NONLINEAR UNMIXING OF HYPERSPECTRAL IMAGES USING RADIAL BASIS FUNCTIONS AND ORTHOGONAL LEAST SQUARES

Y. Altmann, N. Dobigeon, J-Y. Tourneret

University of Toulouse
IRIT/INP-ENSEEIHT
France

S. McLaughlin

School of Engineering
University of Edinburgh
U.K.

ABSTRACT

This paper studies a linear radial basis function network (RBFN) for unmixing hyperspectral images. The proposed RBFN assumes that the observed pixel reflectances are nonlinear mixtures of known endmembers (extracted from a spectral library or estimated with an endmember extraction algorithm), with unknown proportions (usually referred to as abundances). We propose to estimate the model abundances using a linear combination of radial basis functions whose weights are estimated using training samples. The main contribution of this paper is to study an orthogonal least squares algorithm which allows the number of RBFN centers involved in the abundance estimation to be significantly reduced. The resulting abundance estimator is combined with a fully constrained estimation procedure ensuring positivity and sum-to-one constraints for the abundances. The performance of the nonlinear unmixing strategy is evaluated with simulations conducted on synthetic and real data.

Index Terms— Radial basis functions, hyperspectral image, spectral unmixing

1. INTRODUCTION

Spectral unmixing (SU) has been widely used in remote sensing for hyperspectral image analysis. SU assumes that the hyperspectral image pixels are mixtures of spectral components associated to pure materials (referred to as endmembers). According to physical considerations, the fractions of these endmembers (called abundances) have to satisfy positivity and sum-to-one constraints. The most common mixture model is the linear mixing model (LMM) that has been widely studied in the literature. However, the LMM has also shown some limitations which has motivated the consideration of nonlinear mixing models. These nonlinear mixing models include models based on intimate mixtures [1] or bilinear models [2, 3]. Unfortunately, these nonlinear mixing models are not appropriate to any kind of nonlinearity. For instance, bilinear models have nice properties to handle multipath effects as light scatters from one object to another. However, they cannot model the nonlinearity due to intimate mixtures [4]. This paper considers a linear radial basis function network (RBFN) that was originally studied in [5] for nonlinear SU. Motivations for considering neural networks for SU include the universal approximation theorem which states that any nonlinear function can be approximated by a neural network structure with any desired precision [6, p. 208].

This paper focuses on a supervised approach where the RBFN is trained using ground truth data. The ground truth data are used to estimate the inverse nonlinear function relating the observations to the abundances. Our main contribution is to study a new supervised training algorithm selecting a limited number of centers from the training data in order to reduce the computational complexity of the RBFN learning. The proposed algorithm has also the nice property to ensure positivity and sum-to-one constraints for the abundances.

Part of this work has been supported by Direction Generale de l'Armement, French Ministry of Defence

2. LINEAR RADIAL BASIS FUNCTION NETWORK

An observed pixel spectrum $\mathbf{y}$ of the hyperspectral image (in L spectral bands) is assumed to be a mixture of R known pure components spectra $\mathbf{m}_r, r = 1, \ldots, R$, according to the following nonlinear model

$$\mathbf{y} = f_{\mathbf{M}}(\mathbf{a}) + \mathbf{n}, \tag{1}$$

where $\mathbf{M} = [\mathbf{m}_1, \ldots, \mathbf{m}_R]$ is a known matrix containing the endmembers, $\mathbf{a} = [a_1, \ldots, a_R]^T$ is the unknown abundance vector, $f_{\mathbf{M}}$ is an unknown nonlinear function from $\mathbb{R}^R$ to $\mathbb{R}^L$ and $\mathbf{n}$ is an additive noise vector. The abundances of any pixel satisfy the following positivity and sum-to-one constraints

$$\sum_{r=1}^{R} a_r = 1, \;\; a_r \geq 0, \forall r \in \{1, \ldots, R\}. \tag{2}$$

