HAL open science

Band selection in RKHS for fast nonlinear unmixing of hyperspectral images

Tales Imbiriba, José Carlos Bermudez, Cédric Richard, Jean-Yves Tourneret

To cite this version:

Tales Imbiriba, José Carlos Bermudez, Cédric Richard, Jean-Yves Tourneret. Band selection in RKHS for fast nonlinear unmixing of hyperspectral images. 23rd European Signal Processing Conference (EUSIPCO 2015), Aug 2015, Nice, France. pp. 1651-1655. hal-01485023

HAL Id: hal-01485023

https://hal.science/hal-01485023v1

Submitted on 8 Mar 2017

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.

Open Archive TOULOUSE Archive Ouverte (OATAO)
OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : http://oatao.univ-toulouse.fr/
Eprints ID : 17107

The contribution was presented at EUSIPCO 2015:
http://www.eusipco2015.org/

To cite this version : Imbiriba, Tales and Bermudez, José Carlos and Richard, Cédric and Tourneret, Jean-Yves *Band selection in RKHS for fast nonlinear unmixing of hyperspectral images.* (2015) In: 23rd European Signal Processing Conference (EUSIPCO 2015), 31 August 2015 - 4 September 2015 (Nice, France).

Any correspondence concerning this service should be sent to the repository administrator: staff-oatao@listes-diff.inp-toulouse.fr

BAND SELECTION IN RKHS FOR FAST NONLINEAR UNMIXING OF HYPERSPECTRAL IMAGES

T. Imbiriba[⋆] *J. C. M. Bermudez*[⋆] *C. Richard*[†] *J.-Y. Tourneret*[‡]

[⋆] Federal University of Santa Catarina, Florianópolis, SC, Brazil
[†]Université de Nice Sophia-Antipolis, CNRS, Nice, France
[‡]University of Toulouse, IRIT-ENSEEIHT, CNRS, Toulouse, France

ABSTRACT

The profusion of spectral bands generated by the acquisition process of hyperspectral images generally leads to high computational costs. Such difficulties arise in particular with nonlinear unmixing methods, which are naturally more complex than linear ones. This complexity, associated with the high redundancy of information within the complete set of bands, make the search of band selection algorithms relevant. With this work, we propose a band selection strategy in reproducing kernel Hilbert spaces that allows to drastically reduce the processing time required by nonlinear unmixing techniques. Simulation results show a complexity reduction of two orders of magnitude without compromising unmixing performance.

***Index Terms*—** Hyperspectral data, nonlinear unmixing, band selection, kernel methods

1. INTRODUCTION

Hyperspectral images (HI) consist of hundreds or even thousands of contiguous spectral samples ranging from the visible to the near infrared portions of the light spectrum. HIs trade spacial for spectral resolution [1], a consequence of which is that observed HI pixels can be mixtures of the spectral signatures of the materials present in the scene. The spectral signature of each material is usually available as a vector whose elements are proportional to the reflectances associated with that material at each frequency band. Such vectors are typically called endmembers due to their geometrical interpretation in the linear mixing case. The analysis of HIs frequently aims at unmixing the spectral information present at each pixel in the image, a study known as hyperspectral unmixing (HU). Unmixing problems can be cast as supervised or unsupervised learning problems depending on whether the endmembers are known or not.

Several models have been proposed in the literature to describe the mixing process of spectral information in HIs. The simplest one is the linear mixing model (LMM), in which each observed pixel spectrum is modeled as a linear combination of the endmembers [2]. Though the LMM simplifies the mathematical treatment of the unmixing problem, it has been recognized that significant nonlinear effects are often present in the spectral mixing occurring in real images [3]. Nonlinear mixing occurs, for instance, due to multiple interactions among light and different materials during the acquisition process. Recognition of such nonlinear effects has led to several nonlinear mixing models for HI processing. These models include adding cross-terms of different endmembers to the LMM [4], using bilinear mixture models [5], post-nonlinear mixing models [6,7], and kernel-based models [8]. Most nonlinear unmixing techniques are based on Bayesian inference [5], on using manifold learning techniques and geodesic distances [9–11], or processing data in reproducing kernel Hilbert spaces (RKHS) [8, 12].

