

# Estimation of Signal Subspace on Hyperspectral Data

José M. Bioucas-Dias<sup>a</sup> and José M. P. Nascimento<sup>b</sup>

<sup>a</sup> Instituto Superior Técnico and Instituto de Telecomunicações,

<sup>b</sup> Instituto Superior de Engenharia de Lisboa and Instituto de Telecomunicações,

Lisbon, Portugal

## ABSTRACT

Dimensionality reduction plays a crucial role in many hyperspectral data processing and analysis algorithms. This paper proposes a new mean squared error based approach to determine the signal subspace in hyperspectral imagery. The method first estimates the signal and noise correlations matrices, then it selects the subset of eigenvalues that best represents the signal subspace in the least square sense. The effectiveness of the proposed method is illustrated using simulated and real hyperspectral images.

**Keywords:** Dimensionality reduction; Model Selection; Order Selection; Hyperspectral Imagery; Linear Mixture; Signal Subspace.

## 1. INTRODUCTION

Hyperspectral remote sensing exploits the fact that all substances scatter electromagnetic energy, at specific wavelengths, in distinctive patterns related to their molecular composition.<sup>1</sup> Hyperspectral sensors use many contiguous bands of high spectral resolution covering the visible, near-infrared, and shortwave infrared spectral bands ( $0.3 - 2.5\mu m$ ).<sup>2</sup> Very often, the resolution cell corresponding to a single pixel in an image contains several substances. In this situation, the scattered energy is a mixing of the endmember spectra.<sup>3</sup>

Each pixel of an hyperspectral image can be represented as a vector in the space  $\mathfrak{R}^L$ , where  $L$  is the number of bands. Under the linear mixing scenario, the spectral vectors are a linear combination of a few vectors, the so-called endmember signatures. Therefore, the dimensionality of data (number of endmembers) is usually much lower than the number of bands.

A key problem in dimensionality reduction in hyperspectral imagery is the determination of the number of endmembers, termed intrinsic dimension (ID) of the data set. The estimation of the ID allows a correct dimension reduction and thus gains in computational time and complexity. Moreover, the projection of spectral vectors onto a subspace of lower dimension improves the signal-to-noise ratio ( $SNR$ ).

There are basically two approaches for estimating ID<sup>5</sup>: global and local. The first estimates ID of data set as a whole. The second estimates ID using information contained in sample neighborhoods. The latter approach avoids the projection of data onto a lower-dimensional space. Projection techniques, which are generally used as global approaches, seek for the best subspace to project data by minimizing an objective function. For example principal component analysis (PCA)<sup>6</sup> seeks the projection that best represents data in the least square sense; maximum noise fraction (MNF)<sup>7</sup> or noise adjusted principal components (NAPC)<sup>8</sup> seeks the projection that optimizes the ratio of noise power to signal power. This is in contrast with PCA where no noise model is used.

---

José M. Bioucas-Dias: Address: Instituto Superior Técnico, Instituto de Telecomunicações, Av. Rovisco Pais, Torre Norte, Piso 10, 1049-001, Lisbon, Portugal, E-mail: bioucas@lx.it.pt, Telephone: +351.21.8418466, +Fax: +351.21.8418472  
José M. P. Nascimento: Address: Instituto Superior de Engenharia de Lisboa, R. Conselheiro Emídio Navarro N 1, edifício DEETC, 1950-062, Lisbon, Portugal, E-mail: zen@isel.pt, Telephone: +351.21.8317237, Fax: +351.21.8317114

Topological methods are local approaches that estimate the topological dimension of a data set.<sup>9</sup> For example curvilinear component analysis (CCA)<sup>10</sup> and curvilinear distance analysis (CDA)<sup>11</sup> are non-linear projections that are based on the preservation of the local topology.

Recently Harsanyi, Farrand, and Chang developed a Neyman-Pearson detector, referred to as HFC, to determine the number of spectral endmembers in hyperspectral data (see<sup>12</sup>).

This paper proposes a new method to estimate the number of endmembers and the signal subspace in hyperspectral linear mixtures. The method begins by estimating signal and noise correlation matrices. The latter is obtained based on multiple regression theory. The signal subspace is then identified by selecting the set of signal eigenvalues that best represents, in the least square sense,<sup>13</sup> the mean value of data set.

The paper is structured as follows. Section 2 describes the fundamentals of the proposed method. Section 3 evaluate the proposed algorithm using simulated and real data. Section 4 ends the paper by presenting some concluding remarks.

