Unsupervised Hyperspectral Signal Subspace Identification

José M. P. Nascimento⋆ , José M. Bioucas-Dias⋆⋆

⋆ Instituto Superior de Engenharia de Lisboa and Instituto de Telecomunicações,
R. Conselheiro Emídio Navarro N 1, DEETC, 1959-007 Lisboa Portugal
Tel.:+351.21.8317180, Fax:+351.21.8317114, E-mail: zen@isel.pt
⋆⋆ Instituto de Telecomunicações and Instituto Superior Técnico, Technical University of Lisbon,
Av. Rovisco Pais, Torre Norte, Piso 10, 1049-001 Lisboa Portugal
Tel.: +351.21.8418466, Fax: +351.21.8418472, E-mail: bioucas@lx.it.pt

***Abstract*— Hyperspectral imaging sensors provide image data containing both spectral and spatial information from the Earth surface. The huge data volumes produced by these sensors put stringent requirements on communications, storage, and processing.**

This paper presents a method, termed *hyperspectral signal subspace identification by minimum error* (HySime), that infer the signal subspace and determines its dimensionality without any prior knowledge. The identification of this subspace enables a correct dimensionality reduction yielding gains in algorithm performance and complexity and in data storage. HySime method is unsupervised and fully-automatic, *i.e.*, it does not depend on any tuning parameters. The effectiveness of the proposed method is illustrated using simulated data based on U.S.G.S. laboratory spectra and real hyperspectral data collected by the AVIRIS sensor over Cuprite, Nevada.

I. INTRODUCTION

Hyperspectral imaging sensors collect two dimensional spatial images from the Earth's surface over many contiguous bands of high spectral resolution covering the visible, near-infrared, and shortwave infrared (wavelengths between $0.3\mu m$ and $2.5\mu m$), in hundreds of narrow (on the order of $10nm$) contiguous spectral bands [1, 2]. For example, AVIRIS collects a 512 (along track) $\times$ 614 (across track) $\times$ 224 (bands) $\times$ 12 (bits) data cube, corresponding to more than 800 Mbits [3]. Such huge data volumes put stringent requirements in what concerns communications, storage, and processing.

Each pixel of an hyperspectral image can be represented as a vector in the Euclidian space $\mathbb{R}^L$, where L is the number of bands and each channel is assigned to one axis of space. Under the linear mixing scenario, the spectral vectors are a linear combination of the so-called endmember signatures. The number of endmembers present in a given scene is, very often, much less than the number of bands L. Therefore, hyperspectral vectors lie in a low dimensional linear subspace. The identification of this subspace enables the representation spectral vectors in a low dimensional subspace, thus yielding gains in computational time and complexity and in data storage.

This work was supported by the Fundação para a Ciência e Tecnologia, under the projects PDCTE/CPS/49967/2003 and POSC/EEA-CPS/61271/2004 and grant SFRH/BPD/39475/2007.

Principal component analysis (PCA)[4] and maximum noise fraction (MNF)[5] are two techniques often used to reduce the dimensionality of hyperspectral data. The first technique seeks the projection that best represents data in the least square sense, whereas the second seeks the projection that optimizes the ratio of noise power to signal power. In addition, MNF method needs to estimate the noise covariance. Minimum description length (MDL) [6, 7] and Akaike information criterion (AIC) [8] have also been used to infer the hyperspectral signal subspace.

Harsanyi, Farrand, and Chang [9] developed a Neyman-Pearson detection theory-based thresholding method (HFC) to determine the number of spectral endmembers in hyperspectral data, referred to in [10] as *virtual dimensionality* (VD). The HFC method uses the eigenvalues to measure signal energies in the detection model. A modified version, termed noise-whitened HFC (NWHFC), includes a noise-whitening step [10].

