1998 International Computer Symposium
Workshop on Image Processing and Character Recognition
December 17-19, 1998, N.C.K.U., Tainan, Taiwan, R.O.C.

A Partition Noise-Adjusted Principal Components Analysis for Determining Data Dimensionality in Hyperspectral Imagery

*Te-Ming Tu, Hsuen-Chyun Shyu, Ching-Hai Lee, and Chien-Ping Chang**

Department of Electrical Engineering,
National Chung Cheng Institute of Technology,
Taoyuan, Taiwan 33509, Republic of China
Emai:tum@cc04.ccit.edu.tw
*Chung-Shan Institute of Science & Technology

ABSTRACT

In this paper, we apply noise-adjusted principal component analysis (*NAPCA*) to a partitioned data space to resolve the inaccuracy of the noise estimation and properly estimate the data dimensionality. This approach is referred to herein as *PNAPCA*. In contrast to the *PCA*-based approaches which consider interrelationships *within* a set of variables, *PNAPCA* focuses on the relationship *between* two distinct subspaces which are partitioned from the data space of the original image by a simultaneous transformation. This partitioning causes the gap between the group of eigenvalues for signal plus noise and noise only to become larger than all other PCA-based approaches. The number of endmembers can then be determined by a designed union-intersection margin testing (*UIMT*).
Keywords: Noise-adjusted principal components analysis (*NAPCA*), partitioned noise-adjusted principal components analysis (*PNAPCA*)

1. INTRODUCTION

Multispectral/hyperspectral imaging spectrometry in Earth remote sensing applications largely focuses on determining the identities and abundances of materials in a geographic area of interest. Reaching this goal largely rests on identifying the number of endmembers in the image, or equivalently, determining the image's intrinsic dimensionality.

All remote sensing data are accompanied by noises. Several methods have been proposed to address the intrinsic dimensionality problem under the assumption of Gaussion white noise. A previous investigation, like principal components analysis (*PCA*) [1]; Wax and Kailath [2] proposed an approach based on Akaike information criterion (*AIC*) [3] and Schwartz-Rissanen's minimum description length (*MDL*) [4] criterion. However, in their approach the environmental noise is often unknown or indeterminable in practice. If the knowledge or an estimation of the noise covariance is available, the minimum noise fraction (*MNF*) transformation proposed by Green et. al. [5] effectively solves this problem. A later investigation[6] further interpreted this transform as the noise-adjusted principal component analysis (*NAPCA*) with a rapid version proposed in [7]. *NAPCA* is largely limited in that its noise whitening process requires complete knowledge of the noise structure for the processed data.

In this work, we perform partitioned noise-adjusted principal components analysis (referred to herein as *PNAPCA*) to resolve the inaccurate estimation of noise in *NAPCA*. Also presented herein is a more effective means of resolving the inherent dimensionality problem. In contrast to the *PCA*-based approaches which consider interrelationships *within* a set of variables, *PNAPCA* focuses on the relationship *between* two distinct subspaces which are partitioned from the original data space by a simultaneous transform. According to an error analysis of *NAPCA*, the diagonal elements of the covariance matrix in *PNAPCA* can be partitioned into two distinct groups of eigenvalues: one associated with large covariance for signal plus noise, and its complementary with less or equal unity covariance for noise only. In doing so, the number of endmembers can be determined by performing a designed union-intersection margin test (*UIMT*) and counting the number of covariance values larger than unity. The performance of *PNAPCA* approach is then

evaluated by real imaging spectrometer data sets collected by the Airborne Visible Infrared Imaging Spectrometer (AVIRIS). Results presented herein demonstrate that this method can aptly solve the intrinsic dimensionality problem for hyperspectral images.