The SU algorithm studied in this paper is based on the RBFN initially introduced in [5]. This network expresses the estimated abundance vector $\hat{\mathbf{a}} \in \mathbb{R}^R$ as a linear combination of N radial basis functions (RBFs)

$$\hat{\mathbf{a}} = \sum_{n=1}^{N} \phi_n(\mathbf{y}) \boldsymbol{w}_n \tag{3}$$

where N is the number of pixels of a ground truth image (image with known abundances), $\boldsymbol{w}_n = [w_{n,1}, \ldots, w_{n,R}]^T$ is the weight vector of the nth element of the ground truth image and $\phi_n(\mathbf{y})$ is the projection of the data vector $\mathbf{y}$ onto the nth basis function. The main motivation for unmixing hyperspectral images using an RBFN is that the proposed structure (linear combination of RBFs) has known capacities to approximate any nonlinearity with good accuracy. In particular, no prior knowledge about $f_{\mathbf{M}}$ is required and the RBFN is sufficiently general to handle linear and nonlinear mixtures. In the context of hyperspectral unmixing, the proposed RBFN is used as an efficient tool to invert the nonlinear relations relating the observation vector $\mathbf{y}$ to the abundance vector $\mathbf{a}$ defined in (1). This study focuses on Gaussian RBFs parameterized by their centers $\boldsymbol{c}_i$ and by a unique dispersion parameter σ^2 such that

$$\phi_n(\mathbf{y}) \triangleq \phi(\mathbf{y}, \boldsymbol{c}_n) = \exp\left(-\frac{\|\mathbf{y}-\boldsymbol{c}_n\|^2}{2\sigma^2}\right). \quad (4)$$

The RBF parameter σ^2 is chosen as the squared average distance between the pixels, ensuring that the individual RBFs are not too peaked or too flat. However, it is interesting to note that this parameter might be estimated jointly with the abundances (see [7, p. 123] for details). The next section describes a procedure for selecting a reduced number of RBF centers $\{\boldsymbol{c}_n\}_{n=1,\ldots,M}$ involved in (3) from the ground truth data, by ensuring $M \leq N$ is as small as possible.

3. ESTIMATING RBF CENTERS USING OLS

The main issue concerning the proposed RBFN is the choice of the centers defining the basis functions. For that purpose, we propose as in [5] to use training data with known abundances. Thus, we consider N training pixels $\mathbf{y}_1, \ldots, \mathbf{y}_N$ with their associated known abundance vectors $\mathbf{a}_1, \ldots, \mathbf{a}_N$. The centers of the proposed RBFN and their associated weights $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_N$ are determined using these training samples and the orthogonal least squares (OLS) procedure described in this section. By transposing (3) for the abundance vectors of the training samples, the following decomposition can be obtained

$$\boldsymbol{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_N]^T = \boldsymbol{\Phi}\boldsymbol{W} + \boldsymbol{E}, \quad (5)$$

where $\boldsymbol{\phi}(\mathbf{y}) = [\phi_1(\mathbf{y}), \ldots, \phi_N(\mathbf{y})]^T \in \mathbb{R}^N$ contains the projection of $\mathbf{y}$ on the N RBFs, $\boldsymbol{\Phi} = [\boldsymbol{\phi}(\mathbf{y}_1), \ldots, \boldsymbol{\phi}(\mathbf{y}_N)]^T = [\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_N]$ is an $N \times N$ matrix, $\boldsymbol{W} = [\boldsymbol{w}_1, \ldots, \boldsymbol{w}_N]^T$ is an $N \times R$ matrix containing the weight vectors of the N centers, $\boldsymbol{A}$ is an $N \times R$ matrix containing the abundances of the training vectors and $\boldsymbol{E}$ is a projection error matrix of size $N \times R$. The learning procedure proposed in [5] consists of estimating $\boldsymbol{W}$ by solving the following least squares problem

$$\min_{\boldsymbol{W}} \|\boldsymbol{A} - \boldsymbol{\Phi}\boldsymbol{W}\|_F^2 \quad (6)$$

where $\|.\|_F$ is the Frobenius norm. However, numerical problems can be obtained for large values of N, e.g., because the matrix $\boldsymbol{\Phi}$ is ill-conditioned. To alleviate this problem, we propose in this paper to reduce the complexity of the RBFN by selecting a reduced number of $M << N$ centers of the RBFN. For this, all the N training samples are first considered as potential centers. The use of a permutation matrix $\boldsymbol{\Pi}$ of size $N \times N$, allows the columns of $\boldsymbol{\Phi}$ to be rearranged in a matrix $\boldsymbol{\Phi}^{(\boldsymbol{\Pi})}$ as follows

$$\boldsymbol{A} = \boldsymbol{\Phi}\boldsymbol{\Pi}\boldsymbol{\Pi}^T\boldsymbol{W} + \boldsymbol{E} = \boldsymbol{\Phi}^{(\boldsymbol{\Pi})}\boldsymbol{W}^{(\boldsymbol{\Pi})} + \boldsymbol{E}, \quad (7)$$

