Unmixing of Hyperspectral Images with Pure Prior Spectral Pixels

Abir Zidi[1], Julien Marot[2], Klaus Spinnler[1] and Salah Bourennane[2]
[1]*Fraunhofer IIS, Erlangen, Germany*
[2]*Institut Fresnel, Marseille, France*

Keywords: Non-negative Matrix Factorization, PARAFAC, Linear Algebra, Hyperspectral Image, Remote Sensing, Tensor.

Abstract: In the literature, there are several methods for multilinear source separation. We find the most popular ones such as nonnegative matrix factorization (NMF), canonical polyadic decomposition (PARAFAC). In this paper, we solved the problem of the hyperspectral imaging with NMF algorithm. We based on the physical property to improve and to relate the output endmembers spectra to the physical properties of the input data. To achieve this,we added a regularization which enforces the closeness of the output endmembers to automatically selected reference spectra. Afterwards we accounted for these reference spectra and their locations in the initialization matrices. To illustrate our methods, we used self-acquired hyperspectral images (HSIs). The first scene is compound of leaves at the macroscopic level. In a controlled environment, we extract the spectra of three pigments. The second scene is acquired from an airplane: We distinguish between vegetation, water, and soil.

1 INTRODUCTION

Unmixing is an important preprocessing step to further analyze a hyperspectral image (HSI). Tensor decomposition method has been adapted to unmix HSIs, for instance PARAFAC model. However, this model has problems: Firstly, PARAFAC algorithms are very slow in certain circumstances. The alternating least squares algorithm (ALS), used in the PARAFAC-decomposition, may require a very large number of iterations before converging. This slowness of convergence can result from large data, improper packaging, or the presence of degeneration. Secondly PARAFAC algorithms are very complicated to use. Indeed, there are several parameters to find. In this paper, a new algorithms developed with the purpose of circumventing these problems. It consists in decomposing a set of spectra contained in a matrix $\mathbf{Y}$ as the following product: $\mathbf{Y} = \mathbf{AX} + \mathbf{N}$, where $\mathbf{X}$ is the endmember matrix containing the 'source' spectra, $\mathbf{A}$ is the mixing matrix containing the mixing coefficients or contribution of each 'source' or endmember, and $\mathbf{N}$ stands for the modeling errors. Finding out an estimate $\hat{\mathbf{A}}$ of the mixing matrix, with the positivity and possibly sum-to-one contraints, and an estimate $\hat{\mathbf{X}}$ of the endmember matrix with the positivity constraint is a low-rank approximation problem commonly called Nonnegative Matrix Factorization (NMF) (H. Kim et al.,2008).

NMF has been applied to HSI data characterization, and exhibited some advantages with respect to existing endm ember extraction methods such as vertex component analysis (VCA) (JMP. Nascimento et al., April 2005). The interest of VCA was illustrated in the frame of plant analysis in (J. Marot et al., June 2013). It turned out that, contrary to NMF, VCA does not ensure the positivity of the mixing coefficients. NMF has been extensively improved for the past few years (A. Cichocki,et al.,2009). In (Ma et al., January 2004), spectral unmixing is related to optimization and signal processing (in particular: convex geometry concepts). But despite its advantages, it still exhibits some drawbacks: Firstly, it is sensitive to initialize (Q. Du, et al., July 2005). Secondly, there exist an infinite number of solutions for $\mathbf{A}$ and $\mathbf{X}$, and not all of them bear a physical significance. Hence, our problematic is two-fold: How to find an appropriate initialization for the endmember matrix and for the mixing matrix? How can we approach a solution with a physical significance?

Firstly, we propose to use the purest spectra in the scene, selected by a geometrical criterion (JMP. Nascimento et al., April 2005), as a subset of initialization endmembers. As sources are often mostly grouped in separate regions (H. Kim , et al.,2008), we propose a sparse and positive initialization of the mixing matrix, which accounts for the location of ini-

Zidi A., Marot J., Spinnler K. and Bourennane S..
Unmixing of Hyperspectral Images with Pure Prior Spectral Pixels.
DOI: 10.5220/0005311101530158
In *Proceedings of the 10th International Conference on Computer Vision Theory and Applications* (VISAPP-2015), pages 153-158
ISBN: 978-989-758-089-5

tialization spectra.

