

# A novel kernel-based nonlinear unmixing scheme of hyperspectral images

Jie Chen<sup>(1,2)</sup>, Cédric Richard<sup>(1)</sup>, Paul Honeine<sup>(2)</sup>

<sup>(1)</sup>Laboratoire Fizeau, Université de Nice Sophia-Antipolis, France

<sup>(2)</sup>Institut Charles Delaunay, Université de Technologie de Troyes, France

E-mail: jie.chen@utt.fr, cedric.richard@unice.fr, paul.honeine@utt.fr

**Abstract**—In hyperspectral images, pixels are mixtures of spectral components associated to pure materials. Although the linear mixture model is the most studied case, nonlinear models have been taken into consideration to overcome some limitations of the linear model. In this paper, nonlinear hyperspectral unmixing problem is studied through kernel-based learning theory. Endmember components at each band are mapped implicitly in a high feature space, in order to address the nonlinear interaction of photons. Experiment results with both synthetic and real images illustrate the effectiveness of the proposed scheme.

## I. INTRODUCTION

By providing images with a large number of spectral bands, hyperspectral imagery has been applied with success to a large number of areas, as it allows to obtain more accurate and detailed information than other types of remotely sensed techniques. It has been applied to a number of areas, including environment, land use and agricultural monitoring.

In hyperspectral imagery, each pixel is a mixture of spectral components associated to a number of pure materials, often present in the scene [1]. The unmixing problem consists of identifying the pure materials, the so-called endmembers, and estimating their contributions (or abundances). The linear mixture model has been widely studied due to its simple physical interpretation and its analytically tractable solution. Although the linear mixture model has several advantages, there are many situations, involving multiple light scattering effects, which requires nonlinear mixture analysis. Researchers have proposed several unmixing methods to address nonlinear mixtures. In [2], [3], the authors designed neural networks for unmixing nonlinearly mixed pixels. However, these methods require their networks to be trained by training data and the quality of training data may affect the performance notably. In [4], a nonlinear unmixing algorithm based on Bayesian inference was proposed. This method has a high computational complexity and it is dedicated to the bilinear model. In [5], the researchers extended the endmember matrix with cross terms of pure signatures to model the light scattering on different materials. However it remains difficult to identify the number and the elements involved in the cross terms.

In the machine learning community, kernel-based algorithms in Hilbert space have been a topic of considerable

interest and have gained notable success to deal with nonlinear problems in various applications. Kernel methods have been studied for classification and detection problems for hyperspectral images [6], [7]. Kernel-based nonlinear unmixing approaches have also been investigated [8], [9], [10], by deriving nonlinear algorithms based on linear ones thanks to the kernel trick. Kernels are however applied to the spectral signature of each pure component, as a whole, and independently of interactions between materials, thus operating as nonlinear distortion functions.

In this paper, we formulate a new paradigm where each kernel is applied at a given wavelength, thus combining the contributions of each endmember via a nonlinear mixing rule. This scheme thus has a clear physical interpretation. The abundances can be determined by solving a constrained kernel regression problem. Computer simulations will show the effectiveness of the proposed approach.

## II. NONLINEAR MIXTURE MODELS

Suppose that  $L$  is the number of spectral bands. Let  $\mathbf{r} = [r_1, r_2, \dots, r_L]^T$  be an  $L \times 1$  observed column pixel which is a mixture of  $R$  endmembers corrupted by an additive noise. Assume that  $\mathbf{M}$  is the  $L \times R$  target endmember matrix denoted by  $[\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_R]$  with each column  $\mathbf{m}_i$  a material signature. For the sake of convenience in the following we also denote the  $\ell^{th}$  row of  $\mathbf{M}$ , namely, the signatures of the  $\ell^{th}$  wavelength band, by  $\mathbf{m}_{\lambda_\ell}^T$ . Let  $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_R]^T$  be a  $R \times 1$  abundance column vector associated with the pixel  $\mathbf{r}$ .