## 2. SUBSPACE ESTIMATION

Let  $\mathbf{Y} = [\mathbf{Y}_1, \mathbf{Y}_2 \dots \mathbf{Y}_N]$  be a  $L \times N$  matrix of spectral vectors, one per pixel, where  $N$  is the number of pixels and  $L$  the number of bands. Assuming a linear mixing scenario, each observed spectral vector is given by

$$\begin{aligned} \mathbf{y} &= \mathbf{x} + \mathbf{n} \\ &= \mathbf{M}\mathbf{s} + \mathbf{n}, \end{aligned} \quad (1)$$

where  $\mathbf{y}$  is an  $L$ -vector,  $\mathbf{M} \equiv [\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_p]$  is the mixing matrix ( $\mathbf{m}_i$  denotes the  $i$ th endmember signature and  $p$  is the number of endmembers present in the covered area),  $\mathbf{s} = [s_1, s_2, \dots, s_p]^T$  is the abundance vector containing the fractions of each endmember (the notation  $(\cdot)^T$  stands for vector transposed) and  $\mathbf{n}$  models system additive noise.

Owing to physical constraints,<sup>14</sup> abundance fractions are non-negative ( $\mathbf{s} \succeq 0$ ) and satisfy the so-called positivity constraint  $\mathbf{1}^T \mathbf{s} = 1$ , where  $\mathbf{1}$  is a  $p \times 1$  vector of ones.

The correlation matrix of vector  $\mathbf{y}$  is  $\mathbf{R}_y = \mathbf{R}_x + \mathbf{R}_n$ , where  $\mathbf{R}_x = \mathbf{M}\mathbf{R}_s\mathbf{M}^T$  is the signal correlation matrix,  $\mathbf{R}_n$  is the noise correlation matrix, and  $\mathbf{R}_s$  is the abundance correlation matrix. An estimate of the signal correlation matrix is given by

$$\widehat{\mathbf{R}}_x = \widehat{\mathbf{R}}_y - \widehat{\mathbf{R}}_n, \quad (2)$$

where  $\widehat{\mathbf{R}}_y = \mathbf{Y}\mathbf{Y}^T/N$  is the sample correlation matrix of  $\mathbf{Y}$ , and  $\widehat{\mathbf{R}}_n$  is an estimate of noise correlation matrix.

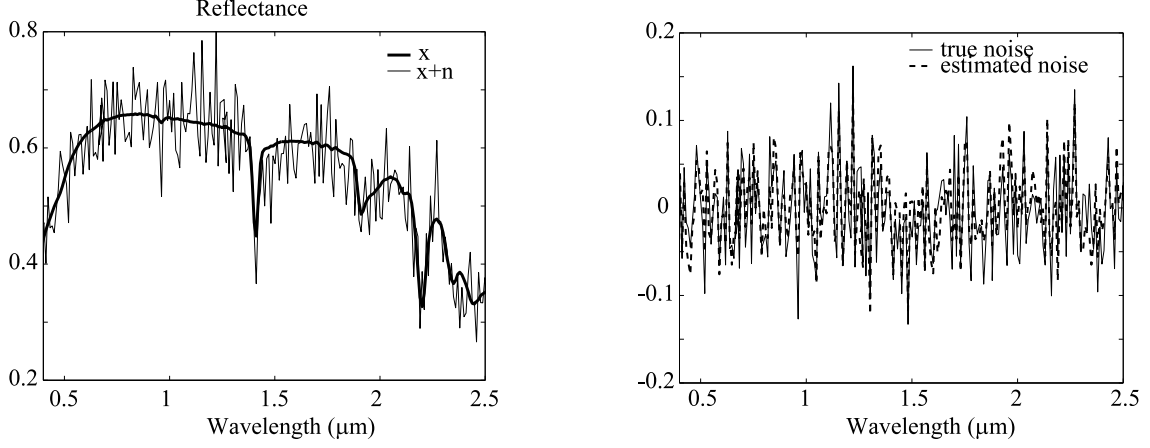
Define  $\mathbf{Y}^i = [Y_{i1}, \dots, Y_{iN}]^T$ , i.e.  $\mathbf{Y}^i$  is the transpose of the  $i$ th line of matrix  $\mathbf{Y}$ , thus containing the data read by the hyperspectral sensor at the  $i$ th band for all image pixels. Define also the matrix  $\mathbf{Y}^{\partial_i} = [\mathbf{Y}^1, \dots, \mathbf{Y}^{i-1}, \mathbf{Y}^{i+1}, \dots, \mathbf{Y}^L]$ .