Secondly, we propose to use the pure spectra as references in the criterion which is minimized to perform NMF. We expect that, as they are selected among the spectral pixels of the HSI, aiming at the closeness to these pure spectra will encourage the physical significance of the endmembers provided by NMF. That is why, for the first time to the best of our knowledge, we introduce these spectra in the criterion which is minimized for the purpose of factorization.

In Section 2, we present the notations which hold throughout the paper, and we show how a HSI is handled to get a set of one-dimensional spectra. The linear mixing model above is then detailed. In Section 3, we propose an innovative initialization for the mixing and endmember matrices, and a new criterion to perform NMF while ensuring the physical significance of the estimated endmember spectra. In Section 4, we detail our implementation of NMF. Section 5 presents the results obtained: Firstly, we extract pigment spectra from a leaf reflectance; secondly, we distinguish between vegetated areas, soil and water in an aerial HSI.

2 NOTATIONS AND DATA MODEL

In the rest of the paper, x denotes a scalar, $\mathbf{x}$ denotes a 1-dimensional vector, $\mathbf{X}$ denotes a 2-dimensional matrix, $\mathcal{X}$ denotes a multidimensional array, also called "tensor" (D. Muti, et al., September 2008). For any vector $\mathbf{x}$, $\mathbf{x}^T$ stands for transpose.

To set the link between algebraic methods and HSIs, a HSI is considered from a mathematical point of view as a tensor of order 3 $\mathcal{T} \in \mathbb{R}^{I_1 \times I_2 \times L}$, where I_1 is the number of rows, I_2 is the number of columns, and L is the number of channels. In the following, we select a subset of $S \leq I_1 I_2$ spectral pixels of $\mathcal{T}$ and set them row-wise in a matrix $\mathbf{Y}$ of size $S \times L$.

Let's consider one row of matrix $\mathbf{Y}$, a spectral pixel denoted by $\mathbf{y}_i$, which is a vector of size $1 \times L$. The model that we adopt for $\mathbf{y}_i$ is the linear combination of J endmembers denoted by $\mathbf{x}_j$ ($j = 1, \ldots, J$). Vector $\mathbf{y}_i$, $i = 1, \ldots, S$ is expressed as:

$$\mathbf{y}_i = \sum_{j=1}^{J} a_{ij}\mathbf{x}_j + \mathbf{n}_i \tag{1}$$

where $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_J$ are the endmember spectra, and $a_{i1}, a_{i2}, \ldots, a_{iJ}$ stand for the abundances of each endmember in the pixel vector $\mathbf{y}_i$.

The term $\mathbf{n}_i$ stands for an additive residual term accounting for the measurement noise and modeling error.

The endmember spectra are supposed to be positive-valued. The abundances a_{ij}, $j = 1, \ldots, J$ are such that:

$$0 \leq a_{ij} \leq 1, \ \forall\, i = 1, \ldots, S \tag{2}$$

$$\sum_{j=1}^{J} a_{ij} = 1 \ \forall\, i = 1, \ldots, S \tag{3}$$

Let $\mathbf{a}_i = [a_{i1}, a_{i2}, \ldots, a_{ij}, \ldots, a_{iJ}]$ be the row vector containing the abundance values associated with $\mathbf{y}_i$. We define the abundance, or mixing matrix as $\mathbf{A}$, whose rows are the abundance vectors $\mathbf{a}_i, i = 1, \ldots, S$ associated with the rows of matrix $\mathbf{Y}$. We define the endmember matrix as $\mathbf{X}$, whose rows are the J endmember spectra. With this formalism, and referring to Eq. (1), we retrieve the linear mixing model presented in the introduction. This data model is in agreement with the one in (A. Cichocki, et al.,2009).

3 NEW CRITERION AND INITIALIZATION MATRICES FOR NMF

The basic NMF optimized function ensures that the two matrix $\mathbf{A}$ and $\mathbf{X}$ are both nonnegatives. Since the NMF solution is not unique, some prior knowledge on HSIs can be introduced to solve this problem.