2. Problem Statement and Signal Model

Linear mixture model for hyperspectral images

Linear spectral mixture model is extensively employed in remotely sensed imagery to determine and quantify multicomponents. Let $\boldsymbol{r}_i$ be an $l\times 1$ column vector denoting the i-th pixel in a hyperspectral image, where l is the number of bands. A linear mixture model for the pixel $\boldsymbol{r}_i$ in a hyperspectral image can be described by [8]

$$\boldsymbol{r}_i = M\boldsymbol{\alpha}_i + \boldsymbol{n}_i \tag{1}$$

Equivalently, (1) can be expressed as a standard signal model:

$$\boldsymbol{r}_i = \boldsymbol{s}_i + \boldsymbol{n}_i. \tag{2}$$

Its true correlation matrix is then defined by

$$\begin{aligned} R &= E[\boldsymbol{r}\boldsymbol{r}^T] \\ &= ME[\boldsymbol{\alpha}\boldsymbol{\alpha}^T]M^T + R_n \\ &= R_s + R_n \end{aligned} \tag{3}$$

where M is an $l\times p$ matrix denoted by $(\boldsymbol{m}_1, \boldsymbol{m}_2, \cdots, \boldsymbol{m}_p)$ and $\boldsymbol{m}_j$ is an $l\times 1$ column vector for the spectral signature of j-th distinct material, p is the number of materials; $\boldsymbol{\alpha}_i$ is a $p\times 1$ column vector given by $(\alpha_1, \alpha_2, \cdots, \alpha_p)^T$ where α_j denotes the fraction of the j-th signature present in $\boldsymbol{r}_i$; $\boldsymbol{n}_i$ is an $l\times 1$ column vector for the combined noise which is assumed to be a wide sense stationary Gaussian process with zero mean and correlation matrix R_n.

Notably, the true correlation matrix $\boldsymbol{R}$ is an $l\times l$ matrix. Meanwhile, the noise correlation matrix R_n is of the full rank l and the signal correlation matrix $R_s = ME[\boldsymbol{\alpha}\boldsymbol{\alpha}^T]M^T$ is of rank p. Therefore, the inherent dimensionality problem attempts to determine the value of p by a given $\boldsymbol{R}$.

3. NAPCA and Its Error Analysis

3.1. NAPCA

The *NAPCA* approach can be regarded as a two stage, cascaded principal component transformation with a diagonalization procedure [9] to achieve the maximum signal-noise-ratio (*MSNR*), i.e. to find a matrix A such that

$$\max_{\mathrm{A}} \frac{\mathrm{A}^T\mathrm{R}\mathrm{A}}{\mathrm{A}^T\mathrm{R}_n\mathrm{A}} = \max_{\mathrm{A}} \frac{\mathrm{A}^T\mathrm{R}_s\mathrm{A}}{\mathrm{A}^T\mathrm{R}_n\mathrm{A}} + I \ (\textit{due to Eq.}(3)). \tag{4a}$$

can be obtained. Equivalently, the above equation means to solve the problem

$$A^T R A = \Lambda \quad \text{and.} \ A^T R_n A = I \tag{4b}$$

To obtain the desired transformation in (4), a whitening process can be designed to simultaneously transform R_n and $\boldsymbol{R}$. Restated,

$$W^T R_n W = I \ \text{ and } \ W^T R W = R_{adj} \tag{5}$$

where $W = \Phi_n \Lambda_n^{-1/2}$ is the whitening matrix, In addition, Λ_n and Φ_n are eigenvalue and eigenvector matrices of R_n, respectively. The adjusted correlation matrix R_{adj} is not a diagonal but a symmetric matrix in general.

Using the eigenvectors of R_{adj}, i.e. Φ_{adj}, as the basis for the second transformation leads to

$$\Phi_{adj}^T I \Phi_{adj} = I \tag{6-a}$$

and

$$\Phi_{adj}^T R_{adj} \Phi_{adj} = \Lambda_{adj} \tag{6-b}$$

Consequently, the desired *NAPCA* transform can be derived by

$$A = \Phi_n \Lambda_n^{-1/2} \Phi_{adj}. \tag{7}$$

The subsequent transformed covariance matrix is then expressed as

$$R_{Y,NAPCA} = A^T R A = diag[\lambda_1, \cdots, \lambda_l] \tag{8}$$

Consider a situation in which the noise covariance matrix R_n is accurately estimated from the data. This allows us to partition the transformed data space into two parts: one consists of eigenvalues larger than one and its complementary of eigenvalues of unity. This observation implies that $\{\lambda_i\}_{i=1}^{p} = \lambda_i' + 1$ and $\{\lambda_i\}_{i=p+1}^{l} = 1$, where λ_i' are the associated

eigenvalues for the signal correlation matrix R_s. Under such a circumstance, the inherent dimensionality of the data can be determined by examining the number of eigenvalues larger than unity.