Assuming that the dimension of the signal subspace is much lower than the number of bands, then the noise correlation matrix  $\widehat{\mathbf{R}}_n$  can be inferred based on multiple regression theory.<sup>15</sup> This consists in assuming that

$$\mathbf{Y}^i = \mathbf{Y}^{\partial_i} \boldsymbol{\beta}_i + \boldsymbol{\epsilon}_i, \quad (3)$$

where  $\mathbf{Y}^{\partial_i}$  is the explanatory data matrix,  $\boldsymbol{\beta}_i = [\beta_1, \dots, \beta_{L-1}]^T$  is the regression vector, and  $\boldsymbol{\epsilon}_i$  are modeling errors. For each  $i \in \{1, \dots, L\}$ , the regression vector is given by  $\boldsymbol{\beta}_i = [\mathbf{Y}^{\partial_i}]^\# \mathbf{Y}^i$ , where  $(\cdot)^\#$  denotes pseudo-inverse matrix. Finally, we compute  $\widehat{\boldsymbol{\epsilon}}_i = \mathbf{Y}^i - \mathbf{Y}^{\partial_i} \widehat{\boldsymbol{\beta}}_i$  and its sample correlation matrix  $\widehat{\mathbf{R}}_n$ .

Figure 1 left shows simulated reflectance  $\mathbf{x}$  and reflectance plus noise  $\mathbf{x} + \mathbf{n}$  for a given pixel. Figure 1 right, presents true and estimated noise for the same pixel. Notice the similarity.



**Figure 1.** Left: Illustration of the noise estimation; Bold line: Reflectance of a pixel; Narrow line: Noise corrupted reflectance; Right: solid line: true noise; dashed line: estimated noise.

Let the singular value decomposition (SVD) of  $\widehat{\mathbf{R}}_x$  be,

$$\widehat{\mathbf{R}}_x = \mathbf{E}\mathbf{\Sigma}\mathbf{E}^T, \quad (4)$$

where  $\mathbf{E} = [\mathbf{e}_1, \dots, \mathbf{e}_k, \mathbf{e}_{k+1}, \dots, \mathbf{e}_L]$  is a matrix with the singular vectors ordered by the descendent magnitude of the respective singular values. The space  $\mathfrak{R}^L$  can be splitted into two orthogonal subspaces:  $\langle E_k \rangle$  spanned by  $\mathbf{E}_k = [\mathbf{e}_1, \dots, \mathbf{e}_k]$  and  $\langle E_k^\perp \rangle$  spanned by  $\mathbf{E}_k^\perp = [\mathbf{e}_{k+1}, \dots, \mathbf{e}_L]$ , where  $k$  is the order of the signal subspace.

Since hyperspectral mixtures have nonnegative components, the projection of the mean value of  $\mathbf{Y}$  onto any eigenvector  $\mathbf{e}_i$ ,  $1 \leq i \leq k$ , is always nonzero. Therefore, the signal subspace can be identified by finding the subset of eigenvalues that best represents, in the least square sense, the mean value of data set.

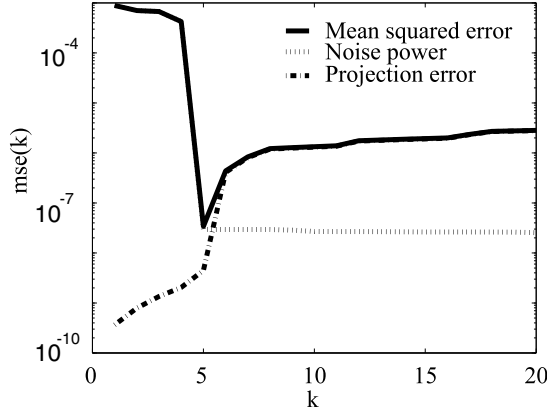
The sample mean value of  $\mathbf{Y}$  is

$$\begin{aligned} \bar{\mathbf{y}} &= \frac{1}{N} \sum_{i=1}^N \mathbf{Y}_i \\ &= \frac{1}{N} \mathbf{M} \sum_{i=1}^N \mathbf{s}_i + \frac{1}{N} \sum_{i=1}^N \mathbf{n}_i \\ &= \mathbf{c} + \mathbf{w}, \end{aligned} \quad (5)$$