One of the problems in practical implementation of unmixing algorithms is the profusion of spectral bands generated in the acquisition process, which leads to high computational costs. This is especially true for nonlinear unmixing algorithms, which are naturally more complex than linear techniques. Such inherent complexity, associated with the high redundancy within the complete set of bands, make the search of band selection techniques natural and relevant [13]. Several band selection algorithms have been proposed for linearly mixed HIs, which generally requires solving an optimization problem [14]. Nonlinear unmixing presents an even more challenging problem for band selection.

In this paper, we propose a technique for band selection in nonlinear supervised HI unmixing problems. This method applies a kernel k-means algorithm to identify nonlinearly separable clusters of spectral bands in the corresponding RKHS. The solution is evaluated using the SK-Hype nonlinear unmixing algorithm [8] on the selected bands. Simulation results indicate a complexity reduction of two orders of magnitude without compromising unmixing performance.

The paper is organized as follows. First, we state the unmixing problem and introduce usual nonlinear mixing models. Then, we describe a nonlinear unmixing algorithm operating in RKHS. Next, we introduce our band selection algorithm based on kernel k-means. We provide promising simulation results to illustrate the performance of our approach. Finally, we present some concluding remarks.

2. MIXTURE MODELS

Each observed HI pixel can be written as a function of the endmembers and a noise component that covers the unknown or unmodeled factors in the system:

$$\boldsymbol{r} = \boldsymbol{\Psi}(\boldsymbol{M}) + \boldsymbol{n} \tag{1}$$

with $\boldsymbol{r} = [r_1, \ldots, r_L]^\top$ the observed pixel vector over L spectral bands, $\boldsymbol{M} = [\boldsymbol{m}_1, \ldots, \boldsymbol{m}_R]$ the $L \times R$ matrix of endmembers $\boldsymbol{m}_i$, $\boldsymbol{n}$ a noise vector, and $\boldsymbol{\Psi}$ an unknown function. Several mixing models have been presented in the literature, depending on the linearity or nonlinearity of $\boldsymbol{\Psi}$, type of mixture, scale, and other properties [3].

2.1. The Linear Mixing Model (LMM)

The LMM assumes that $\boldsymbol{\Psi}$ is a convex combination of the endmembers, that is,

$$\begin{aligned} \boldsymbol{r} &= \boldsymbol{M}\boldsymbol{\alpha} + \boldsymbol{n} \\ \text{subject to } \mathbf{1}^\top \boldsymbol{\alpha} &= 1, \text{ and } \boldsymbol{\alpha} \succeq 0 \end{aligned} \tag{2}$$

where the vector $\boldsymbol{\alpha} = [\alpha_1, \ldots, \alpha_R]^\top$ contains the proportions (or abundances) of each endmember in $\boldsymbol{M}$ and, therefore, cannot be negative and should sum to one. The observation r_ℓ at the ℓ-th spectral band in (2) can be written as

$$r_\ell = \boldsymbol{m}_{\lambda_\ell}^\top \boldsymbol{\alpha} + n_\ell \tag{3}$$

where $\boldsymbol{m}_{\lambda_\ell}^\top$ is the ℓ-th row of $\boldsymbol{M}$. For $n_\ell = 0$ (noiseless case), the sum-to-one and positivity constraints on the abundances in (2) confine the data (pixels) to a simplex, for which the vertices are the endmembers.

2.2. Nonlinear Mixing Models

Different parametric nonlinear models have been proposed in the literature (see, e.g., [3] and references therein). Here we review two popular models that will be used later on to generate synthetic data for evaluation purposes.

The *generalized bilinear model* (GBM) [5] is given by

$$\begin{aligned} \boldsymbol{r} &= \boldsymbol{M}\boldsymbol{\alpha} + \sum_{i=1}^{R-1} \sum_{j=i+1}^{R} \delta_{ij} \alpha_i \alpha_j \boldsymbol{m}_i \odot \boldsymbol{m}_j + \boldsymbol{n} \\ \text{subject to } \mathbf{1}^\top \boldsymbol{\alpha} &= 1, \text{ and } \boldsymbol{\alpha} \succeq 0 \end{aligned} \tag{4}$$

where the parameters $\delta_{ij} \in [0, 1]$ govern the amount of nonlinear contribution, and $\odot$ denotes the Hadamard product. For simplicity, we consider in the following a simplified version of this model, with a single parameter δ to control the nonlinear contribution such that $\delta_{ij} = \delta$ for all (i, j).