3.2 Error analysis for NAPCA

To analyze the error incurred while estimating the noise, (4) can be rewritten as follows:

$$R = R_s + R_n \quad (9a)$$
$$= (R_s + R_\varepsilon) + (R_n - R_\varepsilon) \quad (9b)$$
$$= \tilde{R}_s + \tilde{R}_n \quad (9c)$$

where R_ε represents the correlation matrix of the estimated error. Notably, R_ε has the same characteristic as the noise correlation matrix R_n, implying that R_ε is also a wide sense stationary Gaussian process with zero mean and of the full rank l. In addition, $\tilde{R}_s$ and $\tilde{R}_n$ represent the correlation matrix of the estimated signal and noise, respectively. The expression in (9a) differs from that in (9c) in the sense that the former adopts the complete knowledge of R_n. However, the latter has no prerequisite knowledge so that the true R_n is replaced by an estimate $\tilde{R}_n$ resulting from the observation vector r.

Using a statistical model to describe an observation is a conventional practice in the signal processing community. However, the advantage of using *a priori* (*pr*) model is significantly diminished when the number of observed samples is increased. Under this circumstance, the model described by (9c), *a posterior* (*ps*) observation model, begins to show its dominance and tends to replace the *pr* model in (9a). As the observation process proceeds with, the *ps* model eventually takes over the *pr* model. Consequently, algorithms using the *ps* model generally perform better than those based on a *pr* model. Restated, the information given by the *ps* model provides a better understanding than the *pr* model and facilitates the knowledge about the signal. In a real *ps* model application similar to (9c), an incomplete noise statistics $\tilde{R}_n$ is employed rather than the true noise statistics R_n. Due to the properties of the *pr* and *ps* model, as generally believe (9c) performs better than (9a). Moreover, the *NAPC* process in *a ps* model must be differ from that in the *pr* model. The related changes are illustrated as follows:

Step 1. (4b) is replaced by $A^T R A = \Lambda$ and

$$A^T \tilde{R}_n A = I \quad (10)$$

Step 2. (5) is replaced by

$$\tilde{W}^T \tilde{R}_n \tilde{W} = I \text{ and } \tilde{W}^T R \tilde{W} = \tilde{R}_{adj} \quad (11)$$

where $\tilde{W} = \tilde{\Phi}_n \tilde{\Lambda}_n^{-1/2}$, and $\tilde{\Lambda}_n$ and $\tilde{\Phi}_n$ are eigenvalue and eigenvector matrices of $\tilde{R}_n$, respectively. After the whitening process, the correlation matrix R becomes

$$\begin{aligned} \tilde{R}_{adj} &= \tilde{W}^T \tilde{R}_s \tilde{W} + \tilde{W}^T \tilde{R}_n \tilde{W} \\ &= \tilde{W}^T (R_s + R_\varepsilon) \tilde{W} + I \\ &= \hat{R}_{s,adj} + \hat{R}_{\varepsilon,adj} + I \end{aligned} \quad (12)$$

where $\hat{\Lambda}_n$ and $\hat{\Lambda}_\varepsilon$ are the adjusted correlation matrices.

