UNSUPERVISED NEIGHBOR DEPENDENT NONLINEAR UNMIXING

Rita Ammanouil, André Ferrari, Cédric Richard

Laboratoire Lagrange, Université de Nice Sophia-Antipolis, France
{rita.ammanouil, andre.ferrari, cedric.richard}@unice.fr

ABSTRACT

This communication proposes an unsupervised neighbor dependent nonlinear unmixing algorithm for hyperspectral data. The proposed mixing scheme models the reflectance vector of a pixel as the sum of a linear combination of the endmembers plus a nonlinear function acting on neighboring spectra. The nonlinear function belongs to a reproducing kernel Hilbert space. The observations themselves are considered as the endmember candidates, and the group lasso regularization is used to enable selecting the purest pixels among the candidates. Experiments on synthetic data demonstrate the effectiveness of the proposed approach.

Index Terms— Hyperspectral imaging, nonlinear unmixing, nonlinear endmember extraction, sparse regularization.

1. INTRODUCTION

Each pixel in an hyperspectral image is characterized by a reflectance vector (spectrum), estimated over hundreds of contiguous spectral bands. Due to the relatively low spatial resolution of hyperspectral sensors, the surface imaged by a pixel usually contains more than one constituent material [1, 2]. As a result, the pixel's spectrum is a mixture of the constituent materials spectra, also known as endmembers. The linear mixing model (LMM) [3] is the most prevalent model that represents this mixture. It is given by:

$$\boldsymbol{s}_n^{\text{lin}} = \sum_{i=1}^{M} a_{i,n}\boldsymbol{r}_i + \boldsymbol{e}_n,\ \forall n = 1, \dots N, \qquad (1)$$

where $\boldsymbol{s}_n^{\text{lin}}$ is the L-dimensional spectrum of a linearly mixed pixel indexed by n, L is the number of frequency bands, M denotes the number of endmembers, $a_{i,n}$ is the abundance of the i-th endmember in the n-th pixel, $\boldsymbol{r}_i$ is the L-dimensional spectrum of the i-th endmember, $\boldsymbol{e}_n$ is a vector of Gaussian white noise, and N is the number of observations. Note that all vectors are column vectors. According to the LMM, the spectrum of the observed pixel is equal to a weighted sum of the endmembers' spectra. The weights multiplying the endmembers, also known as abundances, are positive and usually sum to one [4]. The LMM is used in the overwhelming majority of hyperspectral imaging applications. However, there exists certain scenarios where the spectral mixture exhibits strong nonlinear effects [5, 6]. In this case, the LMM is no longer a sufficiently representative model. Several nonlinear mixing models (NLMM's) were proposed to take into account nonlinearity. They usually take the form of a linear mixture of the endmembers plus a nonlinear one. For example, in bilinear models [7–10] the nonlinear term is a weighted sum over the pairwise products of endmembers. Recently, the authors of [11] proposed a generalization of bilinear models. Finally, non parametric approaches to NLMM's estimation based on reproducing kernel Hilbert spaces (RKHS's) [12], and geometry have been proposed [13, 14]. We consider the following new kernel-based NLMM:

$$\boldsymbol{s}_n = \sum_{i=1}^{M} a_{i,n}\boldsymbol{r}_i + \boldsymbol{f}(\boldsymbol{v}_n) + \boldsymbol{e}_n,\ \forall n = 1, \dots N, \qquad (2)$$

where $\boldsymbol{s}_n$ represents the L-dimensional spectrum of a nonlinearly mixed pixel, $\boldsymbol{f}$ is a nonlinear and non parametric function in a vector-valued RKHS [15], $\boldsymbol{v}_n = \text{vec}(\{\boldsymbol{s}_i\}_{i\in\mathcal{N}_n})$, and $\mathcal{N}_n$ contains c neighboring pixels indices, in other terms $\boldsymbol{v}_n$ contains the column-wise concatenation of the spectra of c neighbors of the n-th pixel. Compared to (1), the additional term $\boldsymbol{f}(\boldsymbol{v}_n)$ takes into account the nonlinear fluctuation present in $\boldsymbol{s}_n$. Unlike NLMM's in the literature, the nonlinear term involves the neighbors spectra rather than the endmembers. The choice to act on neighboring spectra is motivated by the adjacency effect [16, 17] which states that the neighboring surfaces of the target pixel also contribute to its spectrum. Even after atmospheric correction aiming at removing adjacent contributions, errors occurred in the estimated spectrum $\boldsymbol{s}_n$ can damage the quality of the extracted information if not taken into account [18]. Similarly to [12], the use of functions in RKHS's allows to model the nonlinearity in a rather general and non parametric way. This is especially interesting when dealing with nonlinearities due to atmospheric correction where it is difficult if not impossible to have an explicit expression. In addition to this, $\boldsymbol{f}$ being a vector-valued function distinguishes the proposed model from the NLMM's in the literature. The nonlinear contribution at each spectral band, i.e. each element of $\boldsymbol{f}(\boldsymbol{v}_n)$, is a function of $\boldsymbol{v}_n$ containing reflectances at all bands. Therefore, it is possible to incorporate inter-band nonlinear interactions.