The *post nonlinear mixing model* (PNMM) [15] is

$$\boldsymbol{r} = \boldsymbol{g}(\boldsymbol{M}\boldsymbol{\alpha}) + \boldsymbol{n} \tag{5}$$

where $\boldsymbol{g}(\cdot)$ is a nonlinear function applied to the result of a linear mixing. The PNMM can represent a wide range of nonlinear mixing models, depending on $\boldsymbol{g}(\cdot)$. For instance, the PNMM considered in [8] is defined as

$$\boldsymbol{r} = (\boldsymbol{M}\boldsymbol{\alpha})^\xi + \boldsymbol{n} \tag{6}$$

where $(\boldsymbol{v})^\xi$ denotes the exponential value ξ applied to each entry of the input vector $\boldsymbol{v}$. The PNMM was explored in other works using different forms for $\boldsymbol{g}(\cdot)$ applied to HU [6, 16, 17].

The GBM (4) and the PNMM (6) nonlinear mixing models consider mainly the scattering effect in light interactions with endmembers. Other models consider different nonlinear effects, depending on the characteristics of the application, such as the type of mixture, the types of materials, the geometry of the reflection surface, and the constraints over the abundances [18–22]. More importantly, these informations are usually missing in HU problems. Therefore, it makes sense to consider nonlinear unmixing strategies that do not make strong assumptions about the nonlinearity in the mixture.

3. NONLINEAR UNMIXING STRATEGY

This section reviews the SK-Hype algorithm[1] for nonlinear unmixing of HIs [8]. It considers the mixing model consisting of a linear trend parametrized by the abundance vector $\boldsymbol{\alpha}$ and a nonlinear residual component ψ. This model is given by

$$r_\ell = u\, \boldsymbol{\alpha}^\top \boldsymbol{m}_{\lambda_\ell} + (1-u)\, \psi(\boldsymbol{m}_{\lambda_\ell}) + n_\ell \tag{7}$$

where $u \in [0, 1]$ controls the amount of linear contribution to the model and $\psi(\cdot)$ is an unknown function in an RKHS $\mathcal{H}$.

Kernel methods are efficient machine learning techniques that were initially introduced for solving nonlinear classification and regression problems. They consist of linear algorithms operating in high dimensional feature spaces into which the data have been mapped using kernel functions [23]. These approaches are based on the framework of reproducing kernels which states that, for any positive kernel $\kappa(\boldsymbol{m}_{\lambda_i}, \boldsymbol{m}_{\lambda_j})$, there exists a Hilbert space $\mathcal{H}$ with inner product $\langle \cdot\,, \cdot \rangle_{\mathcal{H}}$ and a mapping

$$\boldsymbol{\Phi}: \quad \mathbb{R}^L \longrightarrow H \tag{8}$$

$$\boldsymbol{m}_{\lambda_i} \longmapsto \kappa(\cdot, \boldsymbol{m}_{\lambda_i}) \tag{9}$$

such that $\kappa(\boldsymbol{m}_{\lambda_i}, \boldsymbol{m}_{\lambda_j}) = \langle \boldsymbol{\Phi}(\boldsymbol{m}_{\lambda_\mathbf{i}}), \boldsymbol{\Phi}(\boldsymbol{m}_{\lambda_\mathbf{j}}) \rangle_{\mathcal{H}}$. This last property allows to implicitly compute inner products in $\mathcal{H}$ by evaluating a real function, $\kappa(\boldsymbol{m}_{\lambda_i}, \boldsymbol{m}_{\lambda_j})$, in the input space. Other useful properties are $\psi(\boldsymbol{m}_{\lambda_j}) = \langle \psi, \kappa(\cdot, \boldsymbol{m}_{\lambda_j}) \rangle_{\mathcal{H}}$ for all ψ in $\mathcal{H}$, and the reproducing property $\kappa(\boldsymbol{m}_{\lambda_i}, \boldsymbol{m}_{\lambda_j}) = \langle \kappa(\cdot, \boldsymbol{m}_{\lambda_i}), \kappa(\cdot, \boldsymbol{m}_{\lambda_j}) \rangle_{\mathcal{H}}$.