Step 3. Using the eigenvectors of $\tilde{R}_{adj}$, $\tilde{\Phi}_{adj}$, as the basis for the second transformation, i.e. (6b) is replaced by

$$\begin{aligned} &\tilde{\Phi}_{adj}^T \tilde{R}_{adj} \tilde{\Phi}_{adj} \\ &= \tilde{\Lambda}_{adj} \\ &= diag(\tilde{\lambda}_1, \tilde{\lambda}_2, \cdots, \tilde{\lambda}_p, \tilde{\lambda}_{p+1}, \cdots, \tilde{\lambda}_l) \\ &= diag[(\hat{\lambda}_1 + \hat{\varepsilon}_1 + 1), \cdots, (\hat{\lambda}_p + \hat{\varepsilon}_p + 1), (\hat{\varepsilon}_{p+1} + 1), \cdots, (\hat{\varepsilon}_l + 1)] \end{aligned} \quad (13)$$

where, $\hat{\lambda}_i$ and $\hat{\varepsilon}_i$ are the associated eigenvalues for $\hat{R}_{s,adj}$ and $\hat{R}_{\varepsilon,adj}$, respectively, and "1" is a constant energy resulted in the whitening process.

Step 4. The subsequent transformed kernel is then given by

$$A = \tilde{\Phi}_n \tilde{\Lambda}_n^{-1/2} \tilde{\Phi}_{adj} \quad (14)$$

and the real *NAPCA* can be derived by

$$R_{Y,NAPCA} = A^T R A = \tilde{\Lambda}_{adj} \quad (15)$$

In (13), three energies, $\tilde{\lambda}_i$, ε_i, and "1" are independent of each another. Notably, the more incomplete this noise statistics implies larger error eigenvalues. Therefore, our problem largely focuses on accurately finding the inherent dimensionality in a low *SNR* situation where

some materials may have a low probability of occurrence within the scene, implying that they only appear in a small number of pixels or mixed pixels. When the signal energy of materials is smaller than the total noise energy in the entire image, the intrinsic dimensionality is generally underestimated. Restated, the intrinsic dimensionality may be underestimated when some eigenvalues in the estimated signal-subspace are extremely close to an increasing estimated noise variance.

4.PARTITIONED SUBSPACE APPROACHES

In contrast to the *PCA*-based approaches which consider interrelationships *within* a set of variables, the notion presented herein focuses on the relationship *between* two distinct subspaces which are partitioned from the original data space by a simultaneous transform. The underlying motivation of this notion is that the intrinsic dimensionality is invariant to the number of processed bands if this number markedly exceeds that of endmembers. Hence, although the gap between the group of eigenvalues for signals and noises is difficult to detect in the entire data space, such a gap provide valuable insight into two smaller, partitioned, and distinct subspaces. Moreover, the intrinsic dimensionality of these two subspaces is identical.

To investigate the degree of association between two distinct subspaces, the noise-adjusted correlation matrix $\tilde{R}_{adj}$ given in (12) should be partitioned as

$$\tilde{R}_{adj} = \begin{bmatrix} R_1 & R_{12} \\ R_{21} & R_2 \end{bmatrix} = \begin{bmatrix} \hat{R}_{s,1} + \hat{\Lambda}_{\varepsilon,1} + I & \hat{R}_{s,12} \\ \hat{R}_{s,21} & \hat{R}_{s,2} + \hat{\Lambda}_{\varepsilon,2} + I \end{bmatrix} \quad (16)$$

where R_i is an $ki \times ki$ sub-matrix with $i = 1, 2$ and $k1+k2=l$, and l is the dimension of R. Without a loss of generality, we assume that $p << k1 \le k2$. For an extreme case where the noise matrix is band independent, i.e. $R_n = \begin{bmatrix} R_{n,1} & \underline{0} \\ \underline{0} & R_{n,2} \end{bmatrix}$, the error correlation matrices $\hat{R}_{\varepsilon,adj,12}$ and $\hat{R}_{\varepsilon,adj,21}$ equal zero. In such case R_{12} contains only signal information of $\hat{R}_{s,adj,12}$, and the estimation of the rank of $\hat{R}_{s,adj,12}$ allows us to determine the number of endmembers within the scene. However, since the actual value of R and R_n are unknown in practice, the average of the outer products of the sample data is used as an estimate of R. That is,

$$\bar{R} = \frac{1}{N}\sum_{i=1}^{N} r_i r_i^T \quad (17)$$

where $\{r_i\}_{i=1}^{N}$ represents the i-th, $1 \le i \le N$, observation vector. The rank of the estimated submatrix $\hat{R}_{s,12}$ is obviously no longer p, but $k1 (\le k2)$ due to the inaccuracy of the estimation.