where  $\mathbf{c}$  is in the signal subspace and  $\mathbf{w} \sim \mathcal{N}(0, \mathbf{R}_n/N)$  [the notation  $\mathcal{N}(\mu, \mathbf{C})$  stands for normal density function with mean  $\mu$  and covariance  $\mathbf{C}$ ]. Let  $\mathbf{c}_k$  be the projection of  $\mathbf{c}$  onto  $\langle E_k \rangle$ . The estimation of  $\mathbf{c}_k$  can be obtained by projecting  $\bar{\mathbf{y}}$  onto the signal subspace  $\langle E_k \rangle$ , i.e.,  $\widehat{\mathbf{c}}_k = \mathbf{P}_k \bar{\mathbf{y}}$ , where  $\mathbf{P}_k = \mathbf{E}_k \mathbf{E}_k^T$  is the projection matrix onto  $\langle E_k \rangle$ .

The first and second order moments of the estimated error  $\mathbf{c} - \widehat{\mathbf{c}}_k$  are

$$\begin{aligned} E[\mathbf{c} - \widehat{\mathbf{c}}_k] &= \mathbf{c} - E[\widehat{\mathbf{c}}_k] \\ &= \mathbf{c} - E[\mathbf{P}_k \bar{\mathbf{y}}] \\ &= \mathbf{c} - \mathbf{P}_k \mathbf{c} \\ &= \mathbf{c} - \mathbf{c}_k \\ &\equiv \mathbf{b}_k, \end{aligned} \quad (6)$$



**Figure 2.** Mean squared error versus  $k$ , with  $SNR = 35$  dB,  $p = 5$ ; (first experiment)

$$E[(\mathbf{c} - \hat{\mathbf{c}}_k)(\mathbf{c} - \hat{\mathbf{c}}_k)^T] = \mathbf{b}_k \mathbf{b}_k^T + \mathbf{P}_k \mathbf{R}_n \mathbf{P}_k^T / N, \quad (7)$$

where the bias  $\mathbf{b}_k = \mathbf{P}_k^\perp \mathbf{c}$  is the projection of  $\mathbf{c}$  onto the space  $\langle E_k^\perp \rangle$ . Therefore the density of the estimated error  $\mathbf{c} - \hat{\mathbf{c}}_k$  is  $\mathcal{N}(\mathbf{b}_k, \mathbf{b}_k^T \mathbf{b}_k + \mathbf{P}_k \mathbf{R}_n \mathbf{P}_k^T / N)$ ,

The mean squared error between  $\mathbf{c}$  and  $\hat{\mathbf{c}}_k$  is

$$\begin{aligned} \text{mse}(k) &= E[(\mathbf{c} - \hat{\mathbf{c}}_k)^T (\mathbf{c} - \hat{\mathbf{c}}_k)] \\ &= \text{tr}\{E[(\mathbf{c} - \hat{\mathbf{c}}_k)(\mathbf{c} - \hat{\mathbf{c}}_k)^T]\} \\ &= \mathbf{b}_k^T \mathbf{b}_k + \text{tr}(\mathbf{P}_k \mathbf{R}_n \mathbf{P}_k^T / N), \end{aligned} \quad (8)$$

where  $\text{tr}(\cdot)$  denote the trace operator. Since we do not know the bias  $\mathbf{b}_k$ , an approximation of Eq. (8) can be achieved by using the bias estimate  $\hat{\mathbf{b}}_k = \mathbf{P}_k^\perp \bar{\mathbf{y}}$ . However,  $E[\hat{\mathbf{b}}_k] = \mathbf{b}_k$  and  $E[\hat{\mathbf{b}}_k^T \hat{\mathbf{b}}_k] = \mathbf{b}_k^T \mathbf{b}_k + \text{tr}(\mathbf{P}_k^\perp \mathbf{R}_n \mathbf{P}_k^\perp / N)$ , i.e., an unbiased estimate of  $\mathbf{b}_k^T \mathbf{b}_k$  is  $\hat{\mathbf{b}}_k^T \hat{\mathbf{b}}_k - \text{tr}(\mathbf{P}_k^\perp \mathbf{R}_n \mathbf{P}_k^\perp / N)$ . The criteria for the signal subspace order determination is then