This work was partly supported by the Agence Nationale pour la Recherche, France, (Hypanema project, ANR-12-BS03-003), and the regional council of Provence-Alpes-Côte d'Azur.

Naturally, a NLMM requires a nonlinear endmember extraction algorithm (EEA). The authors of [13, 14] propose geometry-based nonlinear EEA's. They express linear algorithms in terms of Euclidean distances rather than spectral coordinates and then replace them by non-Euclidean ones. The authors of [19] use non negative matrix factorization to jointly estimate the endmembers and abundances for the Fan model [8]. The authors of [20] propose a Bayesian framework to estimate the endmembers and abundances for the post-nonlinear mixing model (PPNM) [10]. The algorithm proposed in this work also enables the extraction of the endmembers and the estimation of the abundances and the nonlinear contributions according to model (2). This is possible through the use of collaborative sparse regression [21] to identify the endmembers within the observations [22]. The optimization problem is convex and solved using the alternating direction method of multipliers (ADMM) [23].

The paper is organized as follows. Section 2 describes the NLMM, section 3 developps the estimation algorithm, and section 4 validates the proposed model using simulated data.

2. NONLINEAR FUNCTION

The nonlinear function $\boldsymbol{f}$ used in (2) is a vector-valued function, $\boldsymbol{f}(\boldsymbol{v}_n)$ is a vector with L components representing the nonlinear spectral contribution at each band:

$$\begin{array}{rccc} \boldsymbol{f}: & \mathbb{R}^{cL} & \rightarrow & \mathbb{R}^{L} \\ & \boldsymbol{v}_n & \rightarrow & \boldsymbol{f}(\boldsymbol{v}_n). \end{array} \tag{3}$$

Ideally, $\boldsymbol{f}$ should be a function of the noiseless linear mixtures present in neighboring pixels. However, this information is unknown a priori, and would lead to a non convex optimization problem. A possible solution, is to approximate the linear mixtures by the pixels spectra themselves through the use of $\boldsymbol{v}_n$. The function $\boldsymbol{f}$ belongs to a RKHS $\mathcal{H}_k$ of vector-valued functions [15] associated to the kernel function:

$$\begin{array}{rccc} \boldsymbol{k}: & \mathbb{R}^{cL} \times \mathbb{R}^{cL} & \rightarrow & \mathbb{R}^{L\times L} \\ & (\boldsymbol{v}_n, \boldsymbol{v}_m) & \rightarrow & \boldsymbol{k}(\boldsymbol{v}_n, \boldsymbol{v}_m). \end{array} \tag{4}$$

Unlike a scalar-valued RKHS, $\boldsymbol{k}(\boldsymbol{v}_n, \boldsymbol{v}_m) \in \mathbb{R}^{L\times L}$ rather than $\mathbb{R}$. $\boldsymbol{K} = \{\boldsymbol{k}(\boldsymbol{v}_n, \boldsymbol{v}_m)\}_{n,m=1}^{N}$ is the Gram matrix resulting from the evaluation of the kernel function at all observation couples. It is a $N \times N$ block matrix where each block has $L \times L$ components. As a result, $\boldsymbol{K}$ is a $L\,N \times L\,N$ scalar-valued matrix. The authors of [15] provide several examples for defining $\boldsymbol{k}(\boldsymbol{v}_n, \boldsymbol{v}_m)$ based on scalar-valued kernels such as the Gaussian kernel. According to the representer theorem [15], the evaluation of $\boldsymbol{f}$ at a point can be expressed as an expansion of the kernel function over all training points:

$$\boldsymbol{f}(\boldsymbol{v}_n) = \textstyle\sum_{i=1}^{N} \boldsymbol{k}(\boldsymbol{v}_n, \boldsymbol{v}_i)\boldsymbol{\alpha}_i, \tag{5}$$

where $\boldsymbol{\alpha}_i \in \mathbb{R}^L$ are the kernel coeficients. Estimating the nonlinear function reduces to estimating these coefficients.