When the noise matrix is band independent the true rank p through matrix decompositions can be estimated in several ways, e.g. the singular value decomposition (*SVD*) can be used to estimate the signal subspace and its orthogonal complement. However, *SVD* cannot give a correct estimation for unknown correlation matrices. To improve this drawback, a partitioned version of the *NAPCA* (*PNAPCA*) transformation is developed to estimate the true rank p from R_{12} by utilizing the error analysis discussed in discussed in Section3.2.

Since R_1 and R_2 are symmetric and positive definite matrices, each of these matrices can be diagonalized by a unitary similarity transformation

$$\Phi_{R1}^T R_1 \Phi_{R1} = \Lambda_{R1} \quad (18\text{-a})$$

and $$\Phi_{R2}^T R_2 \Phi_{R2} = \Lambda_{R2} \quad (18\text{-b})$$

where $\Phi_{R1} = [p_1, p_2, \ldots, p_k]$ and $\Phi_{R2} = [q_1, q_2, \ldots, q_k]$ are unitary matrices whose columns are the eigenvectors of R_1 and R_2, respectively. The diagonal of matrices Λ_{R1} and Λ_{R2} are the respective eigenvalues of matrices R_1 and R_2 arranged in a descending order.

Next, an $l \times l$ unitary transformation matrix Q is constructed to rotate the noise-adjusted correlation matrix $\tilde{R}_{adj}$. This unitary matrix is defined by

$$Q = \begin{bmatrix} \Phi_{R1} & \underline{0} \\ \underline{0} & \Phi_{R2} \end{bmatrix} \quad (19)$$

Therefore, the *PNAPCA* transformation is defined by

$$R_{Y,PNAPCA} = Q^T \tilde{R}_{adj} Q = \begin{bmatrix} Q_{R1}^T R_1 Q_{R1} & Q_{R1}^T R_{12} Q_{R2} \\ Q_{R2}^T R_{21} Q_{R1} & Q_{R2}^T R_2 Q_{R2} \end{bmatrix} = \begin{bmatrix} \Lambda_{R1} & \Sigma_{12} \\ \Sigma_{21} & \Lambda_{R2} \end{bmatrix} \quad (20)$$

with

$$\Lambda_{R1} = diag[(\dot{\lambda}_1 + \dot{\varepsilon}_1 + 1), \cdots, (\dot{\lambda}_p + \dot{\varepsilon}_p + 1), (\dot{\varepsilon}_{p+1} + 1), \cdots, (\dot{\varepsilon}_{k1} + 1)] \quad (21a)$$

$$\Lambda_{R2} = diag[(\ddot{\lambda}_1 + \ddot{\varepsilon}_1 + 1), \cdots, (\ddot{\lambda}_p + \ddot{\varepsilon}_p + 1), (\ddot{\varepsilon}_{p+1} + 1), \cdots, (\ddot{\varepsilon}_{k2} + 1)] \quad (21b)$$

$$\Sigma_{12} = \begin{bmatrix} \gamma_{1,1} & \gamma_{1,2} & \cdots & & \gamma_{1,k2} \\ \gamma_{2,1} & \gamma_{2,2} & \cdots & & \gamma_{2,k2} \\ & & \vdots & & \\ \gamma_{p,1} & \gamma_{p,2} & \cdots \gamma_{p,p} \cdots & & \gamma_{p,k2} \\ \gamma_{p+1,1} & \gamma_{p+1,2} & \cdots & \gamma_{p+1,p+1} \cdots & \gamma_{p+1,k2} \\ & & \vdots & & \\ \gamma_{k1,1} & \gamma_{k1,2} & \cdots & & \gamma_{k1,k2} \end{bmatrix} \quad (21c)$$