This paper presents a minimum mean squared error based approach to determine the signal subspace in hyperspectral imagery. The method, termed *hyperspectral signal subspace identification by minimum error* (HySime)[11], starts by estimating the signal and the noise correlation matrices using multiple regression. A subset of eigenvectors of the signal correlation matrix is then used to represent the signal subspace. This subspace is inferred by minimizing the sum of the projection error power with the noise power, which are, respectively, decreasing and increasing functions of the subspace dimension. Therefore, if the subspace dimension is overestimated the noise power term is dominant, whereas if the subspace dimension is underestimated the projection error power term is the dominant. The overall scheme is computationally efficient, unsupervised, and fully-automatic in the sense that it does not depend on any tuning parameters.

The remainder of the paper is structured as follows. Section II describes the fundamentals of the HySime method. Sections III and IV evaluate the proposed algorithm using simulated and real data, respectively. Section V ends the paper by presenting some concluding remarks.

II. HySime Method Description

Let us assume an hyperspectral image where each pixel $\mathbf{y}$ can be represented as a spectral vector in $\mathbb{R}^L$ (L is the number of bands), i.e.,

$$\mathbf{y} = \mathbf{x} + \mathbf{n}, \tag{1}$$

where $\mathbf{x}$ and $\mathbf{n}$ are L-dimensional vectors standing for signal and additive noise, respectively.

Assume also that $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \widehat{\mathbf{R}}_n)$, *i.e.*, the noise is zero-mean Gaussian distributed with covariance matrix $\widehat{\mathbf{R}}_n$. A computational efficient method based on the multiple regression theory [12] is used to estimate the noise as shown in Ref. [11].

Let the eigen-decomposition of the signal sample correlation matrix $\widehat{\mathbf{R}}_x = [\widehat{\mathbf{x}}_1, \ldots, \widehat{\mathbf{x}}_N]\,[\widehat{\mathbf{x}}_1, \ldots, \widehat{\mathbf{x}}_N]^T/N$ be written as

$$\widehat{\mathbf{R}}_x = \mathbf{E}\boldsymbol{\Sigma}\mathbf{E}^T, \tag{2}$$

where $\mathbf{E} \equiv [\mathbf{e}_1, \ldots, \mathbf{e}_L]$ is a matrix with the eigenvectors of $\widehat{\mathbf{R}}_x$. Given a permutation $\pi = \{i_1, \ldots, i_L\}$ of indices $i = 1, \ldots, L$, let us decompose the space $\mathbb{R}^L$ into two orthogonal subspaces: the k-dimensional subspace $\langle \mathbf{E}_k \rangle$ spanned by $\mathbf{E}_k \equiv [\mathbf{e}_{i_1}, \ldots, \mathbf{e}_{i_k}]$ and $\langle \mathbf{E}_k \rangle^\perp$ spanned by $\mathbf{E}_k^\perp \equiv [\mathbf{e}_{i_{k+1}}, \ldots, \mathbf{e}_{i_L}]$, *i.e.*, the orthogonal complement of subspace $\mathbf{E}_k$.

Let $\mathbf{U}_k = \mathbf{E}_k\mathbf{E}_k^T$ be the projection matrix onto $\langle \mathbf{E}_k \rangle$ and $\widehat{\mathbf{x}}_k \equiv \mathbf{U}_k\mathbf{y}$ be the projection of the observed spectral vector $\mathbf{y}$ onto the subspace $\langle \mathbf{E}_k \rangle$. The first and the second-order moments of $\widehat{\mathbf{x}}_k$ given $\mathbf{x}$ are[1]

$$\begin{aligned}
\mathbb{E}[\widehat{\mathbf{x}}_k|\mathbf{x}] &= \mathbf{U}_k\mathbb{E}[\mathbf{y}|\mathbf{x}] \\
&= \mathbf{U}_k\mathbb{E}[\mathbf{x}+\mathbf{n}|\mathbf{x}] \\
&= \mathbf{U}_k\mathbf{x} \\
&\equiv \mathbf{x}_k,
\end{aligned} \tag{3}$$