There are multiple situations where light scattering effects can be observed. For instance, multiple scattering effects can occur on complex vegetated surfaces where it is assumed that incident solar radiation is scattered by the scene through multiple bounces involving several endmembers. Bilinear mixture models account for the presence of multiple photon bounces by introducing additional interaction terms in the linear model. The generalized bilinear model studied in [4] assumes that the observed pixel can be expressed as

$$\mathbf{r} = \mathbf{M}\boldsymbol{\alpha} + \sum_{i=1}^{R-1} \sum_{j=i+1}^R \gamma_{ij} \alpha_i \alpha_j \mathbf{m}_i \otimes \mathbf{m}_j + \mathbf{n} \quad (1)$$

with  $\mathbf{n}$  the noise vector and  $\otimes$  the Hadamard product. The following constraints are considered for the parameters

$$\alpha_i \geq 0, \forall i \in i, \dots, R \quad \text{and} \quad \sum_{i=1}^R \alpha_i = 1$$

and  $0 \leq \gamma_{ij} \leq 1$ .

If the components of interest are in an intimate association, light interacts with more than one component as it is multiply scattered. Hapke developed a bidirectional reflectance model based on the fundamentals principles of radiative transfer theory [11]. This model is usually referred to as the intimate mixture model. The reflectance at a given wavelength is approximately given by

$$\mathbf{r} \approx \frac{\mathbf{w}}{4(\mu_0 + \mu_s)} H(\mathbf{w}, \mu_0) H(\mathbf{w}, \mu_s) \quad (2)$$

where  $\mathbf{w}$  is the average single-scattering albedo,  $\mu_0$  is the cosine of the angle of emergence,  $\mu_s$  is the cosine angle of incidence, and  $H$  is Chandrasekhar's function for isotropic scattering. The average single-scattering albedo of a homogeneous surface composed by large spherical particles is

$$\mathbf{w} = \sum_{i=1}^R \mathbf{w}_i \frac{G_i / (\rho_i D_i)}{\sum_{j=1}^R G_j / (\rho_j D_j)}$$

where  $G_i$  is the mass fraction for material  $i$ ,  $D_i$  is the diameter of the particles of material  $i$ ,  $\rho_i$  is the solid density of the particles of material  $i$ , and  $\mathbf{w}_i$  is the single-scattering albedo of material  $i$ . Their abundances can be estimated from the relative geometric cross section  $\alpha_i = \frac{M_i / (\rho_i D_i)}{\sum_j M_j / (\rho_j D_j)}$ . Note that the measures must be non-negative and sum-to-one similarly to the constraints in linear mixing model.

### III. KERNEL BASED NONLINEAR ANALYSIS OF HYPERSPECTRAL IMAGES

#### A. Presentation of the method

Considering the linear mixing model where an observed pixel is a linear combination of the endmembers weighted by fractional abundances, each pixel is given by

$$\mathbf{r} = \mathbf{M} \boldsymbol{\alpha} + \mathbf{n} \quad (3)$$

We estimate the abundance vector by minimizing the quadratic cost function

$$\begin{aligned} J(\boldsymbol{\alpha}) &= \|\mathbf{r} - \mathbf{M} \boldsymbol{\alpha}\|^2 \\ &= \sum_{\ell=1}^L \|r_\ell - \boldsymbol{\alpha}^\top \mathbf{m}_{\lambda_\ell}\|^2 \end{aligned} \quad (4)$$

under the non-negativity and sum-to-one constraints. Now consider a general mixing mechanism specified by

$$\mathbf{r} = \mathbf{F}(\boldsymbol{\alpha}, \mathbf{M}) + \mathbf{n} \quad (5)$$

where  $\mathbf{F}$  is an unknown nonlinear function that defines the interaction parameterized by  $\mathbf{M}$  and  $\boldsymbol{\alpha}$ . This requires us to consider a more general problem of the form

$$\psi^* = \arg \min_{\psi \in \mathcal{H}_\psi} \sum_{\ell=1}^L (r_\ell - \psi(\boldsymbol{\alpha}, \mathbf{m}_{\lambda_\ell}))^2 \quad (6)$$

where  $\psi$  is element of the functional space  $\mathcal{H}_\psi$ .