$$\begin{aligned} \hat{k} &= \arg \min_k (\hat{\mathbf{b}}_k^T \hat{\mathbf{b}}_k + \text{tr}(\mathbf{P}_k \mathbf{R}_n \mathbf{P}_k^T / N) - \text{tr}(\mathbf{P}_k^\perp \mathbf{R}_n \mathbf{P}_k^\perp / N)) \\ &= \arg \min_k (\bar{\mathbf{y}}^T \mathbf{P}_k^\perp \mathbf{P}_k^\perp \bar{\mathbf{y}} + 2\text{tr}(\mathbf{P}_k \mathbf{R}_n / N) - \text{tr}(\mathbf{R}_n / N)) \\ &= \arg \min_k (\bar{\mathbf{y}}^T \mathbf{P}_k^\perp \bar{\mathbf{y}} + 2\text{tr}(\mathbf{P}_k \mathbf{R}_n / N)), \end{aligned} \quad (9)$$

where we have used  $\mathbf{P} = \mathbf{P}^T$  and  $\mathbf{P}^2 = \mathbf{P}$  for any projection matrix.

Each term of (9) have a clear meaning: the first accounts for projection error power and it is a decreasing function of  $k$ ; the second accounts for noise power and it is an increasing function of  $k$ .

### 3. EXPERIMENTS

#### 3.1. Computer Simulations

In this section we test the proposed method in simulated scenes. The spectral signatures are selected from the U.S. geological survey (USGS) digital spectral library.<sup>16</sup> Abundance fractions are generated according to a Dirichlet distribution given by

$$p(\alpha_1, \alpha_2, \dots, \alpha_p) = \frac{\Gamma(\mu_1 + \mu_2 + \dots + \mu_p)}{\Gamma(\mu_1) \Gamma(\mu_2) \dots \Gamma(\mu_p)} \alpha_1^{\mu_1 - 1} \alpha_2^{\mu_2 - 1} \dots \alpha_p^{\mu_p - 1}, \quad (10)$$

**Table 1.** Signal subspace dimension  $\hat{k}$  as function of  $SNR$  and of  $p$ ; Bold: Proposed method; In brackets: VD estimation with NWHFC method and  $P_f = 10^{-4}$ .

		Signal subspace dimension $\hat{k}$				
Method		<b>New</b> (VD)	<b>New</b> (VD)	<b>New</b> (VD)	<b>New</b> (VD)	<b>New</b> (VD)
$SNR$ (in dB)		50	35	25	15	5
$p = 3$		<b>3</b> (3)	<b>3</b> (3)	<b>3</b> (4)	<b>3</b> (4)	<b>3</b> (2)
$p = 5$		<b>5</b> (6)	<b>5</b> (6)	<b>5</b> (6)	<b>5</b> (6)	<b>4</b> (3)
$p = 10$		<b>10</b> (11)	<b>10</b> (11)	<b>10</b> (9)	<b>8</b> (8)	<b>6</b> (2)
$p = 15$		<b>15</b> (16)	<b>15</b> (15)	<b>13</b> (13)	<b>9</b> (9)	<b>5</b> (2)

where  $0 \leq \alpha_i \leq 1$ ,  $\sum_{i=1}^p \alpha_i = 1$ , and  $\Gamma(\cdot)$  denotes the Gamma function. The mean value of the  $i$ th endmember fraction  $\alpha_i$  is  $E[\alpha_i] = \mu_i / \sum_{k=1}^p \mu_k$ .

The results next presented are organized into two experiments: in the first experiment the method is evaluated with respect to the  $SNR$  and to the number of endmembers  $p$ . We define  $SNR$  as

$$SNR \equiv 10 \log_{10} \frac{E[\mathbf{x}^T \mathbf{x}]}{E[\mathbf{n}^T \mathbf{n}]} \quad (11)$$

In the second experiment, the method is evaluated when a subset of the endmembers are present only in a few pixels of the scene.

In the first experiment, the hyperspectral scene has  $10^4$  pixels and the numbers of endmembers varies from 3 to 15. The abundance fractions are Dirichlet distributed with mean value  $\mu_i = 1/p$ , for  $i = 1, \dots, p$ .