The definition of $\boldsymbol{k}(\boldsymbol{v}_n, \boldsymbol{v}_m)$ allows to incorporate or ignore inter-band interactions. For example, letting the (i, j)-th component in $\boldsymbol{k}(\boldsymbol{v}_n, \boldsymbol{v}_m)$ be a function of the elements in $\boldsymbol{v}_n$ at band i and those in $\boldsymbol{v}_m$ at band j incorporates inter-band interactions. Whereas setting the off-diagonal components to zero, and letting the diagonal components be functions of $\boldsymbol{v}_n$ and $\boldsymbol{v}_m$ at the same band ignores inter-band interactions. The same nonlinear function $\boldsymbol{f}$ is defined for all pixels rather than having a different one per-pixel. As a result, the estimation algorithm is coupled w.r.t. all the pixels. On the one hand, this coupling increases the complexity of the algorithm. On the other hand, it allows to examine and possibly interpret the nonlinear interactions between the pixels.

3. ESTIMATION ALGORITHM

3.1. Optimization problem

The optimization problem associated with model (2) is obtained by minimizing a data fidelity error, and incorporating a priori information through regularizations and constraints:

$$\begin{array}{ll} \min\limits_{\boldsymbol{f}\in\mathcal{H}_k, \boldsymbol{A}} & \frac{1}{2}\sum\limits_{n=1}^{N} \|\boldsymbol{s}_n - \boldsymbol{R}\boldsymbol{a}_n - \boldsymbol{f}(\boldsymbol{v}_n)\|^2 + \lambda\mathcal{J}_1(\boldsymbol{f}) \\ & + \mu\mathcal{J}_2(\boldsymbol{A}) \\ \text{subject to} & a_{i,n} \succeq 0 \quad \forall\, i = 1, \cdots M\,, n = 1, \cdots N, \\ & \sum_{i=1}^{M} a_{i,n} = 1 \quad \forall\, n = 1, \cdots, N, \end{array} \tag{6}$$

where $\boldsymbol{a}_n = [a_{1,n}, \ldots, a_{M,n}]^\top$, $\boldsymbol{R} = [\boldsymbol{r}_1, \ldots, \boldsymbol{r}_M]$, and $\boldsymbol{A} = [\boldsymbol{a}_1, \ldots, \boldsymbol{a}_N]$, λ and μ are tuning parameters. $\boldsymbol{A}$ is constrained to be positive, and each column of $\boldsymbol{A}$ has to sum to one. $\mathcal{J}_1(\boldsymbol{f})$ is set to the squared ℓ_2-norm of $\boldsymbol{f}$ in $\mathcal{H}_k$:

$$\mathcal{J}_1(\boldsymbol{f}) = \frac{1}{2}\|\boldsymbol{f}\|^2_{\mathcal{H}_k}. \tag{7}$$

The ℓ_2-norm of $\boldsymbol{f}$ in $\mathcal{H}_k$ constrains the regularity and smoothness of $\boldsymbol{f}$ [24]. As for $\mathcal{J}_2(\boldsymbol{A})$, the choice of this regularisation is adapted to the unsupervised and supervised cases, i.e. when the endmembers are unknown and known respectively.