In this section, Accordance with valid knowledge of the data, we add constraints (itemized in section 4) to improve the result of deconvolution, using the pure spectrum provided by an innovative initialization.

3.1 Minimized Criterion

To get an estimate of the endmember matrix and the mixing matrix, we seek to minimize the criterion $D(\mathbf{Y}||\mathbf{A},\mathbf{X})$: $\hat{\mathbf{A}}, \hat{\mathbf{X}} = \underset{(\mathbf{A},\mathbf{X})}{argmin} D(\mathbf{Y}||\mathbf{A},\mathbf{X})$. In the simplest versions of the NMF, assuming that the modeling error $\mathbf{N}$ is independent identically distributed, the problem of estimating $\mathbf{A}$ and $\mathbf{X}$ is formulated as the maximization of a likelihood function (see (A. Cichocki, et al.,2009), chapter 3), or equivalently, the minimization of the criterion $D_F(\mathbf{Y}||\mathbf{A},\mathbf{X}) = ||\mathbf{Y} - \mathbf{AX}||_F^2$, where $||\cdot||_F$ denotes Frobenius norm. In real-world data, the actual mixing matrix is rather sparse, owing to the spatial repartition of the materials in the scene: They are often grouped in regions. As advised in (H. Kim et al.,2008), to induce sparsity in the mixing matrix, we add an l_1-norm regularization term. In addition to this term which is related to the spatial properties of the data, we also propose a regularization term which is related to the shape of the spectra:

It enforces the endmembers to approach a set of so-called 'pure' spectra selected from the spectral pixels of the HSI. This selection can be performed by vertex component analysis, pixel purity index, or N-Finder methods (JMP, et al.,April 2005). We choose vertex component analysis because of its low complexity (JMP, et al.,April 2005). Adding the spatial and the spectral regularization terms yields:

$$D(\mathbf{Y}||\mathbf{A},\mathbf{X}) = ||\mathbf{Y}-\mathbf{A}\mathbf{X}||_F^2 + \alpha||\mathbf{A}||_1 + \beta||\mathbf{X}-\mathbf{X}_{pure}||_F^2 \quad (4)$$

where matrix $\mathbf{X}_{pure}$'s rows are the J pure spectra, and α and β are regularization coefficients.

3.2 Initialization Mmatrices

3.2.1 Endmember Matrix

Let $J' \leq J$ be the number of spectra, among the rows of matrix $\mathbf{Y}$, which are assumed to result from a pure material. Typically, one of these spectra can be issued from the light reflected by a pure metal, or by a green section of a leaf supposed to contain only chlorophyll. We choose the first J' pure spectra provided by VCA in subsection 3.1. Let $\mathbf{X}_{1_{Init}}$ be the matrix whose rows contain these spectra. The initialization endmember matrix is defined as:

$$\mathbf{X}_{Init} = \begin{bmatrix} \mathbf{X}_{1_{Init}} \\ \mathbf{X}_{2_{Init}} \end{bmatrix} \quad (5)$$

where $\mathbf{X}_{2_{Init}}$ is a random matrix compound of $J-J'$ rows and L columns. Then, as explained in Section 4, we scale each row of $\mathbf{X}_{Init}$ to unit ℓ_2 norm because we choose a hierarchical alternating least squares (HALS) algorithm.

3.2.2 Abundance Matrix

Let $k_1,...,k_{J'}$ be the subset of row indices associated with the initialization pure spectra $\{\mathbf{y}_{k_1},...,\mathbf{y}_{k_{J'}}\}$ in $\mathbf{Y}$. The coefficients $\{a_{ij}\}$, for $i = 1,...,S$, are initialized as follows:

for $j \leq J'$, $a_{ij} = 1$ if $i = k_j$; and $a_{ij} = 0$ if $i \neq k_j$
for $j > J'$, $a_{ij} = 0$ if $i = k_j$; and $a_{ij} = \lambda$ if $i \neq k_j$;

where, apart from $\lambda < 1$, we impose a sparsity constraint: $\lambda \approx 1$ for one value of i, and $\lambda \approx 0$ for all other values of i. Then, we scale each row of $\mathbf{A}$ to unit sum. It is worth noticing that a factor 1 is set for the location of each spectrum of $\mathbf{X}_{1_{Init}}$. For example, with $J = 5$ and $J' = 3$:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0.01 & 0.95 & 0.02 & 0.01 & 0.01 \\ \dots & \dots & \dots & \dots & \dots \\ 0.03 & 0.93 & 0.01 & 0.02 & 0.01 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix} \quad (6)$$