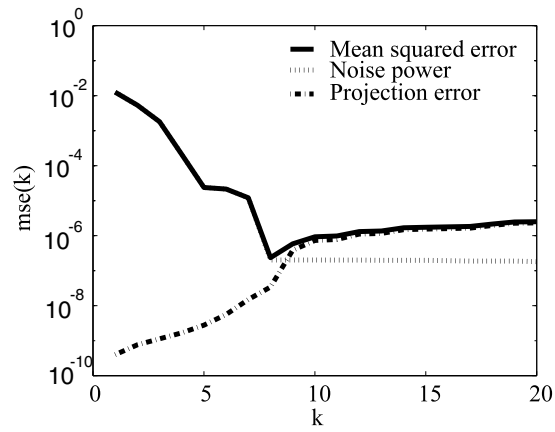
Fig. 2 left shows the evolution of the mean squared error, i.e., of  $\bar{\mathbf{y}}^T \mathbf{P}_k^\perp \bar{\mathbf{y}} + 2\text{tr}(\mathbf{P}_k \mathbf{R}_n / N)$  as a function of the parameter  $k$ , for  $SNR = 35$  dB and  $p = 5$ . The minimum of the mean squared error occurs for  $k = 5$ , which is exactly the number of endmembers present in the image.

Table 1 presents the signal subspace order estimate as function of the  $SNR$  and of  $p$ . In this table we compare the proposed method and the virtual dimensionality (VD), recently proposed in.<sup>12</sup> The VD was estimated by the NWHFC based eigen-thresholding method using the Neyman-Pearson test with the false-alarm probability set to  $P_f = 10^{-4}$ . The proposed method finds the correct ID for  $SNR$  larger than 25 dB, and underestimates ID as the  $SNR$  decreases. In comparison with the NWHFC algorithm, the proposed approach yields systematically equal or better results.

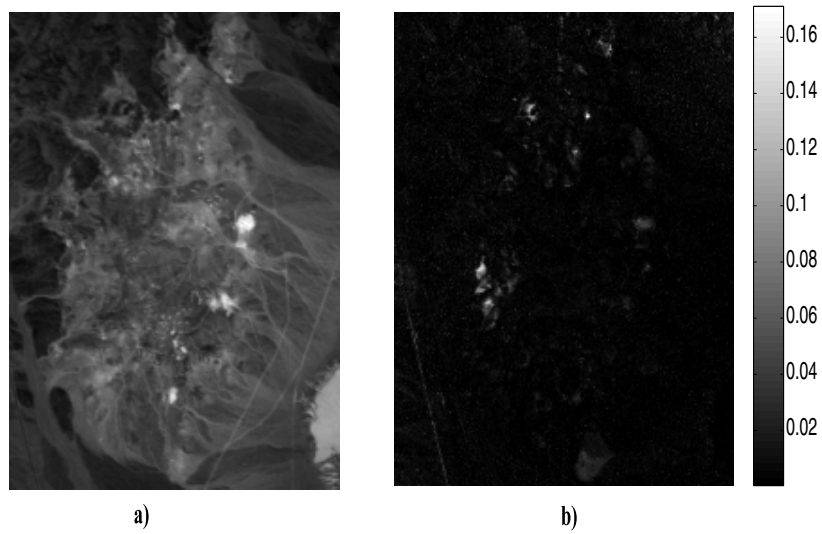
In the second experiment  $SNR = 35$  dB and  $p = 8$ . The first five endmembers have a Dirichlet distribution as in the previous experiment and the other three are forced to appear only in 4 pixels each one. Fig. 3 shows the mean squared error versus  $k$ , when  $p = 8$ . The minimum of  $\text{mse}(k)$  is achieved with  $k = 8$ . This means that the method is able to detect rare endmembers in the image. However, this ability degrades as  $SNR$  decreases, as expected.