With a slight abuse of notations, we assume in the unsupervised case that $\boldsymbol{R}$ is a matrix containing a large number of candidate endmembers, among which are the actual endmembers. $\boldsymbol{R}$ can be a dictionary of measured spectra [25] or a set of potential endmembers spectra selected from the available observations themselves [22, 26]. As a result, the abundance matrix is row sparse. In accordance with this a priori information, the group lasso regularization promotes zero rows. Let $\boldsymbol{a}_{\lambda_i}$ denote the i-th row in $\boldsymbol{A}$, this regularization is given by:

$$\mathcal{J}_2(\boldsymbol{A}) = \sum_{i=1}^{M} \|\boldsymbol{a}_{\lambda_i}\|_2, \tag{8}$$

In the supervised case, $\boldsymbol{R}$ is considered to contain the actual endmembers, and the Tikhonov regularization that promotes smoothness is used:

$$\mathcal{J}_2(\boldsymbol{A}) = \|\boldsymbol{A}\|^2_F. \tag{9}$$

3.2. Iterative algorithm

In this section, we develop the steps in the alternating direction method of multipliers (ADMM) [23] used to solve problem (6). The ADMM splits the optimization problem into simpler sub-problems. First, problem (6) is reformulated as:

$$\begin{array}{ll}\min\limits_{\boldsymbol{f}\in\mathcal{H}_k,\boldsymbol{X},\boldsymbol{Z}} & \frac{1}{2}\sum_{n=1}^{N}\|\boldsymbol{s}_n-\boldsymbol{R}\boldsymbol{x}_n-\boldsymbol{f}(\boldsymbol{v}_n)\|^2+\lambda\mathcal{J}_1(\boldsymbol{f})\\ & +\mu\mathcal{J}_2(\boldsymbol{Z})+\mathcal{I}_{\mathbb{R}^+}(\boldsymbol{Z})\\ \text{subject to} & \boldsymbol{A}\boldsymbol{X}+\boldsymbol{B}\boldsymbol{Z}=\boldsymbol{C},\end{array} \tag{10}$$

with

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{I}_M\\ \boldsymbol{1}^\top\end{pmatrix},\ \boldsymbol{B}=\begin{pmatrix}-\boldsymbol{I}_M\\ \boldsymbol{0}^\top\end{pmatrix},\ \boldsymbol{C}=\begin{pmatrix}\boldsymbol{0}\\ \boldsymbol{1}^\top\end{pmatrix}, \tag{11}$$

where $\boldsymbol{X}$ and $\boldsymbol{Z}$ are the ADMM variables. Compared to the original problem (6), the data fidelity term and $\mathcal{J}_2(\boldsymbol{Z})$ are functions of two different variables, $\boldsymbol{X}$ and $\boldsymbol{Z}$ respectively. A consensus constraint between $\boldsymbol{X}$ and $\boldsymbol{Z}$ is added. The augmented Lagrangian associated to (10) is given by:

$$\begin{array}{c}\mathcal{L}_\rho(\boldsymbol{f},\boldsymbol{X},\boldsymbol{Z},\boldsymbol{V})=\frac{1}{2}\sum_{n=1}^{N}\|\boldsymbol{s}_n-\boldsymbol{R}\boldsymbol{x}_n-\boldsymbol{f}(\boldsymbol{v}_n)\|^2\\ +\lambda\mathcal{J}_1(\boldsymbol{f})+\mu\mathcal{J}_2(\boldsymbol{Z})+\text{tr}(\boldsymbol{V}^\top(\boldsymbol{A}\boldsymbol{X}+\boldsymbol{B}\boldsymbol{Z}-\boldsymbol{C}))\\ +\mathcal{I}_{\mathbb{R}^+}(\boldsymbol{Z})+\frac{\rho}{2}\|\boldsymbol{A}\boldsymbol{X}+\boldsymbol{B}\boldsymbol{Z}-\boldsymbol{C}\|_{\text{F}}^2,\end{array}$$

where $\boldsymbol{V}$ is the matrix of Lagrange multipliers, and ρ is the penalty parameter. The ADMM algorithm consists of iterating 3 main steps developed hereafter.

$\{\boldsymbol{X},\boldsymbol{f}\}$ **minimisation step**: This step consists of minimizing the augmented Lagrangian with respect to $\{\boldsymbol{X},\boldsymbol{f}\}$. To handle this problem, we rewrite it in an equivalent form:

$$\begin{array}{ll}\min\limits_{\boldsymbol{f}\in\mathcal{H}_k,\boldsymbol{X},\boldsymbol{E}} & \frac{1}{2}\sum\limits_{n=1}^{N}\|\boldsymbol{e}_n\|^2+\frac{\lambda}{2}\|\boldsymbol{f}\|_{\mathcal{H}_k}^2+\text{tr}(\boldsymbol{V}^\top\boldsymbol{A}\boldsymbol{X})\\ & +\frac{\rho}{2}\|\boldsymbol{A}\boldsymbol{X}+\boldsymbol{B}\boldsymbol{Z}-\boldsymbol{C}\|_{\text{F}}^2\\ \text{subject to} & \boldsymbol{e}_n=\boldsymbol{s}_n-\boldsymbol{R}\boldsymbol{x}_n-\boldsymbol{f}(\boldsymbol{v}_n).\end{array} \tag{12}$$