$$\begin{aligned}
\mathbb{E}\left[(\widehat{\mathbf{x}}_k - \mathbf{x}_k)(\widehat{\mathbf{x}}_k - \mathbf{x}_k)^T|\mathbf{x}\right] &= \mathbb{E}[(\mathbf{U}_k\mathbf{y} - \mathbf{U}_k\mathbf{x}) \\
&\quad \times(\mathbf{U}_k\mathbf{y} - \mathbf{U}_k\mathbf{x})^T|\mathbf{x}] \\
&= \mathbb{E}\left[(\mathbf{U}_k\mathbf{n}\mathbf{n}^T\mathbf{U}_k^T)|\mathbf{x}\right] \\
&= \mathbf{U}_k\widehat{\mathbf{R}}_n\mathbf{U}_k^T.
\end{aligned} \tag{4}$$

The mean squared error between $\mathbf{x}$ and $\widehat{\mathbf{x}}_k$ is

$$\begin{aligned}
\mathrm{mse}(k|\mathbf{x}) &= \mathbb{E}\{(\mathbf{x} - \widehat{\mathbf{x}}_k)^T(\mathbf{x} - \widehat{\mathbf{x}}_k)|\mathbf{x}\} \\
&= \mathbb{E}\{(\underbrace{\mathbf{x} - \mathbf{x}_k}_{\mathbf{b}_k} - \mathbf{U}_k\mathbf{n})^T(\underbrace{\mathbf{x} - \mathbf{x}_k}_{\mathbf{b}_k} - \mathbf{U}_k\mathbf{n})|\mathbf{x}\} \\
&= \mathbf{b}_k^T\mathbf{b}_k + \mathrm{tr}(\mathbf{U}_k\widehat{\mathbf{R}}_n\mathbf{U}_k^T),
\end{aligned} \tag{5}$$

Computing the mean of expression (5) with respect to $\mathbf{x}$, noting that $\mathbf{b}_k = \mathbf{x} - \mathbf{x}_k = \mathbf{U}_k^\perp\mathbf{x}$, and using the properties $\mathbf{U} = \mathbf{U}^T$, $\mathbf{U}^2 = \mathbf{U}$, and $\mathbf{U}^\perp = \mathbf{I} - \mathbf{U}$ of the projection matrices, we get

$$\begin{aligned}
\mathrm{mse}(k) &= \mathbb{E}\{(\mathbf{U}_k^\perp\mathbf{x})^T(\mathbf{U}_k^\perp\mathbf{x})\} + \mathrm{tr}(\mathbf{U}_k\widehat{\mathbf{R}}_n\mathbf{U}_k^T) \\
&= \mathrm{tr}(\mathbf{U}_k^\perp\mathbf{R}_x) + \mathrm{tr}(\mathbf{U}_k\widehat{\mathbf{R}}_n) \\
&= \mathrm{tr}(\mathbf{U}_k^\perp\mathbf{R}_y) + 2\,\mathrm{tr}(\mathbf{U}_k\widehat{\mathbf{R}}_n) + c,
\end{aligned} \tag{6}$$

[1] $\mathbb{E}\{\cdot\}$ denotes the expectation operator.

where c is an irrelevant constant. The criterion to estimate the signal subspace, let us call it X, is the minimization of $\mathrm{mse}(k)$ given by expression (6), with respect to all the permutations $\pi = \{i_1, \ldots, i_L\}$ of size L and to k, with the correlation matrix $\mathbf{R}_y$ replaced with the sample correlation matrix $\widehat{\mathbf{R}}_y = [\mathbf{y}_1, \ldots, \mathbf{y}_N]\,[\mathbf{y}_1, \ldots, \mathbf{y}_N]^T/N$; *i.e.*,