### 3.2. Cuprite Experiments

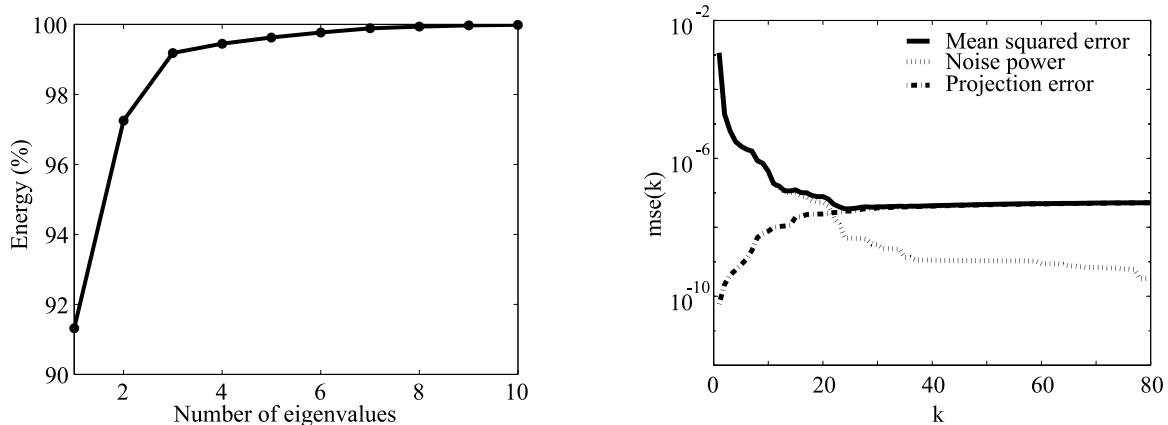
In this section, we apply the proposed method to real hyperspectral data collected by the AVIRIS<sup>17</sup> sensor over Cuprite, Nevada. Cuprite is a mining area in southern Nevada with mineral and little vegetation.<sup>18</sup> Cuprite test site, located approximately 200 Km northwest of Las Vegas is a relatively undisturbed acid-sulfate hydrothermal system near highway 95. The geology and alteration were previously mapped in detail.<sup>19, 20</sup> A geologic summary and a mineral map can be found in.<sup>18</sup> This site has been extensively used for remote sensing experiments over the past years.<sup>21, 22</sup> This study is based on subimage (  $250 \times 190$  pixels and 224 bands) of a data set acquired



**Figure 3.** Mean squared error versus  $k$ , with  $SNR = 35$  dB,  $p = 8$  (3 spectral vectors occur only on 4 pixels each; (second experiment))



**Figure 4.** Left: Band 30 (wavelength  $\lambda = 667.3nm$ ) of the subimage of AVIRIS cuprite Nevada data set; Right: percentage of energy in the subspace  $E_{9:23}$ .



**Figure 5.** Left: Percentage of signal energy as function of the number of eigenvalues; Right: mean squared error versus  $k$  for cuprite data set.

on the AVIRIS flight 19 June 1997 (see Fig. 4 left). AVIRIS instrument covers the spectral region from  $0.41\mu\text{m}$  to  $2.45\mu\text{m}$  in 224 bands with  $10\text{nm}$  bands. Flying at an altitude of  $20\text{km}$ , it has an instantaneous field of view (IFOV) of  $20\text{m}$  and views a swath over  $10\text{km}$  wide.

The proposed method when applied to this data set estimates  $\hat{k} = 23$  (see Fig. 5 right). According to the truth data presented in,<sup>18</sup> there are 8 materials in these area. This difference is due to a) the presence of rare pixels not accounted for in the truth data<sup>18</sup> and b) spectral variability.

The bulk of spectral energy is explained with only a few eigenvectors. This can be observed from Fig. 5 left, where the accumulated signal energy is plotted as function of the eigenvalue index. The energy contained in the first 8 eigenvalues is 99.94% of the total signal energy. This is further confirmed in Fig. 4, right, where we show, in gray level and for each pixel, the percentage of energy contained in the subspace  $\langle E_{9:23} \rangle = \langle [\mathbf{e}_9, \dots, \mathbf{e}_{23}] \rangle$ . Notice that only a few (rare) pixels contain energy in  $\langle E_{9:23} \rangle$ . Furthermore, these energies are a very small percentage of the correspondent spectral vector energies (less than 0.16%) in this subspace.

The VD estimated by the HFC based eigen-thresholding method<sup>12</sup> ( $P_f = 10^{-3}$ ) on the same data set yields  $\hat{k} = 20$ . A lower value of  $P_f$  would lead to a lower number of endmembers. This result seems to indicate that the proposed method performs better than the HFC with respect to rare materials.

## 4. CONCLUSIONS

The determination of the signal subspace dimensionality is a difficult and challenging task. In this paper, we have proposed a method to estimate the dimensionality of hyperspectral linear mixtures. The method is based on the mean squared error criteria.

A set of experiments with simulated and real data leads to the conclusion that the method is an useful tool in hyperspectral data analysis, yielding comparable or better results than the state-of-the-art methods.

## ACKNOWLEDGMENTS

This work was supported by the Fundação para a Ciência e Tecnologia, under the projects POSC/EEA-CPS/61271/2004 and PDCTE/CPS/49967/2003 and by Departamento de Engenharia de Electrónica e Telecomunicações e de Computadores of the Instituto Superior de Engenharia de Lisboa.