The solution of (12) can be found by solving its dual:

$$\max_{\bar{\boldsymbol{w}}}\quad -\tfrac{1}{2}\bar{\boldsymbol{w}}^\top\boldsymbol{Q}\bar{\boldsymbol{w}}+\bar{\boldsymbol{w}}^\top\boldsymbol{p}, \tag{13}$$

with:

$$\begin{cases}\boldsymbol{Q} &= \boldsymbol{I}_{L\,N}+\frac{1}{\lambda}\boldsymbol{K}+\frac{1}{\rho}\boldsymbol{I}_N\otimes(\boldsymbol{R}(\boldsymbol{A}^\top\boldsymbol{A})^{-1}\boldsymbol{R}^\top),\\ \boldsymbol{p} &= \text{vec}(\boldsymbol{S}+\frac{1}{\rho}\boldsymbol{R}(\boldsymbol{A}^\top\boldsymbol{A})^{-1}\boldsymbol{A}^\top(\boldsymbol{V}+\rho(\boldsymbol{B}\boldsymbol{Z}-\boldsymbol{C}))),\end{cases}$$

where $\boldsymbol{W}=[\boldsymbol{w}_1,\cdots,\boldsymbol{w}_N]$ is a matrix of Lagrange multipliers, $\bar{\boldsymbol{w}}=\text{vec}(\boldsymbol{W})$. After solving (13), the optimal variables are deduced using the following equations:

$$\begin{cases}\boldsymbol{E}^* &= \boldsymbol{W}^*,\\ \boldsymbol{X}^* &= (\boldsymbol{A}^\top\boldsymbol{A})^{-1}(\frac{\boldsymbol{R}^\top\boldsymbol{W}^*}{\rho}-\boldsymbol{A}^\top(\frac{\boldsymbol{V}}{\rho}+\boldsymbol{B}\boldsymbol{Z}-\boldsymbol{C})),\\ \boldsymbol{f}^*(\cdot) &= \frac{1}{\lambda}\sum_{i=1}^{N}\boldsymbol{k}(\cdot,\boldsymbol{v}_i)\boldsymbol{w}_i^*.\end{cases}$$

In order to reduce the computational complexity of this step, the conjugate gradient (CG) method can be used to solve (13) as will be done in all the experiments.

$\boldsymbol{Z}$ **minimization step**: In the case of (8), the optimization problem in this step is separable for each row in $\boldsymbol{Z}$:

$$\min_{\boldsymbol{z}}\quad \tfrac{1}{2}\|\boldsymbol{z}-\boldsymbol{v}\|_2^2+\alpha\|\boldsymbol{z}\|_2+\mathcal{I}_{\mathbb{R}^+}(\boldsymbol{z}) \tag{14}$$

where $\boldsymbol{v}=\boldsymbol{x}+\rho^{-1}\boldsymbol{\lambda}$, $\alpha=\rho^{-1}\mu$. The vectors $\boldsymbol{\lambda}$, $\boldsymbol{x}$ and $\boldsymbol{z}$ correspond to line in $\boldsymbol{\Lambda}$, $\boldsymbol{X}$ and $\boldsymbol{Z}$, respectively. The minimizer of (14) has a unique solution given by the proximal operator:

$$\begin{cases}\boldsymbol{z}^*=\boldsymbol{0} & \text{si } \|(\boldsymbol{v})_+\|_2<\alpha\\ \boldsymbol{z}^*=\left(1-\frac{\alpha}{\|(\boldsymbol{v})_+\|_2}\right)(\boldsymbol{v})_+ & \text{sinon}\end{cases} \tag{15}$$

where $(\cdot)_+=\max(\boldsymbol{0},\cdot)$. In the case of (9), the solution is given by a simple projection on the positive orthant. The solution in this case is not reported for conciseness.