In order to extract the fraction of the pure endmembers, we specify the function  $\psi$  as follows

$$\psi(\boldsymbol{\alpha}, \mathbf{m}_{\lambda_\ell}) = \boldsymbol{\alpha}^\top \mathbf{m}_{\lambda_\ell} + \varphi(\mathbf{m}_{\lambda_\ell}) \quad (7)$$

where  $\varphi$  is an element of the functional space  $\mathcal{H}_\varphi$ . This means that the mixture mechanism consists of a linear mixing and some kind of nonlinear fluctuation defined by  $\varphi(\cdot)$ . Let  $\mathcal{H}_\varphi$  be a reproducing kernel Hilbert space (RKHS) of real-valued function on a compact  $\mathcal{M}$  and  $\langle \cdot, \cdot \rangle$  be the inner product in this space. We denote  $\kappa_\varphi : \mathcal{M} \times \mathcal{M} \rightarrow \mathbb{R}$  as the kernel function of this space, which satisfies the reproducing property  $\varphi(\mathbf{m}_\lambda) = \langle \varphi(\cdot), \kappa_\varphi(\cdot, \mathbf{m}_\lambda) \rangle$  for every function  $\varphi \in \mathcal{H}_\varphi$  and  $\mathbf{m}_\lambda \in \mathcal{M}$ . It can be found that the functional space  $\mathcal{H}_\psi$  of (7) is also a RKHS whose kernel function is

$$\kappa_\psi(\mathbf{m}_{\lambda_\ell}, \mathbf{m}_{\lambda_p}) = \mathbf{m}_{\lambda_\ell}^\top \mathbf{m}_{\lambda_p} + \kappa_\varphi(\mathbf{m}_{\lambda_\ell}, \mathbf{m}_{\lambda_p}) \quad (8)$$

By virtue of the Riesz representation theorem, the function  $\psi(\cdot)$  of  $\mathcal{H}_\psi$  minimizing (6) can be written as a kernel expansion in terms of available data, namely, the endmember components at each wavelength

$$\psi(\cdot) = \sum_{\ell=1}^L \beta_\ell \kappa_\psi(\cdot, \mathbf{m}_{\lambda_\ell}) \quad (9)$$

Replacing this expansion in (6), it can be shown the problem becomes linear with respect to the parameters to estimate

$$\boldsymbol{\beta}^* = \arg \min_{\boldsymbol{\beta} \in \mathbb{R}^L} \|\mathbf{r} - \mathbf{K}_\psi \boldsymbol{\beta}\|^2 \quad (10)$$

where  $\mathbf{K}_\psi$  is the Gram matrix whose  $(\ell, p)$ -th element is  $\kappa_\psi(\mathbf{m}_{\lambda_\ell}, \mathbf{m}_{\lambda_p})$ . With (8), it is can be found that

$$\mathbf{K}_\psi = \mathbf{M} \mathbf{M}^\top + \mathbf{K}_\varphi$$

where  $\mathbf{K}_\varphi$  is the Gram matrix corresponding to the nonlinear map  $\varphi$ . The pixel vector can be thus reconstructed by

$$\begin{aligned} \hat{\mathbf{r}} &= \mathbf{K}_\psi \boldsymbol{\beta}^* \\ &= \mathbf{M}(\mathbf{M}^\top \boldsymbol{\beta}^*) + \mathbf{K}_\varphi \boldsymbol{\beta}^* \end{aligned}$$

Comparing this expression with (7), it can be observed that the second term  $\mathbf{K}_\varphi \boldsymbol{\beta}^*$  approximates the nonlinear effect defined by  $\varphi(\cdot)$ . The first term, namely, the linear component, allows us to determine the abundance vector  $\boldsymbol{\alpha}^*$  with

$$\boldsymbol{\alpha}^* = \mathbf{M}^\top \boldsymbol{\beta}^* \quad (11)$$

As abundances are determined in this form, taking the non-negativity and sum-to-one constraints into consideration in this context yields

$$\begin{aligned} \mathbf{M}^\top \boldsymbol{\beta} &\geq \mathbf{0} \\ \mathbf{1}^\top (\mathbf{M}^\top \boldsymbol{\beta}) &= 1 \end{aligned}$$

where  $\mathbf{1}$  is an  $R \times 1$  vector of ones.