$$\begin{aligned}
\widehat{X} &= \left\langle \left[\mathbf{e}_{\widehat{i}_1}, \ldots, \mathbf{e}_{\widehat{i}_{\widehat{k}}}\right] \right\rangle && (7) \\
(\widehat{k}, \widehat{\pi}) &= \arg\min_{k,\pi} \left\{ \mathrm{tr}(\mathbf{U}_k^\perp\widehat{\mathbf{R}}_y) + 2\,\mathrm{tr}(\mathbf{U}_k\widehat{\mathbf{R}}_n) \right\}, && (8)
\end{aligned}$$

where the dependence on the permutation π is through $\mathbf{U}_k = \mathbf{E}_k\mathbf{E}_k^T$. For a given permutation π, each term of expression (8) has a clear meaning: the first term accounts for the projection error power and is a decreasing function of k; the second term accounts for the noise power and is an increasing function of k.

By exploiting, again, the fact that the $\mathbf{U}_k$ is a projection matrix and that $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$, for $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{L\times L}$, the minimization (8) can be rewritten as

$$(\widehat{k}, \widehat{\pi}) = \arg\min_{k,\pi} \left\{ c + \sum_{j=1}^{k} \delta_{i_j} \right\}, \tag{9}$$

where c is an irrelevant constant and $\delta_{i_j} = \mathbf{e}_{i_j}^T\widehat{\mathbf{R}}_y\mathbf{e}_{i_j} + \mathbf{e}_{i_j}^T\widehat{\mathbf{R}}_n\mathbf{e}_{i_j}$. Based the right hand side of (9), it follows that the corresponding minimization is achieved simply by including all the negative terms δ_i, for $i = 1, \ldots, L$, and only these, in the sum.

III. Experiments

In this section, we apply the proposed HySime algorithm to simulated scenes and compare it with the NWHFC eigen-based Neyman-Pearson detector [10]. As concluded in [10], these algorithms are the state-of-the-art in hyperspectral signal subspace identification, outperforming the information theoretical criteria approaches; namely, the minimum description length (MDL) [6, 7] and the Akaike information criterion (AIC) [8].

A simulated hyperspectral image composed of 10^4 spectral vectors, each one following expression (1), is generated. Assume that signal vectors are in an p-dimensional subspace, *i.e.*,

$$\mathbf{x} = \mathbf{M}\mathbf{s},$$

where $\mathbf{M} \equiv [\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_p]$ is a full-rank $L \times p$ matrix ($\mathbf{m}_i$ denotes the ith endmember signature), p is the number of endmembers present in the covered area, and $\mathbf{s} = [s_1, s_2, \ldots, s_p]^T$ is the abundance vector containing the fractions of each endmember (notation $(\cdot)^T$ stands for vector transposed).

The spectral signatures are selected from the USGS digital spectral library [13]. The abundance fractions are generated according to a Dirichlet distribution given by

$$D(s_1, \ldots, s_p|\theta_1, \ldots, \theta_p) = \frac{\Gamma(\sum_{j=1}^{p}\theta_j)}{\prod_{j=1}^{p}\Gamma(\theta_j)} \prod_{j=1}^{p} s_j^{\theta_j - 1}, \tag{10}$$

where $\{s_1, \ldots, s_p\}$ are subject to nonnegativity and constant sum constraints, *i.e.*, $\{\mathbf{s} \in \mathbb{R}^p : s_j \geq 0, \sum_{j=1}^{p} s_j = 1\}$.

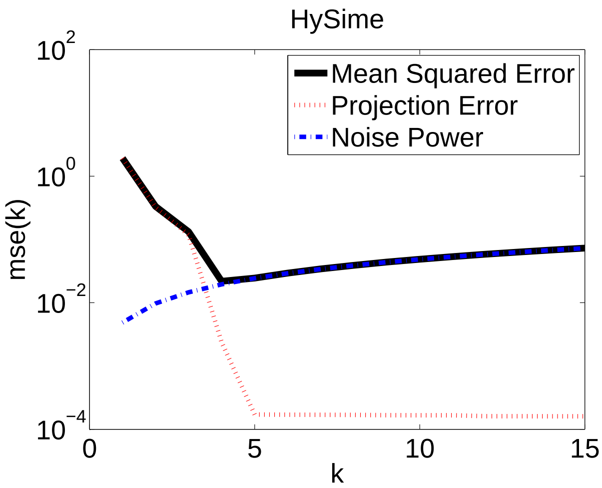

Fig. 1. Mean squared error versus k, with $SNR = 20\,\text{dB}$ and $p = 5$.