Update of the Lagrange multipliers: The last step in the algorithm consists of updating the Lagrange multipliers:

$$\boldsymbol{V}^{k+1}=\boldsymbol{V}^k+\rho(\boldsymbol{A}\boldsymbol{X}+\boldsymbol{B}\boldsymbol{Z}-\boldsymbol{C}). \tag{16}$$

Note that each step uses the latest estimate of the variables.

4. EXPERIMENTS

The performance of the proposed approach is evaluated using synthetic data sets. The number of simulated observations is $N=10^2$, and the abundances are generated using Dirichlet distributions with shape parameter 2. The endmembers are extracted from the USGS library, and have $L=112$ spectral bands. 50% of the observations are linearly mixed (1), and the remaining 50% are generated according to the PPNM:

$$\boldsymbol{s}_n=\sum_{i=1}^{M}a_{i,n}\boldsymbol{r}_i+u(\sum_{i=1}^{M}a_{i,n}\boldsymbol{r}_i).^2+\boldsymbol{e}_n, \tag{17}$$

where u controls the strength of the nonlinear part, and the dot before the power of two indicates that the operation is applied element wise. The same value for u is considered for all nonlinearly mixed pixels. Finally, the endmembers were inserted among the observations to ensure the pure pixel assumption.

4.1. Unsupervised case

We first examine the endmembers extraction accuracy of the proposed nonlinear approach denoted as UNDU (unsupervised and nonlinear neighbor dependent unmixing), NFINDR [27], and DMaxD [14] in the noiseless case. NFINDR and DMaxD require knowing the number of endmembers to be extracted. DMaxD also requires defining an appropriate distance metric. The PPNM distance defined and proposed in [14] is used. UNDU, does not require knowing the number of endmembers. It requires tuning λ and μ. A Gaussian kernel per band is used to define the diagonal elements of (4), and the off-diagonal elements are set to zero.

The observations themselves serve as the candidate endmembers in $\boldsymbol{R}$. Table 1 reports the average of the M spectral angles (SA's) in radian between the identified and the true endmembers. The average SA's are reported for the 3 algorithms for different values of M and u. For UNDU, $\lambda = 0.01$ and μ is set to values in the interval $[0.1\ 1]$. Table (1) shows that UNDU had a zero average SA in all cases, meaning that it correctly identified all the endmembers. DMaxD sometimes failed in correctly identifying all the endmembers resulting in nonzero average SA's. In those cases, it misidentified only 1 out of M endmembers. NFINDR misidentified at least 1 and at most 2 endmembers in all scenarios.

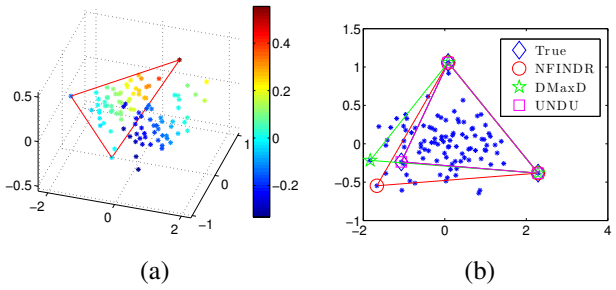
Figure 1 (a) shows the 3 dimensional scatter obtained using principal component analysis (PCA) of the data set with $M = 3$ and $u = 0.2$. The $3D$ scatter shows that the linearly mixed pixels lie in the triangle formed by the three endmembers, whereas the nonlinearly mixed pixels are shifted away (down) from the triangle's surface. Figure 1 (b) shows the 2 dimensional scatter of the same data set in addition to the simplexes formed by the true and identified endmembers for each algorithm. UNDU correctly identified the true endmembers. Both NFINDR and DMaxD mistakened one of the endmembers with a candidate that gives a simplex bigger than the true one. The reason why DMaxD had misidentifications is that only 50% of the pixels were nonlinearly mixed rather than 100% as tested in [14]. UNDU has a relatively high computational complexity due to the $\boldsymbol{X}$-minimization step. The average execution time for UNDU was 150 sec. whereas DMaxD and NFINDR required 0.01 and 0.05 sec. respectively.