In Eq. (6), we choose solely coefficients which are close to 0 or close to 1, to respect the sparsity constraint.

The criterion presented in this section is minimized starting with the initialization matrices above, with the algorithm whose implementation is detailed in Section 4.

4 SPARSE HALS-NMF IMPLEMENTATION

In this section, we present explanations of the criteria, taking into account the physical realities, proposed in the previous section and the implemented algorithm.

We propose a hierarchical implementation of the combined generalized alternating least squares algorithm proposed in (A. Cichocki, et al.,2009) (chapter 4). With a HALS scheme, we encourage the sparsity of the mixing matrix and the smoothness of the endmember spectra (A. Cichocki, et al.,2009).
Also, the HALS convergence speed outperforms the one of projected gradient (W. Chen, et al., March 2012). However it is sensitive to the scaling of the initial matrices (N. Gillis, et al., April 2010). For example, as explained in (N. Gillis, et al., April 2010), if the magnitude of the coefficients in the initial $\mathbf{A}$ and $\mathbf{X}$ are not of the same order of magnitude as the values in $\mathbf{Y}$, this will lead to rank deficient approximations and numerical problems.
The update rules derived from Eq. (4) are as follows:

$$\mathbf{X} \Leftarrow \frac{\mathbf{A}^T\mathbf{Y} + \beta\mathbf{X}_{pure}}{||\mathbf{A}||_F^2 + \beta} \quad (7)$$

$$\mathbf{A} \Leftarrow \frac{\mathbf{Y}\mathbf{X}^T - \frac{\alpha}{2}\mathbf{1}}{||\mathbf{X}||_F^2} \quad (8)$$

where $\mathbf{1}$ denotes a matrix with 1-valued coefficients.

Let ξ be a scalar whose magnitude is smaller than any other value in the considered problem. We ensure the non-negativity of any data 'd' while performing the following operation: $d \leftarrow \max[\xi, d]$. In the following, we denote this operation as $[d]_+$. The following algorithm is a hierarchical implementation of the update rules of Eqs. (7) and (8):

Algorithm 1: *HALS – NMF*.

Set convergence parameter tol.
Initialize the criterion C to a large value.
Y is the observation matrix,
Add in **Y** a column of **1**'s: $\mathbf{Y} = [\mathbf{1}, Y]$.
Initialize **X** as in Eq. (5) and insert in **X** a column of **1**'s: $\mathbf{X} = [\mathbf{1}, X]$.
Initialize **A** as in Eq. (6).
Set $\mathbf{B} = \mathbf{X}^T$ and $\mathbf{B}_{pure} = \mathbf{X}_{pure}^T$;
while C > tol **do**
  update **B**:
  $\mathbf{W} = \mathbf{Y}^T\mathbf{A}$;
  $\mathbf{V} = \mathbf{A}^T\mathbf{A}$;
  for $j \in [1, J[$ **do**
    Update rule for $\mathbf{b}_j$ (see Eq. (9))
  end for
  Update the columns of **B** as $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_J$.
  update **A**;
  $\mathbf{P} = \mathbf{YB}$
  $\mathbf{Q} = \mathbf{B}^T\mathbf{B}$
  for $j \in [1, J[$ **do**
    Update rule for $\mathbf{a}_j$ (see Eqs. (10) and (11))
  end for
  Update the columns of **A** as $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_J$.
  Set $\mathbf{X} = \mathbf{B}^T$;
  Compute $C = ||\mathbf{Y} - \mathbf{AX}||_F^2$.
end while

We use the term hierarchical because the columns $\mathbf{b}_j$ of matrix **B** and $\mathbf{a}_j$ of matrix **A** are estimated successively (see the two 'for' loops in algorithm 1).

