthesis_tools_for_MATLAB`, Last accessed: 2022-02-25.

[20] N. Gillis and S. A. Vavasis, "Semidefinite programming based preconditioning for more robust near-separable nonnegative matrix factorization," *SIAM Journal on Optimization*, vol. 25, pp. 677–698, jan 2015.