4.2. Supervised case

Finally, the proposed approach was evaluated in the supervised case. To this end, we used the data set created with $M = 4$ endmembers and $u = 0.1$ and added Gaussian noise to have SNR $= 50$ dB. The tuning parameters were set to $\lambda = 0.01$ and $\mu = 0.001$. The first row of Figure 2 shows the true and estimated spectra for the linear and nonlinear parts in two of the nonlinearly pixels. In particular, the results are shown for pixels 17 and 14, the pixels with the best and worst reconstruction errors respectively. The second row of Figure 2 shows the image of the true (left) and estimated (right) nonlinear part for all pixels. The proposed approach scored a root mean square error (RMSE) equal to 0.0196 over the abundances and 0.0070 over the nonlinear spectra. Whereas FCLS scored a RMSE of 0.1150 over the abundances.

5. CONCLUSION

This communication proposes a novel unsupervised and non parametric neighbor dependent nonlinear unmixing algorithm. The performances were validated on synthetic data. Preliminary and promising results were obtained on real data but not reported due to lack of space.



(a) (b)

Fig. 1. Results obtained with $M = 3$ and $u = 0.2$. Left: 3D scatter of observations using PCA with 3 bands. Right: 2D scatter of points and identified endmembers.

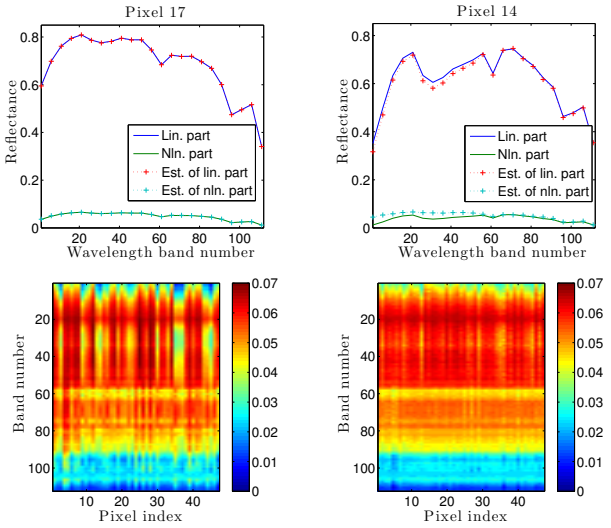


Fig. 2. First row: True and estimated spectra for linear and nonlinear parts for 2 pixels. Second row: True (left) and estimated (right) nonlinear contribution in all observations.

Table 1. SA averaged over the M estimated endmembers.

	$M = 3$	$M = 4$	$M = 5$
		$u = 0.1$	
NFINDR	0.0086	0.0274	0.0038
DMaxD	0.0086	**0.0000**	**0.0000**
UNDU	**0.0000**	**0.0000**	**0.0000**
		$u = 0.2$	
NFINDR	0.0115	0.0203	0.0086
DMaxD	0.0163	**0.0000**	0.0115
UNDU	**0.0000**	**0.0000**	**0.0000**
		$u = 0.3$	
NFINDR	0.0128	0.0595	0.0116
DMaxD	0.0128	0.0233	0.0143
UNDU	**0.0000**	**0.0000**	**0.0000**

6. REFERENCES

[1] G. Shaw and H. Burke, "Spectral imaging for remote sensing," *Lincoln Laboratory Journal*, vol. 14, no. 1, pp. 3–28, 2003.

[2] C. Chang, *Hyperspectral data exploitation: theory and applications*, John Wiley and Sons, 2007.

[3] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Transactions on Signal Processing*, vol. 19, no. 1, pp. 44–57, 2002.

[4] D. C. Heinz and C. I. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hypersectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 3, pp. 529–545, 2001.

[5] N. Dobigeon, L. Tits, B. Somers, Y. Altmann, and P. Coppin, "A comparison of nonlinear mixing models for vegetated areas using simulated and real hyperspectral data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 6, pp. 1869–1878, 2014.

[6] I. Meganem, P. Deliot, X. Briottet, Y. Deville, and S. Hosseini, "Linear–quadratic mixing model for reflectances in urban environments," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 1, pp. 544–558, 2014.

[7] J. Nascimento and J. Bioucas-Dias, "Nonlinear mixture model for hyperspectral unmixing," in *SPIE*, 2009, pp. 74770I–74770I.