The update rule for $\mathbf{b}_j$ is as follows:

$$\mathbf{b}_j \Leftarrow \left[\frac{\mathbf{b}_j + \mathbf{w}_j - \mathbf{B}\mathbf{v}_j + \beta\mathbf{b}_{pure_j}}{\mathbf{v}_{jj} + \beta}\right]_+ \quad (9)$$

where $\mathbf{b}_{pure_j}$ refers to the columns of $\mathbf{B}_{pure}$.

The update rule for $\mathbf{a}_j$ is as follows:

$$\mathbf{a}_j \Leftarrow \left[\frac{\mathbf{a}_j q_{jj} + \mathbf{p}_j - \mathbf{A}\mathbf{q}_j - \alpha/2}{\mathbf{q}_{jj}}\right]_+ ; \quad (10)$$

$$\mathbf{a}_j \Leftarrow \mathbf{a}_j / \|\mathbf{a}_j\|_2 \quad (11)$$

The normalization in Eq. (11) helps mitigating the effects of rotation indeterminacies on matrix **A**.

In the next section, we exemplify the proposed method on hypespectral acquisitions of scenes containing vegetation. We first consider artificial mixtures, and then spectra extracted from an aerial HSI.

5 RESULTS

This section presents some experiments performed on data to illustrate the performance of the proposed non-negative matrix factorization algorithm.

In the first place, we extract pigment spectra from a leaf reflectance; afterwards, we distinguish between vegetated areas, soil and water in an aerial HSI.

To achieve our acquisitions, we have based on some facts: In (N. Dobigeon, et al., June 2013) , Dobigeon *et. al.* clearly emphasized the degradation of a mixing model when the wavelengths of interest cover both visible and infra-red (IR) domains:

Indeed, the pigment concentration rules the vegetation reflectance in the visible domain, whereas the internal structure of the leaf does it in the IR domain. We acquired HSIs with L=500 to 830 bands, between 400 and 700 to 900 nm: In the visible and the very near IR domains. Some experiments permitted to rule the regularization parameters to $\alpha = 0.2$ and $\beta = 0.6$.

5.1 Artificial Spectral Mixtures

Figure 1a) presents three leaves with a homogeneous color. We consider the content of these leaves as being as pure as possible, to get a first set of three reference spectra. These three spectra distinguish clearly from each other: The green spectrum of Fig. 1b) is the one of chlorophyll, also represented in (AK. Mahlein et al., 2005), the yellow one is the spectrum of carotenoid, which bears high values in the yellow and orange wavelengths, the red spectrum is the one of the anthocyanin. Indeed the red color of the leaf in Fig. 1a) is characteristic of this phenolic compound, that young leaves produce to protect themselves from UV rays.

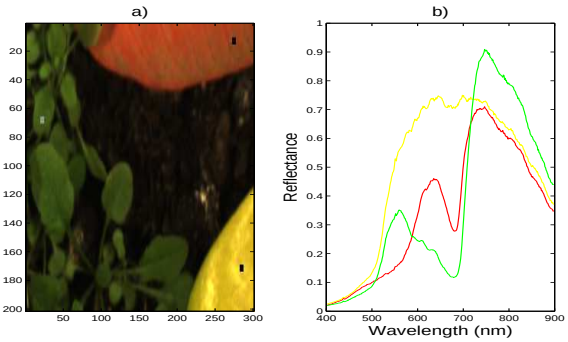

Figure 1: HSI data: a) three leaves with homogeneous color containing pure pigments: chlorophyll, carotenoid, anthocyanin; b) spectra extracted from each leaf.