This density, besides enforcing positivity and full additivity constraints, displays a wide range of shapes, depending on the parameters of the distribution. On the other hand, as noted in [14], the Dirichlet density is suited to model fractions.

Consider that the noise correlation matrix is $\mathbf{R}_n = \text{diag}(\sigma_1^2, \ldots, \sigma_L^2)$ and that the diagonal elements follow a Gaussian shape centered at the band $L/2$, *i.e.*,

$$\sigma_i^2 = \sigma^2 \frac{e^{-\frac{(i-L/2)^2}{(2\eta^2)}}}{\sum_{j=1}^{L} e^{-\frac{(j-L/2)^2}{(2\eta^2)}}}, \tag{11}$$

for $i = 1, \ldots, L$. Parameter η plays the role of variance in the Gaussian shape ($\eta \to \infty$ corresponds to white noise; $\eta \to 0$ corresponds to *one-band* noise). Parameter σ^2 controls the total noise power.

The method is evaluated with respect to the number of endmembers p, to the spectral noise shape (white and nonwhite), and to the SNR defined by

$$\text{SNR} \equiv 10 \log_{10} \frac{\mathbb{E}\left[\mathbf{x}^T\mathbf{x}\right]}{\mathbb{E}\left[\mathbf{n}^T\mathbf{n}\right]}. \tag{12}$$

Fig. 1 shows the evolution of the mean squared error for HySime algorithm as a function of the parameter k, for $p = 5$, SNR$=$ 20 dB and $\eta = 0$. The minimum of the mean squared error occurs at $k = 5$, which is exactly the number of endmembers present in the image. As expected, the projection error power and of noise power display decreasing and increasing behaviors, respectively, as a function of the subspace dimension k.

Table I presents the signal subspace order estimates yielded by HySime algorithm and the virtual dimensionality (VD) determined by the NWHFC algorithm [10], as a function of the SNR, of the number of endmembers, p, and of the noise shape.

NWHFC algorithm is basically the HFC one [9] preceded by a noise-whitening step, based on the estimated noise correlation matrix. In implementing this step, we got poor results in very high SNRs and colored noise scenarios. This is basically because the noise estimation step in NWHFC needs to invert the noise correlation matrix, which gives inaccurate results when the noise power is small. For this reason, we have used both the true and estimated noise correlation matrices. The results based on the true correlation matrix are in brackets. We stress that, for the setting of this experiment, HySime method yields the same results, whether using the estimated or the true noise correlation matrices.

Another central issue of NWHFC algorithm is the false-alarm probability P_f it is parameterized with. This probability is used in a series of Neyman-Pearson tests, each one designed to detect a different orthogonal signal subspace direction. It is necessary, therefore, to specify the false-alarm probability P_f of the tests. Based on the hints given in [10] and in our own results, we choose $P_f \in \{10^{-3}, 10^{-4}, 10^{-5}\}$.

The figures shown in Table I, based on 50 Monte Carlo runs, have the following behavior:

i) HySime and NWHFC algorithms parameterized with $P_f = 10^{-3}$ display similar performances at low subspace dimension, say $p \leq 5$, and white noise. This is also true for colored noise and NWHFC working with known noise covariance matrix. However, if the noise statistics is unknown, NWHFC performs much worse than HySime;

ii) HySime performs better that NWHFC for high space dimensions, say $p > 5$.