where $\boldsymbol{\Phi}^{(\boldsymbol{\Pi})} = \boldsymbol{\Phi}\boldsymbol{\Pi}$ and $\boldsymbol{W}^{(\boldsymbol{\Pi})} = \boldsymbol{\Pi}^T\boldsymbol{W}$. Note that $\boldsymbol{\Pi}$ is chosen so that the M first columns of $\boldsymbol{W}^{(\boldsymbol{\Pi})}$ contain the projections of $\mathbf{y}_1, \ldots, \mathbf{y}_N$ on the M centers of highest interest. By decomposing $\boldsymbol{\Phi}^{(\boldsymbol{\Pi})}$ as a concatenation of two submatrices $\boldsymbol{\Phi}^{(\boldsymbol{\Pi})}_{1:M}$ and $\boldsymbol{\Phi}^{(\boldsymbol{\Pi})}_{M+1:N}$ containing the M relevant and $N-M$ non-relevant columns of $\boldsymbol{\Phi}^{(\boldsymbol{\Pi})}$, the abundance matrix can be written

$$\boldsymbol{A} = \boldsymbol{\Phi}^{(\boldsymbol{\Pi})}_{1:M}\boldsymbol{W}^{(\boldsymbol{\Pi})}_M + \boldsymbol{E}_M, \quad (8)$$

where $\boldsymbol{\Phi}^{(\boldsymbol{\Pi})}_{1:M}$ (resp. $\boldsymbol{\Phi}^{(\boldsymbol{\Pi})}_{M+1:N}$) contains the M first relevant (resp. the last non-relevant) columns of $\boldsymbol{\Phi}^{(\boldsymbol{\Pi})}$, $\boldsymbol{W}^{(\boldsymbol{\Pi})}_M$ (resp. $\boldsymbol{W}^{(\boldsymbol{\Pi})}_{N-M}$) contains the M first (resp. the $P-M$ last) rows of $\boldsymbol{W}^{(\boldsymbol{\Pi})}$ and $\boldsymbol{E}_M = \boldsymbol{\Phi}^{(\boldsymbol{\Pi})}_{M+1:N}\boldsymbol{W}^{(\boldsymbol{\Pi})}_{N-M} + \boldsymbol{E}$.

The permutation matrix $\boldsymbol{\Pi}$ and the number of centers M can be determined using an exhaustive search to minimize $\|\boldsymbol{E}_M\|_F^2$. However, this method requires prohibitive computational cost. As an alternative, the proposed center selection procedure is achieved in a forward regression manner (as in [5]) using a QR decomposition of $\boldsymbol{\Phi}^{(\boldsymbol{\Pi})}$. Indeed, the abundance matrix can be rewritten[1]

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R}\boldsymbol{W}^{(\boldsymbol{\Pi})} + \boldsymbol{E} = \boldsymbol{Q}\boldsymbol{\Theta} + \boldsymbol{E}, \quad (9)$$

where $\boldsymbol{Q} = [\boldsymbol{q}_1, \ldots, \boldsymbol{q}_N] =$ is an $N \times N$ matrix with orthogonal columns and $\boldsymbol{\Theta} = \boldsymbol{R}\boldsymbol{W}^{(\boldsymbol{\Pi})}$ is a new unknown weight matrix of size $N \times R$. After decomposing the matrix $\boldsymbol{Q}$ as the concatenation of two sub-matrices $\boldsymbol{Q}_{1:M}$ and $\boldsymbol{Q}_{M+1:P}$ containing the relevant and non-relevant columns of $\boldsymbol{Q}$, the abundance matrix can be written

$$\boldsymbol{A} = \boldsymbol{Q}_{1:M}\boldsymbol{\Theta}_M + \boldsymbol{E}_M, \quad (10)$$

where $\boldsymbol{\Theta}_M$ contains the M first rows of $\boldsymbol{\Theta}$. For any permutation matrix $\boldsymbol{\Pi}$, allowing relevant centers to be selected, (10) is used to estimate $\boldsymbol{\Theta}_M$ according to the least-squares principle.