The optimization problem proposed in [8] for estimating the unknown variables $\boldsymbol{\alpha}$, $\psi(\cdot)$ and u in (7) is

$$\begin{aligned} \min_{\boldsymbol{\alpha}, \psi, u} \; & \frac{1}{2} \left(\frac{1}{u} \|\boldsymbol{\alpha}\|^2 + \frac{1}{1-u} \|\psi\|_{\mathcal{H}}^2 \right) \\ & + \frac{1}{2\mu} \sum_{\ell=1}^{L} (r_\ell - \boldsymbol{\alpha}^\top \boldsymbol{m}_{\lambda_\ell} - \psi(\boldsymbol{m}_{\lambda_\ell}))^2 \end{aligned} \tag{10}$$

subject to $\boldsymbol{\alpha} \succeq \mathbf{0}$ and $\mathbf{1}^\top \boldsymbol{\alpha} = 1$. This convex problem can be solved using a two stage alternating iterative optimization procedure with respect to $(\boldsymbol{\alpha}, \psi)$ and u.

3.1. Solving with respect to $(\boldsymbol{\alpha}, \psi)$

Introducing the Lagrange multipliers $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, the dual problem of (10) for fixed u is given by [8]

$$\begin{aligned} & \max_{\boldsymbol{\beta}, \boldsymbol{\gamma}} \; G(u, \boldsymbol{\beta}, \boldsymbol{\gamma}) = \\ & \quad -\frac{1}{2} \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{pmatrix}^\top \left(\begin{array}{c|c} \boldsymbol{K}_u + \mu \boldsymbol{I} & u\boldsymbol{M} \\ \hline u\boldsymbol{M}^\top & u\boldsymbol{I} \end{array} \right) \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{pmatrix} \\ & \quad + \begin{pmatrix} \boldsymbol{r} \\ \mathbf{0} \end{pmatrix}^\top \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{pmatrix} \\ & \quad \text{subject to} \quad \boldsymbol{\gamma} \succeq \mathbf{0} \end{aligned} \tag{11}$$

with $\boldsymbol{K}_u = u\boldsymbol{M}\boldsymbol{M}^\top + (1-u)\boldsymbol{K}$, and $\boldsymbol{K}$ the Gram matrix such that $\boldsymbol{K}_{ij} = \kappa(\boldsymbol{m}_{\lambda_i}, \boldsymbol{m}_{\lambda_j})$. This leads to the following solution of primal problem:

$$\begin{cases} \boldsymbol{\alpha}^\star = \frac{\boldsymbol{M}^\top \boldsymbol{\beta}^\star + \boldsymbol{\gamma}^\star}{\mathbf{1}^\top (\boldsymbol{M}^\top \boldsymbol{\beta}^\star + \boldsymbol{\gamma}^\star)} \\ \psi^\star = (1-u) \sum_{\ell=1}^{L} \beta_\ell^\star \, \kappa(\cdot, \boldsymbol{m}_{\lambda_\ell}) \\ e_\ell^\star = \mu \, \beta_\ell^\star \end{cases} \tag{12}$$

where $\boldsymbol{\beta}^\star$ and $\boldsymbol{\gamma}^\star$ are the solutions of (11).

[1] Matlab code available at www.cedric-richard.fr

3.2. Solving with respect to u

Using the stationary conditions (12), the optimum value for u given $\boldsymbol{\beta}^\star$ and $\boldsymbol{\gamma}^\star$ can be computed each iteration as [24]

$$u^\star = \left(1 + (1 - u^\star_{-1})\sqrt{\frac{\boldsymbol{\beta}^{\star\top}\boldsymbol{K}\boldsymbol{\beta}^\star}{\|\boldsymbol{M}^\top\boldsymbol{\beta}^\star + \boldsymbol{\gamma}^\star\|^2}}\right) \tag{13}$$

where $u^\star_{-1}$ is the optimum $u^\star$ obtained at the previous iteration.