We conclude, therefore, that HySime algorithm yields systematically equal or better results than NWHFC algorithm. Another advantage of HySime approach is that it does not depend on any tunable parameter.

IV. Experiments with Real Hyperspectral Data

In this section, the proposed method, HySime, is applied to a subimage (350×350 pixels and 224 bands) of the Cuprite data set acquired by the AVIRIS sensor on June 19, 1997 (see Fig. 2). The AVIRIS instrument covers the spectral region from $0.41\mu m$ to $2.45\mu m$ in 224 bands with a $10nm$ band width. Flying at an altitude of $20km$, it has an IFOV of $20m$ and views a swath over $10km$ wide. This site has been extensively used for remote sensing experiments over the past years and its geology was previously mapped in detail [15]. The HySime method when applied to the AVIRIS data set, estimates a subspace dimension of $\widehat{k} = 20$. According to the ground truth presented in [15], there are 18 materials in this area. This difference is due to a) the presence of rare pixels not accounted for in [15] and b) spectral variability.

The VD estimated by the NWHFC method [10] ($P_f = 10^{-3}$) on the same data set yields $\widehat{k} = 23$. A lower value of P_f would lead to a higher number of endmembers. According to the ground truth presented in [15], the estimates yielded by HySime and NWHFC methods overestimate the number of endmembers in the Cuprite data set. The mainly reason for this difference, as we have explained, is the presence of rare pixels present in the data set not accounted for in [15].

Table I
Signal subspace dimension $\widehat{k}$, based on 50 Monte Carlo runs, as function of SNR, p, and η (noise shape). Figures in brackets were computed based on the true noise statistics.

		White Noise ($\eta = 0$)			
SNR	Method	$p = 3$	$p = 5$	$p = 10$	$p = 15$
	HySime	**3**	**5**	**10**	**15**
	NWHFC				
50 dB	($P_f = 10^{-3}$)	3 (3)	5 (5)	7 (7)	10 (11)
	($P_f = 10^{-4}$)	3 (3)	5 (5)	7 (7)	8 (8)
	($P_f = 10^{-5}$)	3 (3)	4 (4)	7 (6)	8 (8)
	HySime	**3**	**5**	**10**	**15**
	NWHFC				
35 dB	($P_f = 10^{-3}$)	3 (3)	4 (4)	7 (7)	9 (9)
	($P_f = 10^{-4}$)	3 (3)	4 (4)	7 (6)	8 (8)
	($P_f = 10^{-5}$)	3 (3)	4 (4)	6 (6)	8 (8)
	HySime	**3**	**5**	**10**	**14**
	NWHFC				
25 dB	($P_f = 10^{-3}$)	3 (3)	5 (5)	6 (6)	9 (8)
	($P_f = 10^{-4}$)	3 (3)	5 (5)	6 (6)	7 (7)
	($P_f = 10^{-5}$)	3 (3)	4 (4)	5 (5)	7 (7)
	HySime	**3**	**5**	**8**	**12**
	NWHFC				
15 dB	($P_f = 10^{-3}$)	3 (3)	5 (5)	5 (4)	5 (5)
	($P_f = 10^{-4}$)	3 (3)	4 (4)	3 (3)	3 (2)
	($P_f = 10^{-5}$)	3 (3)	4 (4)	3 (3)	2 (2)