To determine the best permutation matrix and a stopping rule for the proposed algorithm, an appropriate error criterion (referred to as error reduction ratio) has to be defined. Using the orthogonality of $\boldsymbol{Q}$, the output energy can be written

$$\begin{aligned}\boldsymbol{A}^T\boldsymbol{A} &= \sum_{n=1}^{N} \boldsymbol{\theta}_n^T\boldsymbol{q}_n^T\boldsymbol{q}_n\boldsymbol{\theta}_n + \boldsymbol{E}^T\boldsymbol{E} \\ &= \sum_{m=1}^{M} \boldsymbol{\theta}_m^T\boldsymbol{q}_m^T\boldsymbol{q}_m\boldsymbol{\theta}_m + \boldsymbol{E}_M^T\boldsymbol{E}_M. \end{aligned} \quad (11)$$

In order to measure the contribution of the first M orthogonal regressors, it makes sense to consider the following error reduction rate due to $\boldsymbol{q}_1, ..., \boldsymbol{q}_M$

$$\epsilon_M = \frac{\left\|\sum_{m=1}^{M} \boldsymbol{\theta}_m^T\boldsymbol{q}_m^T\boldsymbol{q}_m\boldsymbol{\theta}_m\right\|_F}{\left\|\boldsymbol{A}^T\boldsymbol{A}\right\|_F}. \quad (12)$$

This procedure is repeated in a forward regression manner (i.e., $M = 1, 2, \ldots$) until the relative error $\Delta_M = |\epsilon_{M-1} - \epsilon_M| / \epsilon_{M-1}$ is less than a given threshold $\rho > 0$. The resulting algorithm is summarized in Table 1. Note that the algorithm provides M relevant centers $\tilde{\boldsymbol{c}}_1 = \mathbf{y}_{n_1}, \ldots, \tilde{\boldsymbol{c}}_M = \mathbf{y}_{n_M}$ (out of the initial N training data) that will be used for abundance estimation. The next section proposes a modification of this algorithm to ensure positivity and sum-to-one constraints for the abundances.

4. CONSTRAINED ABUNDANCE ESTIMATION

Once the M most relevant centers $\{\tilde{\boldsymbol{c}}_m\}_{m=1,\ldots,M}$ associated with $\{\boldsymbol{\phi}_{n_m}\}_{m=1,\ldots,M}$ and the corresponding weight matrix $\boldsymbol{W}^{(\boldsymbol{\Pi})}_M$ have been selected (using the OLS procedure), the RBFN has to be used to estimate the abundance vector associated with a new observation vector $\mathbf{y}$. The estimated abundance vector of $\mathbf{y}$ might be defined as the output $\hat{\mathbf{a}}$ of the RBFN defined by

$$\hat{\mathbf{a}} = \sum_{m=1}^{M} \tilde{\phi}_m(\mathbf{y})\boldsymbol{w}_m^{(\boldsymbol{\Pi})} = \boldsymbol{W}_M^{(\boldsymbol{\Pi})T}\tilde{\boldsymbol{\phi}}(\mathbf{y}) \quad (13)$$

[1]The QR decomposition of $\boldsymbol{\Phi}^{(\boldsymbol{\Pi})}$ should be denoted $\boldsymbol{\Phi}^{(\boldsymbol{\Pi})} = \boldsymbol{Q}^{(\boldsymbol{\Pi})}\boldsymbol{R}^{(\boldsymbol{\Pi})}$. However, the upperscript $(\boldsymbol{\Pi})$ has been removed to simplify notations.