The whole optimization problem is now expressed as

$$\min_{\boldsymbol{\beta}} \|\mathbf{r} - \mathbf{K}_\psi \boldsymbol{\beta}\|^2 \quad (12)$$

$$\text{subject to } \mathbf{M}^\top \boldsymbol{\beta} \geq \mathbf{0} \quad (13)$$

$$\mathbf{1}^\top (\mathbf{M}^\top \boldsymbol{\beta}) = 1 \quad (14)$$

Once  $\boldsymbol{\beta}^*$  is calculated, the abundances can be estimated by (11). The problem defined by (12) to (14) is a quadratic programming problem with both equality and inequality constraints. Numerous candidate methods can be found to solve this problem and determine  $\boldsymbol{\beta}^*$ , as presented in [12], [13]. In this paper, we do not focus our interests on detailed solution approaches of this optimization problem.

### B. Kernel selection and interpretation

The kernel  $\kappa_\varphi$  is an important element that should be defined in this approach and it modelizes the nonlinearity in the mixture model. There is no limit in the choice of this kernel function provided that  $\kappa_\varphi$  is an inner product in the given Hilbert space. Note that it is not necessary to know explicitly the nonlinear map  $\varphi$ . However it is possible to choose a kernel which is closely associated to the application or to the inherent mixing model.

A widely used kernel in machine learning community is the Gaussian kernel  $\kappa(\mathbf{m}_{\lambda_\ell}, \mathbf{m}_{\lambda_p}) = \exp(-\|\mathbf{m}_{\lambda_\ell} - \mathbf{m}_{\lambda_p}\|^2/\sigma^2)$  with bandwidth  $\sigma$ . Some well-known kernels are homogeneous polynomial kernels  $\kappa(\mathbf{m}_{\lambda_\ell}, \mathbf{m}_{\lambda_p}) = \langle \mathbf{m}_{\lambda_\ell}, \mathbf{m}_{\lambda_p} \rangle^q$ . Polynomial kernels map the input data into a finite dimensional polynomial space. It can be noticed that the second-order polynomial kernel ( $q = 2$ ) is tightly related to the bilinear mixture model. Moreover, a higher order polynomial implies that multiple scatters between materials is considered.

In what follows, we shall illustrate the relation between selection of  $\kappa_\varphi$  and physical models using a simple example. Consider the case where there are three endmembers in the scene and the generalized bilinear mixture model is the mixing mechanism. A pixel is thus expressed by [4]

$$\begin{aligned} \mathbf{r} = \mathbf{M} \boldsymbol{\alpha} + \gamma_{12} \alpha_1 \alpha_2 \mathbf{m}_1 \otimes \mathbf{m}_2 + \gamma_{13} \alpha_1 \alpha_3 \mathbf{m}_1 \otimes \mathbf{m}_3 \\ + \gamma_{23} \alpha_2 \alpha_3 \mathbf{m}_2 \otimes \mathbf{m}_3 \end{aligned} \quad (15)$$

with constraints

$$\begin{aligned} \alpha_1, \alpha_2 \text{ and } \alpha_3 &\geq 0 \\ \sum_{i=1}^3 \alpha_i &= 1 \end{aligned}$$