Starting from these three spectra, we create artificially mixed spectra. The mixing matrix is chosen

as:

$$\mathbf{A} = \begin{pmatrix} 0 & 1.0000 & 0 \\ 0 & 0 & 1.0000 \\ 1.0000 & 0 & 0 \\ 0.7000 & 0.2000 & 0.1000 \\ 0.6000 & 0.2500 & 0.2500 \\ 0.2200 & 0.7000 & 0.0800 \\ 0.1000 & 0.6500 & 0.2500 \\ 0.1500 & 0.0100 & 0.8400 \\ 0.0300 & 0.0200 & 0.9500 \\ 0.9800 & 0.0100 & 0.0100 \end{pmatrix} \quad (12)$$

The estimated mixing matrix is:

$$\hat{\mathbf{A}} = \begin{pmatrix} 0.0000 & 1.0000 & 0.0000 \\ 0.0252 & 0.0000 & 0.9748 \\ 0.8143 & 0.0594 & 0.1264 \\ 0.5688 & 0.2469 & 0.1843 \\ 0.4460 & 0.2633 & 0.2907 \\ 0.1755 & 0.7296 & 0.0949 \\ 0.0838 & 0.6680 & 0.2483 \\ 0.1427 & 0.0046 & 0.8527 \\ 0.0485 & 0.0059 & 0.9456 \\ 0.7979 & 0.0684 & 0.1336 \end{pmatrix} \quad (13)$$

One column corresponds to one endmember, and one row to one location in the leaf. Figure 2 shows the mixed and unmixed spectra. The unmixed spectra are almost perfectly superimposed to the chosen source spectra. The relative error between actual and estimated spectra is 3.8%.

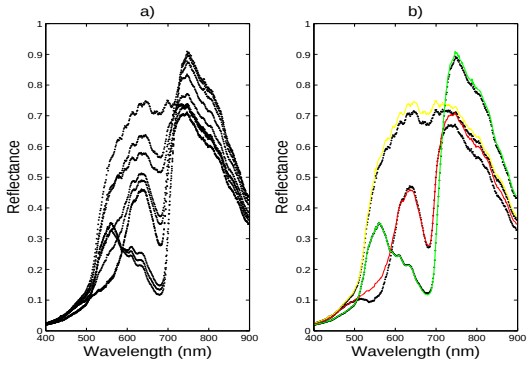

Figure 2: Mixed and unmixed spectra.

5.2 Vegetated Area Delimitation in Aerial HSI

We now process an aerial HSI, that we acquired from an airplane at the height of 500 meters. The spatial resolution is as follows: 1 row accounts for 1 m., and 1 column accounts for 0.25 m. We choose $S = I_1 I_2$: we process all spectral pixels. The whole image is provided at (Institut Fresnel). Fig. 3 shows the results obtained on a sub-image with 160 rows and 500 columns and $L = 830$ spectral bands. We fixed the number of endmembers to four.

The left column in Fig. 3 presents the four estimated endmembers; the right column displays the abundance maps from low contribution (blue) to high contribution (red). The first endmember is similar in shape to the spectrum of chlorophyll obtained in subsection 5.1. The gap around 760 nm is due to oxygen absorbance. So we can assert that, in the regions corresponding to abundance values above a certain threshold, such as 70%, the vegetation dominates. The second endmember is associated with high abundance values on the water of the river, and on the dirt roads, from which we infer that it represents the spectrum of sunlight: It is most reflected in these regions of the scene. The third endmember accounts for soil.

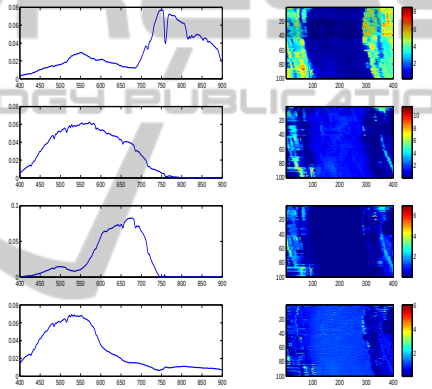

Figure 3: Estimated endmember spectra and abundance map.