		Gaussian shaped noise ($\eta = 1/18$)			
SNR	Method	$p = 3$	$p = 5$	$p = 10$	$p = 15$
	HySime	**3**	**5**	**10**	**15**
	NWHFC				
50 dB	($P_f = 10^{-3}$)	59 (3)	41 (5)	61 (10)	45 (10)
	($P_f = 10^{-4}$)	48 (3)	33 (5)	54 (10)	34 (10)
	($P_f = 10^{-5}$)	43 (3)	28 (5)	41 (9)	27 (10)
	HySime	**3**	**5**	**10**	**15**
	NWHFC				
35 dB	($P_f = 10^{-3}$)	9 (3)	10 (5)	12 (10)	10 (10)
	($P_f = 10^{-4}$)	9 (3)	9 (5)	11 (10)	8 (10)
	($P_f = 10^{-5}$)	7 (3)	7 (5)	10 (9)	8 (10)
	HySime	**3**	**5**	**10**	**15**
	NWHFC				
25 dB	($P_f = 10^{-3}$)	4 (3)	5 (5)	11 (10)	9 (11)
	($P_f = 10^{-4}$)	4 (3)	5 (5)	11 (10)	9 (10)
	($P_f = 10^{-5}$)	4 (3)	5 (5)	11 (9)	8 (10)
	HySime	**3**	**5**	**8**	**12**
	NWHFC				
15 dB	($P_f = 10^{-3}$)	4 (3)	5 (5)	11 (10)	10 (10)
	($P_f = 10^{-4}$)	4 (3)	5 (5)	11 (10)	8 (10)
	($P_f = 10^{-5}$)	4 (3)	5 (5)	11 (9)	8 (10)

V. Conclusions

The huge volumes and rates of data generated by hyperspectral sensors demand expensive processors with very high performance and memory capacities. Dimensionality reduction is, therefore, a relevant first step in the hyperspectral data processing chain. This paper introduces the HySime algorithm, a new approach to estimate the signal subspace in hyperspectral imagery. HySime algorithm estimates the signal and the noise correlation matrices and then selects the subset of eigenvalues that best represents the signal subspace in the minimum mean squared error sense. A set of experiments with simulated and real data leads to the conclusion that the HySime algorithm is an effective and useful tool, yielding comparable or better results than the state-of-the-art algorithms.

Fig. 2. False-color subimage of AVIRIS Cuprite Nevada data set.

References

[1] B. Hapke, *Theory of Reflectance and Emmittance Spectroscopy*. Cambridge, U. K.: Cambridge Univ. Press, 1993.
[2] T. M. Lillesand, R. W. Kiefer, and J. W. Chipman, *Remote Sensing and Image Interpretation*, 5th ed. John Wiley & Sons, Inc., 2004.
[3] J. P. Kerekes and J. E. Baum, "Spectral imaging system analytical model for subpixel object detection," *IEEE Trans. Geosci. Remote Sensing*, vol. 40, no. 5, pp. 1088–1101, 2002.
[4] I. T. Jolliffe, *Principal Component Analysis*. New York: Spriger Verlag, 1986.
[5] 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*, vol. 26, no. 1, pp. 65–74, 1988.
[6] G. Schwarz, "Estimating the dimension of a model," *Annals of Statistics*, vol. 6, pp. 461–464, 1978.
[7] J. Rissanen, "Modeling by shortest data description," *Automatica*, vol. 14, pp. 465–471, 1978.
[8] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Automat. Contr.*, vol. 19, no. 6, pp. 716–723, 1974.
[9] J. Harsanyi, W. Farrand, and C.-I. Chang, "Determining the number and identity of spectral endmembers: An integrated approach using neyman-pearson eigenthresholding and iterative constrained rms error minimization," in *Proc. 9th Thematic Conf. Geologic Remote Sensing*, 1993.
[10] C.-I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sensing*, vol. 42, no. 3, pp. 608–619, 2004.
[11] J. M. Bioucas-Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sensing*, vol. 46, no. 8, pp. 2435–2445, 2008.
[12] R. Roger and J. Arnold, "Reliably estimating the noise in aviris hyperspectral imagers," *Int. J. Rem. Sens.*, vol. 17, no. 10, pp. 1951–1962, 1996.
[13] 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 μm," U.S. Geological Survey, Open File Report 93-592, 1993.
[14] T. Minka, "Estimating a dirichlet distribution," M.I.T., Tech. Rep., 2000.
[15] 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*, 1992, pp. 47–49.