4. BAND SELECTION AND UNMIXING

Band selection is the strategy of choice for reducing the complexity of HI processing when the original spectral information of each pixel needs to be preserved [25]. Most existing band selection algorithms have been derived assuming linearly mixed HIs. To the best of our knowledge, the problem of band selection for a single nonlinearly mixed pixel and with the objective of reducing the complexity of the unmixing process is still largely untreated in the literature. As each pixel may be better characterized by a different set of bands, we propose a simple supervised band selection approach to be applied to individual pixels so that only those endmembers present in each pixel, that is, the matrix $\boldsymbol{M}$ for that pixel, are considered. The clustering technique proposed in the following is based on the *kernel k-means* (KKM) algorithm [26], and uses the fact that each band of the HI is a function of the elements in one row of matrix $\boldsymbol{M}$. The choice of the KKM algorithm is due to the practical assumption that we lack information about the nonlinearity associated to the endmember mixing in each HI.

4.1. Kernel k-means for band-selection

For band selection we consider each row of $\boldsymbol{M}$ as an element of a vector space, and search for a set of K disjoint clusters $\mathcal{C}_1, \ldots, \mathcal{C}_K$ in that space. Then, a unique wavelength is chosen using the KKM algorithm to represent each cluster. KKM is an extension of the standard k-means clustering algorithm that identifies nonlinearly separable clusters [26]. It maps the data implicitly to a RKHS $\mathcal{H}$ where it performs a conventional k-means algorithm. Since the computation of the centroids in space $\mathcal{H}$ is intractable, KKM algorithms use the reproducing property [27] to directly compute squared distances between points in a cluster $\mathcal{C}_k$. Therefore, given a cluster $\mathcal{C}_k$ enclosing points $\{\kappa(\cdot, \boldsymbol{m}_{\lambda_\ell})\}_{\ell\in\mathcal{C}_k}$, we can write the centroid μ_k as

$$\mu_k = \frac{1}{N_k}\sum_{i\in\mathcal{C}_k}\kappa(\cdot, \boldsymbol{m}_{\lambda_i}) \tag{14}$$

where N_k is the number of points in $\mathcal{C}_k$. The squared distances to the centroid of $\mathcal{C}_k$ are then computed as

$$\begin{aligned}\|\kappa(\cdot, \boldsymbol{m}_{\lambda_\ell}) - \mu_k\|^2_{\mathcal{H}} = \kappa(\boldsymbol{m}_{\lambda_\ell}, \boldsymbol{m}_{\lambda_\ell}) \\ - \frac{1}{N_k}\sum_{i\in\mathcal{C}_k}\kappa(\boldsymbol{m}_{\lambda_\ell}, \boldsymbol{m}_{\lambda_i}) \\ + \frac{1}{N_k^2}\sum_{i\in\mathcal{C}_k}\sum_{j\in C_k}\kappa(\boldsymbol{m}_{\lambda_i}, \boldsymbol{m}_{\lambda_j})\end{aligned} \tag{15}$$

and the cluster error is defined as

$$E(\mu_1, \ldots, \mu_K) = \sum_{k=1}^{K}\sum_{\ell\in\mathcal{C}_k}\|\kappa(\cdot, \boldsymbol{m}_{\lambda_\ell}) - \mu_k\|^2_{\mathcal{H}}. \tag{16}$$

To preserve the original HI band information, we propose to represent cluster $\mathcal{C}_k$ by the band ℓ_k corresponding to the closest point to the centroid μ_k:

$$\ell_k = \arg\min_{\ell\in\mathcal{C}_k}\|\kappa(\cdot, \boldsymbol{m}_{\lambda_\ell}) - \mu_k\|^2_{\mathcal{H}}. \tag{17}$$

The *global kernel k-means* (GKKM) algorithm uses the principles above for incremental clustering [26]. GKKM avoids poor local minima and produces near-optimal solutions that are robust to cluster initialization. A fast GKKM (FGKKM) version that performs a unique KKM run and greatly reduces the complexity of the algorithm can also be used. Algorithm 1 details the application of FGKKM algorithm using (17) for band selection. We refer to Algorithm 1 as KKMBS for short.