[8] W. Fan, B. Hu, J. Miller, and M. Li, "Comparative study between a new nonlinear model and common linear model for analysing laboratory simulated-forest hyperspectral data," *International Journal of Remote Sensing*, vol. 30, no. 11, pp. 2951–2962, 2009.

[9] A. Halimi, Y. Altmann, N. Dobigeon, and J-Y Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 11, pp. 4153–4162, 2011.

[10] Y. Altmann, A. Halimi, N. Dobigeon, and J-Y Tourneret, "Supervised nonlinear spectral unmixing using a postnonlinear mixing model for hyperspectral imagery," *IEEE Transaction on Image Processing*, vol. 21, no. 6, pp. 3017–3025, 2012.

[11] R. Heylen and P. Scheunders, "A multilinear mixing model for nonlinear spectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. PP, no. 99, pp. 1–12, 2015.

[12] J. Chen, C. Richard, and P. Honeine, "Nonlinear estimation of material abundances in hyperspectral images with l1-norm spatial regularization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 5, pp. 2654 – 2665, 2014.

[13] R. Heylen, D. Burazerovic, and P. Scheunders, "Non-linear spectral unmixing by geodesic simplex volume maximization," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 3, pp. 534–542, 2011.

[14] R. Heylen, P. Scheunders A. Rangarajan, and P. Gader, "Nonlinear unmixing by using different metrics in a linear unmixing chain," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 6, pp. 2655–2664, 2014.

[15] C. Micchelli and M. Pontil, "On learning vector-valued functions," *Neural Computation*, vol. 17, no. 1, pp. 177–204, 2005.

[16] D. Tanre, P. Deschamps, P. Duhaut, and M. Herman, "Adjacency effect produced by the atmospheric scattering in thematic mapper data," *Journal of Geophysical Research: Atmospheres*, vol. 92, no. D10, pp. 12000–12006, 1987.

[17] R. Richter, M. Bachmann, W. Dorigo, and A. Muller, "Influence of the adjacency effect on ground reflectance measurements," *IEEE Geoscience and Remote Sensing Letters*, vol. 3, no. 4, pp. 565–569, 2006.

[18] D. Hadjimitsis, G. Papadavid, A. Agapiou, K. Themistocleous, M. Hadjimitsis, A. Retalis, S. Michaelides, N. Chrysoulakis, L. Toulios, and C. Clayton, "Atmospheric correction for satellite remotely sensed data intended for agricultural applications: impact on vegetation indices," *Natural Hazards and Earth System Science*, vol. 10, no. 1, pp. 89–95, 2010.

[19] O. Eches and M. Guillaume, "A bilinear–bilinear nonnegative matrix factorization method for hyperspectral unmixing," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 4, pp. 778–782, 2014.

[20] Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Unsupervised post-nonlinear unmixing of hyperspectral images using a hamiltonean monte carlo algorithm," *IEEE Transaction on Image Processing*, vol. 23, no. 6, pp. 2663–2675, 2014.

[21] M.-D. Iordache, J. Bioucas-Dias, and A. Plaza, "Collaborative sparse regression for hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 1, pp. 341–354, 2013.

[22] R. Ammanouil, A. Ferrari, C. Richard, and D. Mary, "Blind and fully constrained unmixing of hyperspectral images," *IEEE Transaction on Image Processing*, vol. 23, no. 12, pp. 5510 – 5518, 2014.

[23] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[24] T. Evgeniou, M. Pontil, and T. Poggio, "Regularization networks and support vector machines," *Advances in computational mathematics*, vol. 13, no. 1, pp. 1–50, 2000.

[25] M.-D. Iordache, A. Okujeni, S. van der Linden, J. Bioucas-Dias, A. Plaza, and B. Somers, "A multi-measurement vector approach for endmember extraction in urban environments," in *Image Information Mining Conference: The Sentinels Era*, Bucharest, Romania, 2014.

[26] E. Esser, M. Moller, S. Osher, G. Sapiro, and J. Xin, "A convex model for nonnegative matrix factorization and dimensionality reduction on physical space," *IEEE Transaction on Image Processing*, vol. 21, no. 7, pp. 3239–3252, 2012.

[27] M. E. Winter, "N-FINDR: an algorithm for fast autonomous spectral endmember determination in hyperspectral data," in *Proc. SPIE Imaging Spectrometry*, 1999.