6 CONCLUSIONS

We consider unmixing of hyperspectral pixels by NMF. We introduce spatial and spectral information in the initialization mixing and endmember matrices, and in the criterion which is minimized to perform NMF. For this, we select automatically with vertex component analysis the purest spectra among the spectral pixels. Firstly, these pure spectra are used as a subset of initialization endmembers for NMF. The mixing matrix is initialized as a sparse matrix, and accounts for the location, in the HSI, of the initialization endmembers. Secondly, we use the pure spectra as references in a regularization term of the criterion minimized for NMF. This term introduces some spectral prior: It enforces the endmember spectra estimated by NMF to exhibit some physical significance. Apart from this term, which is the most novel aspect of this work, we use a second regularization term

which accounts for a spatial prior knowledge on the processed HSI: The materials in the scene are often grouped in regions and consequently the mixing matrix should be sparse, which is encouraged through its l_1-norm. We first imaged leaves, and then an aerial partly vegetated scene. Considering their shape, the proposed method permits to better interpret the endmember spectra as a comparative implementation of NMF. We could reliably delimitate the vegetated areas in remote sensing context.

ACKNOWLEDGEMENTS

The authors would like to thank the Bavarian Research Foundation (BFS: Bayerische Forschungsstiftung) for supporting this researches done at the Fraunhofer IIS in Frth.

We thank the firm "Action Air Environnement" for providing us with airplanes.

REFERENCES

H. Kim and K. Park, "Non-negative Matrix Factorization Based on Alternating Non-negativity Constrained Least Squares and Active Set Method", *SIAM J.'l Matrix Analysis and Applications*, vol. 30(2), pp. 713-730, 2008.

A. Cichocki, S. Amari, R. Zduneck, and A.H. Phan, "Non-negative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation", *Wiley- Blackwell*, 2009.

W. Chen and M. Guillaume, "HALS-based NMF with flexible constraints for hyperspectral unmixing", *EURASIP Journal on Advances in Signal Processing*, vol. 2012(54), March 2012.

N. Dobigeon and C. Fevotte, "Robust nonnegative matrix factorization for nonlinear unmixing of hyperspectral images," *IEEE WHISPERS*, Gainesville, FL, June 2013.

Ma et al, "A signal processing perspective on hyperspectral unmixing", *IEEE Signal Processing Magazine*, vol. 31(1), pp. 67-81, January 2014.

JMP. Nascimento and JMB. Dias, "Vertex component analysis: a fast algorithm to unmix hyperspectral data", *IEEE TGRS*, vol. 43(4), pp. 898-910, April 2005.

J. Marot and S. Bourennane, "Leaf marker spectra identification by hyperspectral image acquisition and vertex component analysis", *EUVIP'13*, pp. 190-195, June 2013.

Q. Du, I. Kopriva, and H. Szu, "Investigation on Constrained Matrix Factorization for Hyperspectral Image Analysis", In procs. of *IEEE International Geoscience and Remote Sensing Symposium*, vol. 6, pp. 4304-4306, Seoul, July 2005.

N. Gillis and F. Glineur, "Using underapproximations for sparse nonnegative matrix factorization", *Pattern Recognition*, Vol. 43(4), April 2010, pp. 1676-1687.

P. O. Hoyer, "Non-negative matrix factorization with sparseness constraints", *Journal of Machine Learning Research*, vol. 5, pp. 1457-1469, 2004.

D. Muti, S. Bourennane, and J. Marot, "Lower-Rank Tensor Approximation and Multiway Filtering," *SIAM Journal on Matrix Analysis and Applications (SIMAX)*, vol. 30(3), pp. 1172-1204, September 2008.

AK. Mahlein *et. al*, "Recent advances in sensing plant diseases for precision crop protection," *Eur J Plant Pathol*, vol. 133, pp. 197-209, May 2012.

John P. Kerekes et. al., "SHARE 2012: subpixel detection and unmixing experiments", In Procs. SPIE 8743, *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery*, May 18, 2013.

http://www.fresnel.fr/perso/marot/Documents/Aerial_HSI_EUVIP.bmp

AK. Mahlein et. al., "Hyperspectral Imaging for Small-Scale Analysis of Symptoms Caused by Different Sugar Beet Diseases", *Plant Methods*, vol. 8, no 3, pp. 1-13, 2012.

L. Chaerle et al., "Multicolor fluorescence imaging for early detection of the hypersensitive reaction to tobacco mosaic virus," *Journal of Plant physiology*, vol. 164, pp. 253-262, March 2007.