Algorithm 1: FGKKM Band Selection (KKMBS)

Input : The $L \times R$ matrix $\boldsymbol{M}$ and the desired number of bands N_b.
Output: Selected band indices $\boldsymbol{\ell}$.

```
% Find clusters
[C_1, ..., C_{N_b}] = FGKKM(M, N_b);
% Find the lines of M closer to the centroids in C_1, ..., C_{N_b}
for k = 1 to N_b do
    ℓ_k = arg min_{ℓ∈C_k} ||κ(·, m_{λ_ℓ}) − μ_k||²_H;
end
return ℓ
```

4.2. Unmixing of reduced data

For a given N-pixel HI region with a known endmember matrix $\boldsymbol{M}$, define $\boldsymbol{R} = [\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N]$ as the matrix of observations. The unmixing problem then reduces to an inversion step for which we solve the optimization problem described in Section 3 to obtain the estimated abundances given by $\boldsymbol{A} = [\boldsymbol{\alpha}^\star_1, \ldots, \boldsymbol{\alpha}^\star_N]$. We propose to replace matrices $\boldsymbol{R}$ and $\boldsymbol{M}$ by their reduced versions $\boldsymbol{R}_r$ and $\boldsymbol{M}_r$ containing only N_b selected bands obtained with Algorithm 1. Then, the SK-Hype algorithm can be used to perform data unmixing and to estimate the abundance matrix as shown in Figure 1.

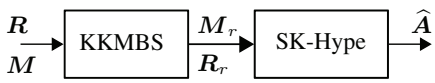


Fig. 1: Band selection and unmixing.

5. SIMULATION RESULTS

This section presents simulation results with synthetic data to illustrate the potential of the band selection technique. The HI was built using measured spectra from eight materials: alunite, calcite, epidote, kaolinite, buddingtonite, almandine, jarosite and lepidolite. Their spectra were selected from the spectral library of the ENVI software and consisted of 420 contiguous bands, covering wavelengths from 0.3951 to 2.56 micrometers. We compared the results obtained with 2 nonlinear mixing models using 5 unmixing algorithm implementations and 2 performance measures: the *root mean square error* (RMSE) and the *relative execution time* (RET).

$$\text{RMSE} = \|\boldsymbol{A} - \widehat{\boldsymbol{A}}\|_F/(NR) \tag{18}$$

Table 2: RMSE and RET for different nonlinearities in each band.

	PNMM			
	5 Endmembers		8 Endmembers	
Alg. (N_b)	RMSE	RET (BS+HU)	RMSE	RET (BS+HU)
FCLS	0.1966	1	0.1521	1
SK-Hype	**0.1127**	3744.9	**0.0765**	3672.2
SK-Hype(10)	0.1131	21.7 (24.3)	0.0827	21.5 (24.4)
SK-Hype(100)	0.1140	153.4 (184.2)	0.0771	140.3 (170.8)
SK-Hype(300)	0.1129	1764.3 (1907.5)	**0.0765**	1601.0 (1743.7)
	GBM			
	5 Endmembers		8 Endmembers	
FCLS	0.0527	1	0.0390	1
SK-Hype	0.1090	3432.5	0.0745	3825.1
SK-Hype(10)	**0.1053**	20.8 (23.6)	**0.0730**	20.9 (23.0)
SK-Hype(100)	0.1105	142.2 (171.5)	0.0748	142.9 (173.8)
SK-Hype(300)	0.1093	1612.8 (1735.2)	0.0745	1637.8 (1778.1)

where $\|\boldsymbol{X}\|_F$ denotes the Frobenius norm of matrix $\boldsymbol{X}$. The RET for the p-th algorithm was determined as $\text{RET}_p = t_p/t_{\text{FCLS}}$, where t_p is the execution time for the p-th algorithm and t_{FCLS} is the execution time for the Fully Constrained Least Squares (FCLS) algorithm.

The data were nonlinearly mixed using the PNMM and GBM models presented in Section 2.2. For PNMM, we used $\xi = 0.7$, and for GBM, we considered $\delta = 1$. We generated 6 databases using the two models and different number of endmembers (5 and 8). Each database had 2000 pixels generated using random abundances drawn uniformly from the simplex defined by (2). The additive noise power σ_n^2 was chosen to produce a 21dB SNR. For each dataset we unmixed the data using FCLS, SK-Hype without band selection and the proposed algorithm for $N_b = [10, 100, 300]$. For both SK-Hype and KKMBS, we used the Gaussian kernel [28] with bandwidth $\sigma_k^2 = 0.3$. Table 1 presents the RMSEs and the RETs obtained for each simulation. The RETs within parentheses include the time for both band separation and unmixing as indicated. These results show comparable RMSEs for all nonlinear unmixing results and very significant improvements in RET when using the proposed band selection approach (about 145 times for $N_b = 10$).