## REFERENCES

1. B. Hapke, *Theory of Reflectance and Emittance Spectroscopy*, Cambridge Univ. Press, Cambridge, U. K., 1993.
2. G. Shaw and D. Manolakis, "Signal processing for hyperspectral image exploitation," *IEEE Signal Processing Mag.* **19**(1), pp. 12–16, 2002.
3. T. M. Lillesand, R. W. Kiefer, and J. W. Chipman, *Remote Sensing and Image Interpretation*, John Wiley & Sons, Inc., fifth ed., 2004.
4. L. Jimenez and D. Landgrebe, "Supervised classification in high-dimensional space: Geometrical, statistical, and asymptotical properties of multivariate data," *IEEE Trans. Syst., Man, Cybern. C* **28**, pp. 39–54, 1998.
5. A. K. Jain and R. C. Dubes, *Algorithms for clustering data*, Prentice Hall, N. J., 1988.
6. I. T. Jolliffe, *Principal Component Analysis*, Springer Verlag, New York, 1986.
7. A. Green, M. Berman, P. Switzer, and M. D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Trans. Geosci. Remote Sensing* **26**(1), pp. 65–74, 1994.
8. J. B. Lee, S. Woodyatt, and M. Berman, "Enhancement of high spectral resolution remote-sensing data by noise-adjusted principal components transform," *IEEE Trans. Geosci. Remote Sensing* **28**, pp. 295–304, 1990.
9. J. Bruske and G. Sommer, "Intrinsic dimensionality estimation with optimally topologic preserving maps," *IEEE Trans. Pattern Anal. Machine Intell.* **20**(5), pp. 572–575, 1998.
10. P. Demartines and J. Héroult, "Curvilinear component analysis : A self-organizing neural network for nonlinear mapping of data sets," *IEEE Trans. Neural Networks* **8**(1), pp. 148–154, 1997.
11. M. Lennon, G. Mercier, M. Mouchot, and L. Hubert-Moy, "Curvilinear component analysis for nonlinear dimensionality reduction of hyperspectral images," in *Proc. of the SPIE Symp. on Remote Sensing Conference on Image and Signal Processing for Remote Sensing VII*, **4541**, 2001.
12. C.-I. Chang, *Hyperspectral Imaging: Techniques for spectral detection and classification*, Kluwer Academic, New York, 2003.
13. L. L. Scharf, *Statistical Signal Processing, Detection Estimation and Time Series Analysis*, Addison-Wesley Pub. Comp., 1991.
14. D. Manolakis, C. Siracusa, and G. Shaw, "Hyperspectral subpixel target detection using linear mixing model," *IEEE Trans. Geosci. Remote Sensing* **39**(7), pp. 1392–1409, 2001.
15. R. Roger and J. Arnold, "Reliably estimating the noise in aviris hyperspectral imagers," *International Journal of Remote Sensing* **17**(10), pp. 1951–1962, 1996.
16. R. N. Clark, G. A. Swayze, A. Gallagher, T. V. King, and W. M. Calvin, "The U.S. geological survey digital spectral library: Version 1: 0.2 to 3.0  $\mu\text{m}$ ," open file report 93-592, U.S. Geological Survey, 1993.
17. G. Vane, R. Green, T. Chrien, H. Enmark, E. Hansen, and W. Porter, "The airborne visible/infrared imaging spectrometer (AVIRIS)," *Remote Sensing of the Environment* **44**, pp. 127–143, 1993.
18. G. Swayze, R. Clark, S. Sutley, and A. Gallagher, "Ground-truthing AVIRIS mineral mapping at cuprite, nevada,," in *Summaries of the Third Annual JPL Airborne Geosciences Workshop*, **1**, pp. 47–49, 1992.
19. R. Ashley and M. Abrams, "Alteration mapping using multispectral images - cuprite mining district, esmeralda county,," open file report 80-367, U.S. Geological Survey, 1980.
20. M. Abrams, R. Ashley, L. Rowan, A. Goetz, and A. Kahle, "Mapping of hydrothermal alteration in the cuprite mining district, nevada, using aircraft scanner images for the spectral region 0.46 to 2.36 $\mu\text{m}$ ," *Geology* **5**, pp. 713–718, 1977.
21. A. Goetz and V. Strivastava, "Mineralogical mapping in the cuprite mining district," in *Proc. of the Airborne Imaging Spectrometer Data Analysis Workshop, JPL Publication 85-41*, pp. 22–29, 1985.
22. F. Kruse, J. Boardman, and J. Huntington, "Comparison of airborne and satellite hyperspectral data for geologic mapping," in *Proc. of the SPIE Aerospace Conference*, **4725**, pp. 128–139, 2002.