First step ($M = 1$)
for $n = 1 : N$ **do**
  Compute $\mathbf{q}_1^{(n)} = \boldsymbol{\phi}_n$
  (Define a permutation matrix $\mathbf{\Pi}$ such that $\mathbf{\Phi}^{(\mathbf{\Pi})} = [\boldsymbol{\phi}_n, \boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_{n-1}, \boldsymbol{\phi}_{n+1}, \ldots, \boldsymbol{\phi}_N]$)
  Compute $\boldsymbol{\theta}_1^{(n)}$ according to (10)
  Compute $\epsilon_1^{(n)}$ according to (12)
end for
Find n_1 such that $\epsilon_1 = \epsilon_1^{(n_1)} = \max_n \epsilon_1^{(n)}$,
Set $\mathbf{q}_1 = \mathbf{q}_1^{(n_1)}$, $\tilde{\mathbf{c}}_1 = \mathbf{y}_{n_1}$
Iterations ($M \geq 2$)
for $M = 1, 2, \ldots$ **do**
  $M = M + 1$
  for $n = 1 : N, n \neq n_1, \ldots, n_{M-1}$ **do**
    Compute $\mathbf{q}_M^{(n)}$ from $\boldsymbol{\phi}_n, \mathbf{q}_1, \ldots, \mathbf{q}_{M-1}$
    (Define a permutation matrix $\mathbf{\Pi}$ such that $\mathbf{\Phi}^{(\mathbf{\Pi})} = [\boldsymbol{\phi}_{n_1}, \ldots, \boldsymbol{\phi}_{n_{M-1}}, \{\boldsymbol{\phi}_i, i \notin \{n_1, \ldots, n_{M-1}\}\}]$)
    Compute $\boldsymbol{\theta}_M^{(n)}$ according to (10)
    Compute $\epsilon_M^{(n)}$ according to (12)
  end for
  Find n_M so that $\epsilon_M = \epsilon_M^{(n_M)} = \max_n \epsilon_M^{(n)}$,
  Set $\mathbf{q}_M = \mathbf{q}_M^{(n_M)}$, $\tilde{\mathbf{c}}_M = \mathbf{y}_{n_M}$
end for

Table 1. Center selection algorithm.

where $\tilde{\boldsymbol{\phi}}(\mathbf{y}) = [\tilde{\phi}_1(\mathbf{y}), \ldots, \tilde{\phi}_M(\mathbf{y})]^T$ is the projection of $\mathbf{y}$ on the M centers $\{\tilde{\mathbf{c}}_m\}_{m=1,\ldots,M}$ and $\boldsymbol{W}_M^{(\mathbf{\Pi})} = [\boldsymbol{w}_1^{(\mathbf{\Pi})}, \ldots, \boldsymbol{w}_M^{(\mathbf{\Pi})}]^T$. However, the positivity and sum-to-one constraints are not necessarily ensured when using (13) (even if these constraints can be satisfied for the training data). To satisfy these constraints, we propose to consider the following constrained optimization problem

$$\min_{\mathbf{a}} \left\| \tilde{\boldsymbol{\phi}}(\mathbf{y}) - \boldsymbol{W}_M^{(\mathbf{\Pi})T\dagger} \mathbf{a} \right\|_2^2 \quad \text{subject to} \quad (2) \tag{14}$$

where $\boldsymbol{W}_M^{(\mathbf{\Pi})T\dagger}$ is the pseudo-inverse of $\boldsymbol{W}_M^{(\mathbf{\Pi})T}$ and $\|.\|_2$ is the standard ℓ_2 norm. The minimization problem (14) has been explicitly addressed in [8] where the FCLS algorithm was introduced. The FCLS algorithm includes the sum-to-one constraint of the abundances as an additional observation equation in the criterion to be minimized. The following optimization problem is then obtained

$$\min_{\mathbf{a}} \left\| \begin{bmatrix} \tilde{\boldsymbol{\phi}}(\mathbf{y}) \\ \delta \end{bmatrix} - \begin{bmatrix} \boldsymbol{W}_M^{(\mathbf{\Pi})T\dagger} \\ \delta \mathbf{1}_R^T \end{bmatrix} \mathbf{a} \right\|_2^2 \tag{15}$$

subject to the non-negativity constraints for the abundance vector, where $\delta \in \mathbb{R}$ controls the impact of the sum-to-one constraint and $\mathbf{1}_R \in \mathbb{R}^R$ is a vector of ones. Note that a high value of δ will enforce the sum-to-one constraint (see [8] for more details). All simulations have been conducted in this study with $\delta = 10^5$.