Table 2 shows the results when the nonlinearity parameters (δ for GBM and ξ for PNMM) were set differently for each band. Parameter δ (ξ) varied in the interval $[0.5, 1]$ ([0.5, 0.9]) with steps of 0.05 (0.04), changing at every 42 bands. These results corroborate the results shown in Table 1. We have also compared the results of KKMBS with those obtained after randomly selecting the bands. Figure 2 shows the histograms of the RMSE when selecting 10 and 100 bands. It is clear that the KKMBS leads to significantly better results when few bands are selected.

Table 1: RMSE and RET for SNR = 21dB and random abundances.

	PNMM			
	5 endmembers		8 endmembers	
Alg. (N_b)	RMSE	RET (BS+HU)	RMSE	RET (BS+HU)
FCLS	0.1893	1	0.1243	1
SK-Hype	0.1136	2690.6	**0.0762**	3028.8
SK-Hype(10)	**0.1114**	16.0 (18.1)	0.0775	16.8 (18.7)
SK-Hype(100)	0.1150	107.7 (129.1)	0.0766	118.6 (144.8)
SK-Hype(300)	0.1139	1226.7 (1327.1)	0.0763	1331.5 (1452.7)
	GBM			
FCLS	0.2419	1	0.1836	1
SK-Hype	0.1080	3320.7	0.0738	3072.6
SK-Hype(10)	**0.1037**	24.1 (26.9)	**0.0712**	18.7 (21.1)
SK-Hype(100)	0.1095	157.0 (194.6)	0.0741	119.7 (148.1)
SK-Hype(300)	0.1083	1548.6 (1676.0)	0.0738	1475.2 (1597.7)

6. CONCLUSIONS

This work proposed a supervised band selection strategy to reduce the complexity of nonlinear hyperspectral data unmixing without compromising the accuracy of abundance estimation. Significant reduction in processing time was achieved in all cases tested. These results suggest the possibility of important complexity reduction for nonlinear HI processing algorithms without performance loss. It is conjectured that presented performance can be further improved by removing redundancy in the data through a more specialized band selection algorithm. This is the topic of a work in progress.

REFERENCES

[1] D. Landgrebe, "Hyperspectral image data analysis," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 17–28, 2002.

[2] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geoscience and Remote Sensing Magazine*, vol. 1, no. 2, pp. 6–36, 2013.

[3] N. Dobigeon, J.-Y. Tourneret, C. Richard, J. C. M. Bermudez, S. McLaughlin, and A. O. Hero, "Nonlinear unmixing of hyperspectral images: Models and algorithms," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 82–94, 2014.

[4] J. Boardman, "Automatic spectral unmixing of AVIRIS data using convex geometry concepts," in *Proc. AVIRIS workshop*, 1993, vol. 1, pp. 11–14.

[5] 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.

[6] Y. Altmann, A. Halimi, N. Dobigeon, and J.-Y. Tourneret, "Supervised nonlinear spectral unmixing using a polynomial post nonlinear model for hyperspectral imagery," in *Proc. IEEE ICASSP*, 2011, pp. 1009–1012.