where  $\gamma_{12}$ ,  $\gamma_{13}$  and  $\gamma_{23}$  are additional attenuation parameters, and  $\otimes$  denotes the Hadamard product. In this case suppose that the second-order homogeneous polynomial kernel is chosen for defining the functional space  $\mathcal{H}_\varphi$ , namely,  $\kappa_\varphi(\mathbf{m}_{\lambda_\ell}, \mathbf{m}_{\lambda_p}) = \langle \mathbf{m}_{\lambda_\ell}, \mathbf{m}_{\lambda_p} \rangle^2$ , which can be written in an inner product form

$$\begin{aligned} \kappa_\varphi(\mathbf{m}_{\lambda_\ell}, \mathbf{m}_{\lambda_p}) \\ = \langle (m_{\lambda_\ell,1}^2, m_{\lambda_\ell,2}^2, m_{\lambda_\ell,3}^2, \sqrt{2}m_{\lambda_\ell,1}m_{\lambda_\ell,2}, \\ \sqrt{2}m_{\lambda_\ell,1}m_{\lambda_\ell,3}, \sqrt{2}m_{\lambda_\ell,2}m_{\lambda_\ell,3}), \\ (m_{\lambda_p,1}^2, m_{\lambda_p,2}^2, m_{\lambda_p,3}^2, \sqrt{2}m_{\lambda_p,1}m_{\lambda_p,2}, \\ \sqrt{2}m_{\lambda_p,1}m_{\lambda_p,3}, \sqrt{2}m_{\lambda_p,2}m_{\lambda_p,3}) \rangle \end{aligned}$$

where  $m_{\lambda_\ell,i}$  is the  $i^{\text{th}}$  component of  $\mathbf{m}_{\lambda_\ell}$ . Thus the feature map from input space to the high dimensional space is then described by

$$\begin{aligned} \Phi : \mathbf{m}_{\lambda_\ell} &= (m_{\lambda_\ell,1}, m_{\lambda_\ell,2}, m_{\lambda_\ell,3}) \\ \mapsto \Phi(\mathbf{m}_{\lambda_\ell}) &= (m_{\lambda_\ell,1}^2, m_{\lambda_\ell,2}^2, m_{\lambda_\ell,3}^2, \sqrt{2}m_{\lambda_\ell,1}m_{\lambda_\ell,2}, \\ &\sqrt{2}m_{\lambda_\ell,1}m_{\lambda_\ell,3}, \sqrt{2}m_{\lambda_\ell,2}m_{\lambda_\ell,3}) \in \mathbb{R}^6 \end{aligned}$$

The basis of this functional space is thus closely related to the bilinear model defined by (15) except for the additional auto-terms and the coefficient  $\sqrt{2}$ .

From this example it is clear that the kernel  $\kappa_\varphi(\cdot, \cdot)$  represents the nonlinear interactions between different material signatures. However, thanks to the kernel, we do not need to map explicitly the original input data into the high dimensional space. This allows us to consider complex photon interaction mechanism just by modifying the kernel  $\kappa_\varphi$ . For example, a polynomial kernel with order  $q = 4$  means that more photon reflectance is modeled. Compared with the method in [5], which inserts products of pure material signatures as new endmembers, we do not need to extend the endmember matrix by adding such terms. The endmember matrix does not suffer from the problem of choosing cross terms and of large expansion in dimension.

## IV. EXPERIMENT RESULTS

### A. Experiments with synthetic images

Firstly, we study the proposed algorithm using synthetic data, generated by the bilinear mixture model and the intimate mixture model using three pure materials. We consider three mineral materials with spectral signatures from the USGC library. These spectra consist of 480 contiguous bands ranging from 0.35 to 2.5  $\mu\text{m}$  (See Fig. 1). Two hyperspectral images composed of 2500 pixels were generated respectively with these two models. All the fractions were subject to nonnegativity and constant sum-to-one constraints. Both images were corrupted by additive noise with SNR equal to 20 dB and 30 dB. The Gaussian kernel function was used for the proposed algorithm and preliminary experiments were conducted to