5. SIMULATIONS

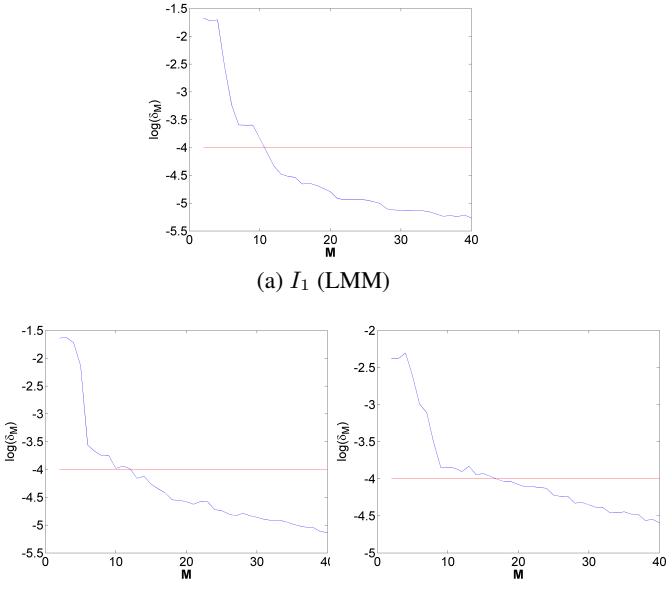
5.1. Synthetic data

The performance of the proposed RBFN is investigated by unmixing three synthetic images. The $R = 3$ endmembers associated with these images have been extracted from the spectral libraries provided with the ENVI software (i.e., green grass, olive green paint and galvanized steel metal). The first synthetic image I_1 has been generated using the standard linear mixing model (LMM). A second image I_2 has been generated according to the bilinear mixing model introduced in [3], referred to as "Fan model" (FM), whereas a third image I_3 has been generated according to the bilinear mixing model presented in [2], referred to as "Nascimento model" (NM). In each case, the learning procedure has been achieved using training sets of 2500 synthetic pixels (denoted as T_1, T_2 and T_3) generated according to the corresponding mixing models. For each image, the abundance vectors $\mathbf{a}_n$, (for $n = 1, \ldots, 2500$) have been uniformly generated in the simplex defined by the positivity and sum-to-one constraints. All images have been corrupted by an additive white Gaussian noise with signal-to-noise ratio $\text{SNR} = L^{-1}\sigma^{-2} \|f_{\text{M}}(\mathbf{a})\|^2 \simeq 15\text{dB}$. The threshold of the OLS procedure has been fixed to $\rho = 10^{-4}$.

Table 2 shows the number of selected centers and the corresponding error reduction ratio (ϵ_M). Fig. 1 shows the evolution of δ_M versus M. From these results, we can notice that less than 20 centers (out of the initial 2500) are sufficient to describe the three model mixtures. Note that the NM requires more centers than the LMM and FM to describe the mixture since the number of parameters involved the NM is equal to $R(R+1)/2$ as opposed to R for the LMM and FM (see [2] for more details). In order to illustrate the center selection procedure, Fig. 2 shows the $M = 13$ centers selected by the OLS algorithm applied to FM (right), out of the initial centers (left). It is important to note that the selected centers correspond to both pure and mixed pixels that are useful to describe the mixing model.

	I_1	I_2	I_3
M	11	13	17
$1 - \epsilon_M$	9.8×10^{-4}	1.3×10^{-3}	9.0×10^{-3}

Table 2. Number of centers after OLS.



(a) I_1 (LMM)

(b) I_2 (FM) (c) I_3 (NM)

Fig. 1. Evolution of δ_M versus M for the three images I_1, I_2 and I_3 (blue lines) and corresponding threshold ρ (red lines).

The quality of the unmixing procedure can be measured by comparing the estimated and actual abundances using the root mean square error $\text{RMSE} = \sqrt{\sum_{n=1}^N \|\tilde{\mathbf{a}}_n - \mathbf{a}_n\|^2 / NR}$ where $\mathbf{a}_n$ and $\tilde{\mathbf{a}}_n$ are the actual and estimated abundance vectors for the nth pixel of

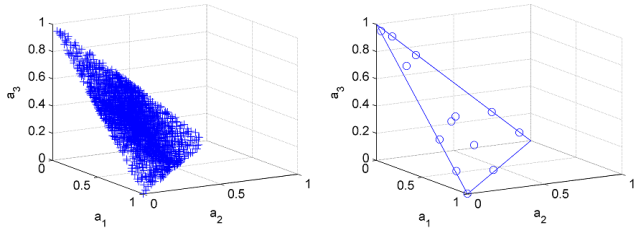


Fig. 2. Left: Initial 2500 centers. Right: $M = 13$ selected centers.

the image. Table 2 shows the RMSEs associated with the test images I_1, I_2, I_3 using the training sets T_1, T_2, T_3. In each case, the standard RBFN algorithm and its constrained version (denoted as CRBFN) have been considered. The unmixing performance in terms of RMSE does not change significantly when the OLS procedure is used. However, the proposed OLS algorithm allows the computational complexity of the learning step to be significantly reduced.