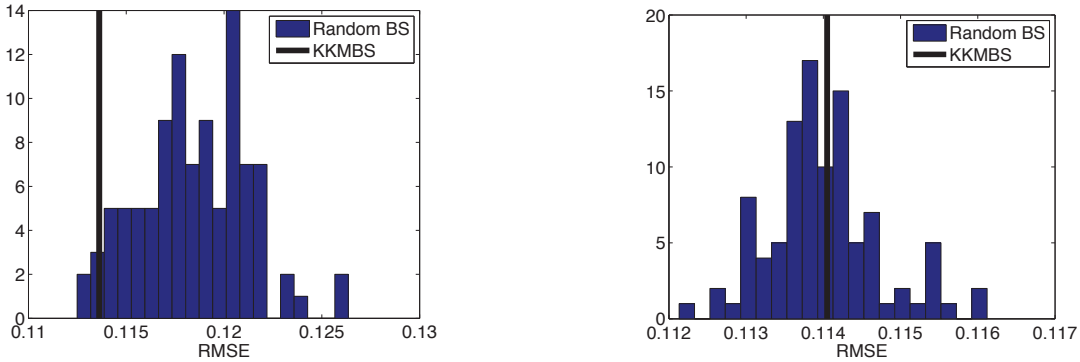
[7] J. Chen, C. Richard, and P. Honeine, "Estimating abundance fractions of materials in hyperspectral images by fitting a post-nonlinear mixing model," in *Proc. IEEE WHISPERS*, 2013, pp. 1–4.

[8] J. Chen, C. Richard, and P. Honeine, "Nonlinear unmixing of hyperspectral data based on a linear-mixture/nonlinear-fluctuation model," *IEEE Transactions on Signal Processing*, vol. 61, pp. 480–492, 2013.

[9] H. N. Nguyen, C. Richard, P. Honeine, and C. Theys, "Hyperspectral image unmixing using manifold learning methods derivations and comparative tests," in *Proc. IEEE IGARSS*, 2012, pp. 3086–3089.

[10] P. Honeine and C. Richard, "Solving the pre-image problem in kernel machines: A direct method," in *Proc. IEEE MLSP*, 2009, pp. 1–6.

[11] P. Honeine and C. Richard, "Preimage problem in kernel-based machine learning," *IEEE Signal Processing Magazine*, vol. 28, no. 2, pp. 77–88, 2011.



(a) 10 bands.

(b) 100 bands.

Fig. 2: RMSE for 10 and 100 bands.

[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] C.-I. Chang, *Hyperspectral data processing - Algorithm design and analysis*, Wiley, 2014.

[14] C.-I. Chang and K.-H. Liu, "Progressive band selection of spectral unmixing for hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 4, pp. 2002–2017, 2014.

[15] C. Jutten and J. Karhunen, "Advances in nonlinear blind source separation," in *Proc. ICA*, 2003, pp. 245–256.

[16] Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinearity detection in hyperspectral images using a polynomial post-nonlinear mixing model.," *IEEE Transactions on Image Processing*, vol. 22, no. 4, pp. 1267–1276, 2013.

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

[18] 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.

[19] B. Hapke, *Theory of Reflectance and Emittance Spectroscopy*, Cambridge University Press, 1993.

[20] C. C. Borel and S. A. W. Gerstl, "Nonlinear spectral mixing models for vegetative and soil surfaces," *Remote Sensing of Environment*, vol. 47, no. 3, pp. 403–416, 1994.

[21] B. Somers, K. Cools, S. Delalieux, J. Stuckens, D. Van der Zande, W. W. Verstraeten, and P. Coppin, "Nonlinear hyperspectral mixture analysis for tree cover estimates in orchards," *Remote Sensing of Environment*, vol. 113, no. 6, pp. 1183–1193, 2009.

[22] T. W. Ray and B. C. Murray, "Nonlinear spectral mixing in desert vegetation," *Remote Sensing of Environment*, vol. 55, no. 1, pp. 59–64, 1996.

[23] V. N. Vapnik, *The nature of statistical learning theory*, Springer, New York, NY, 1995.

[24] J. Chen, C. Richard, A. Ferrari, and P. Honeine, "Nonlinear unmixing of hyperspectral data with partially linear least-squares support vector regression," in *Proc. IEEE ICASSP*, 2013, pp. 2174–2178.

[25] Chein-I Chang and Su Wang, "Constrained band selection for hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 6, pp. 1575–1585, 2006.

[26] G. F. Tzortzis and A. C. Likas, "The global kernel k-means algorithm for clustering in feature space.," *IEEE Transactions on Neural Networks*, vol. 20, pp. 1181–1194, 2009.

[27] N. Aronszajn, "Theory of reproducing kernels," *Transactions of the American Mathematical Society*, vol. 68, 1950.

[28] B. Schlkopf, J. C. Burges, and A. J. Smola, *Advances in kernel methods*, MIT Press, Cambridge, MA, 1999.