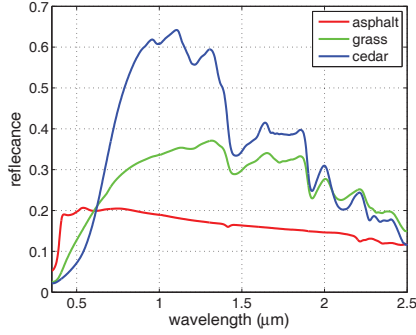


Fig. 1. Three endmembers used in the simulation

finally set the kernel bandwidth  $\sigma = 2$ . For both models, we compared the results of the proposed approach with the following three linear/nonlinear algorithms:

- 1) Fully constrained least square method in [14] (denoted by FCLS). This method was derived based on linear mixture model and provides the optimal solution in the linear sense.
- 2) Nonlinear method in [5] (denoted by NLCT). This method extends the endmember matrix with cross terms of pure signatures, to model the light scatter effects. In the experiment, we added all the second-order cross terms so that it corresponds to the bilinear model.
- 3) Kernel fully constrained least square method in [9] (denoted by KFCLS). This is a kernel method which replaces the inner products in the linear mixture problem by kernel functions.

The root mean square error

$$\text{RMSE} = \sqrt{\frac{1}{LR} \sum_{jk} \|\alpha_{ij} - \hat{\alpha}_{ij}\|^2}$$

was used to compare these algorithms. The simulation results of these algorithm are reported in Table I.

TABLE I  
RMSE COMPARISONS

	SNR = 30 dB		SNR = 20 dB	
	bilinear	intimate	bilinear	intimate
FCLS [14]	0.1218	0.1389	0.1256	0.1421
NLCT [5]	0.0456	0.1356	0.0696	0.1396
KFCLS [9]	0.1608	0.1732	0.1863	0.1800
<b>Proposed</b>	<b>0.0295</b>	<b>0.0711</b>	<b>0.0551</b>	<b>0.0860</b>

As the method FCLS was initially derived for linear mixture models, it has a relatively large RMSE compared with the other nonlinear methods. With added second order cross terms, NLCT reduces notably the RMSE for the bilinear mixture model. However, with the intimate model, it marginally improves the performance. This means it does not match this

model. KFCLS, which maps signatures of different material into a high dimensional space, performs even worth than the linear method. Compared with all these linear/nonlinear methods, the proposed approach clearly shows a notable improvement of the RMSE.

### B. Experiments with a real image

Let us now study these algorithms with a real image, the scene of the Cuprite mining district (Nevada, USA), captured by the airborne visible infrared imaging spectrometer (AVIRIS). A sub-image of size  $81 \times 101$  pixels was chosen here to evaluate the proposed algorithm. The averaged spectral angle between  $\mathbf{r}$  and reconstructed  $\hat{\mathbf{r}}$  was used

$$SA = \frac{1}{N} \sum_{n=1}^N \theta(\mathbf{r}_n, \hat{\mathbf{r}}_n)$$

where  $N$  is the total number of pixels and

$$\theta(\mathbf{r}, \hat{\mathbf{r}}) = \cos^{-1} \left( \frac{\langle \mathbf{r}, \hat{\mathbf{r}} \rangle}{\|\mathbf{r}\| \|\hat{\mathbf{r}}\|} \right)$$

In this experiment, we used VCA algorithm [15] to extract the endmembers, was  $R = 3$  and  $R = 5$  for illustration purpose. The results are shown in the Table II.

TABLE II  
SPECTRAL ANGLE COMPARISONS

	$R = 3$	$R = 5$
FCLS [14]	0.0594	0.0233
NLCT [5]	0.0591	0.0213
KFCLS [9]	-	-
<b>Proposed</b>	<b>0.0281</b>	<b>0.0183</b>

Note that the results for KFCLS can not be listed as there is no reconstruction model for this kernel method scheme. Compared with the others, the proposed algorithm has the lowest reconstruction errors for 3 and 5 endmembers. The estimated abundances of these four algorithms are illustrated in Fig. 2.