	RMSE ($\times 10^{-1}$)			
	without OLS		with OLS	
	RBFN	CRBFN	RBFN	CRBFN
I_1	0.409	0.407	0.411	0.403
I_2	0.391	0.378	0.376	0.393
I_3	0.541	0.532	0.547	0.544

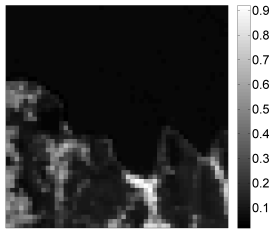
Table 3. RMSEs for the synthetic training data.

5.2. Real data

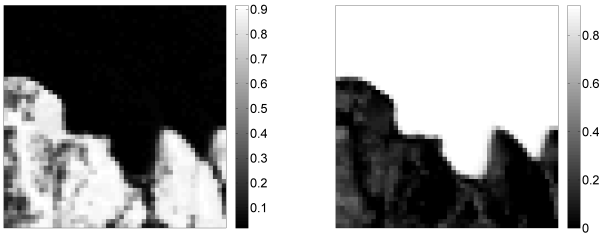
The real image considered in this section is composed of $L = 189$ spectral bands and was acquired in 1997 by the airborne visible infrared imaging spectrometer (AVIRIS) over Moffett Field (CA, USA). A sub-image of size 50×50 pixels has been chosen here to evaluate the proposed unmixing procedure. The scene is mainly composed of water, vegetation and soil. The endmembers have been extracted by the VCA algorithm [9] with $R = 3$. These endmembers have been used to generate 2500 training pixels according to the 3 mixing models investigated in the previous section. The abundance vectors $\mathbf{a}_n, n = 1, \ldots, 2500$, have been uniformly generated in the simplex defined by the positivity and sum-to-one constraints. All generated images have been corrupted by an additive Gaussian noise with SNR $\simeq$ 15dB. Fig. 3 shows typical abundance maps estimated by the CRBFN whose centers have been determined using the OLS procedure (the learning step has been conducted with data generated using the FM). The resulting maps are very encouraging and are close to maps obtained via the optimization method introduced in [3], which assumes a known mixing model. Thus demonstrating that the CRBFN has the capacity to invert the relation between abundance vectors and the observed pixels for this example.

6. CONCLUSION

A new nonlinear unmixing algorithm based on a radial basis function network was presented for hyperspectral imagery. This algorithm selected a limited number of relevant network centers from training data using orthogonal least squares. The orthogonal least squares algorithm allowed the complexity of the network to be reduced without significant degradation of the unmixing quality. A modification of the algorithm was also proposed to enforce positivity and sum-to-one constraints for the abundance vectors. This algorithm provided



(a) Abundance of vegetation



(b) Abundance of soil (c) Abundance of water

Fig. 3. Estimated abundance maps (CRBFN applied to FM).

promising results. Future works include the adaptive update of the weights and centers for unsupervised spectral unmixing. Dictionary learning algorithms [10] could also be investigated for center selection in radial basis function networks.

7. REFERENCES

[1] B. W. Hapke, "Bidirectional reflectance spectroscopy. I. Theory," *J. Geophys. Res.*, vol. 86, pp. 3039–3054, 1981.

[2] J. M. P. Nascimento and J. M. Bioucas-Dias, "Nonlinear mixture model for hyperspectral unmixing," in *Proc. SPIE Image and Signal Processing for Remote Sensing XV*, L. Bruzzone, C. Notarnicola, and F. Posa, Eds., vol. 7477, no. 1. SPIE, 2009, p. 74770I.

[3] 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," *Remote Sensing of Environment*, vol. 30, no. 11, pp. 2951–2962, June 2009.

[4] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Trans. Geosci. and Remote Sensing*, 2011, to appear.

[5] 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 Geosci. and Remote Sensing Lett.*, vol. 39, no. 8, pp. 2314–2318, Aug. 2001.

[6] S. Haykin, *Neural networks: a comprehensive foundation*, 2nd ed. New Jersey: Prentice-Hall, 1999.

[7] S. Theodoridis and K. Koutroumbas, *Pattern recognition*. London: Academic Press Inc, 1999.

[8] D. C. Heinz and C.-I Chang, "Fully constrained least-squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. and Remote Sensing*, vol. 29, no. 3, pp. 529–545, March 2001.

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

[10] C. Richard, J. C. Bermudez, and P. Honeine, "Online prediction of time series data with kernels," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 1058–1067, March 2009.