### REFERENCES

- [1] N. Keshava and J. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 44–57, 2002.
- [2] K. J. Guilfoyle, M. L. Althouse, and C.-I. Chang, "A quantitative and comparative analysis of linear and nonlinear spectral mixture models using radial basis function neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 39, no. 8, pp. 2314–2318, 2001.
- [3] J. Plaza, P. Martínez, R. Pérez, and A. Plaza, "Nonlinear neural network mixture models for fractional abundance estimation in AVIRIS hyperspectral images," in *Proceedings of XIII JPL Airborne Earth Science Workshop*, Pasadena, CA, 2004.
- [4] A. Halimi, Y. Altman, 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.
- [5] N. Raksuntorn and Q. Du, "Nonlinear Spectral Mixture Analysis for Hyperspectral Imagery in an Unknown Environment," *IEEE Geoscience and Remote Sensing Letters*, vol. 7, no. 99, pp. 836–840, 2010.

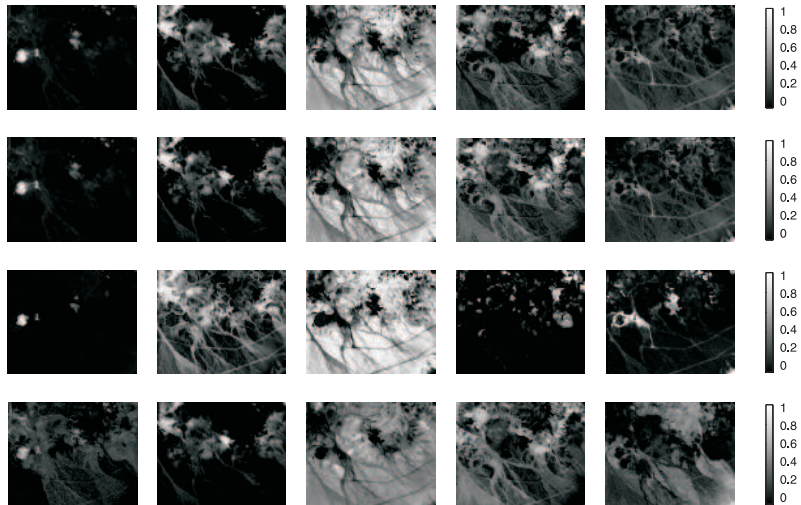


Fig. 2. Abundances estimated by four algorithms (From top to bottom: FCLS, NLCT, KFCLS, proposed)

- [6] G. Camps-Valls and L. Bruzzone, "Kernel-based methods for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 6, pp. 1351–1362, 2005.
- [7] H. Kwon and N. Nasrabadi, "Kernel orthogonal subspace projection for hyperspectral signal classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 12, pp. 2952–2962, 2005.
- [8] J. Broadwater and A. Banerjee, "A comparison of kernel functions for intimate mixture models," in *Proceedings of Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, Grenoble, France, 2009, pp. 1–4.
- [9] J. Broadwater, R. Chellappa, A. Banerjee, and P. Burlina, "Kernel fully constrained least squares abundance estimates," in *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*, Barcelona, Spain, 2007, pp. 4041–4044.
- [10] X. Wu, X. Li, and L. Zhao, "A kernel spatial complexity-based nonlinear unmixing method of hyperspectral imagery," in *Proceedings of Life System Modeling and Intelligent Computing*, ser. Lecture Notes in Computer Science, Wuxi, China, 2010, vol. 6330, pp. 451–458.
- [11] B. Hapke, "Bidirectional reflectance spectroscopy, 1, Theory," *Journal of Geophysical Research*, vol. 86, pp. 3039–3054, 1981.
- [12] J. Chen, C. Richard, J.-C. M. Bermudez, and P. Honeine, "Non-negative least-mean-square algorithm," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5225 – 5235, 2011.
- [13] D. Luenberger and Y. Ye, *Linear and nonlinear programming*, ser. International series in operations research & management science. Springer, 2008.
- [14] D. C. Heinz and C. Chang, "Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, pp. 529–545, 2001.
- [15] J. Nascimento and J. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 4, pp. 898–910, 2005.