and

$$\Sigma_{21} = \Sigma_{12}^T. \quad (21d)$$

In (20) and (21), Λ_{R1} and Λ_{R2} are transformed energy matrices in which $\{\dot{\lambda}_i\}_{i=1}^{p}$ and $\{\ddot{\lambda}_i\}_{i=1}^{p}$ represent the rearrangement of the adjusted signal energy, $\hat{R}_{\varepsilon,adj,1}$ and $\hat{R}_{\varepsilon,adj,2}$, respectively. In addition, $\{\dot{\varepsilon}_i\}_{i=1}^{k1}$ and $\{\ddot{\varepsilon}_i\}_{i=1}^{k2}$ are the rearrangement of the adjusted error energy in $\hat{R}_{\varepsilon,adj,1}$ and $\hat{R}_{\varepsilon,adj,2}$, respectively. Owing to R_{12}, Σ_{12} is the transformed covariance matrix whose diagon alelements $\gamma_{i,i} = p_i^T R_{12} q_i$ for $i = 1,2,\ldots,k$ represent the cross-covariances between these two transformed variances Λ_{R1} and Λ_{R2}. However, some of these variances may be too small to have any statistical significance. For this problem, a conventionally used approach is to test the null hypothesis to find ($k1$-p) smaller cross-covariances of zeroes. That is,

$$H_P: \gamma_p > \gamma_{p+1} = \cdots = \gamma_{k1} = 0. \quad (22)$$

The test is performed in the following manner: Test the null hypothesis $H_j, j = 1,2,\cdots,k1$. If $H_j, j = 1,2,\cdots,k1$ passes the test but H_{p+1} does not. Then, the estimated intrinsic dimensionality is p.

In this work, a simple union-intersection margin test (*UIMT*) is developed for the *PNAPCA* transformation to replace the general null test. Using the Scharwtz inequality and the relationships in (21a) and (21b) leads to

$$[\mathrm{cov}(X,Y)]^2 \le \mathrm{var}(X) \cdot \mathrm{var}(Y) \quad (23)$$

so that

$$\gamma_{i,j}^2 \le (\dot{\lambda}_i + \dot{\varepsilon}_i + 1) \cdot (\ddot{\lambda}_i + \ddot{\varepsilon}_i + 1) \quad i = 1,2,\ldots,p \quad (24a)$$

and

$$\gamma_{i,i}^2 \le (\dot{\varepsilon}_i + 1) \cdot (\ddot{\varepsilon}_i + 1) i = p+1, p+2, \ldots, k1. \quad (24b)$$

According to the results of error analysis and the fact that three energies $\tilde{\lambda}_i$, $\hat{\varepsilon}_i$, and "1" are independent of each other, the set theory is utilized to derive the threshold of the new hypothesis test. From the perspective of the sets, the covariance between two variables can be interpreted as the intersection of these two sets. Meanwhile, the sum of the energies for two distinct variables can be interpreted as the union of these two sets. Hence, (24) can be rewritten as a form of set operation:

$$\begin{aligned} \gamma_{i,i}^2 &\le (\dot{\lambda}_i \cup \dot{\varepsilon}_i \cup 1) \cap (\ddot{\lambda}_i \cup \ddot{\varepsilon}_i \cup 1) \\ &= (\dot{\lambda}_i \cap \ddot{\lambda}_i) \cup (\dot{\lambda}_i \cap \ddot{\varepsilon}) \cup (\dot{\lambda}_i \cap 1) \\ &\quad \cup (\dot{\varepsilon}_i \cap \ddot{\lambda}_i) \cup (\dot{\varepsilon}_i \cap \ddot{\varepsilon}) \cup (\dot{\varepsilon}_i \cap 1) \\ &\quad \cup (1 \cap \ddot{\lambda}_i) \cup (1 \cap \ddot{\varepsilon}) \cup (1 \cap 1) \\ &= (\dot{\lambda}_i \cap \ddot{\lambda}_i) \cup 1 \\ &= (\dot{\lambda}_i \cap \ddot{\lambda}_i) + 1 \quad i = 1,2,\ldots,p \end{aligned} \quad (25a)$$

and

$$\begin{aligned} \gamma_{i,i}^2 &\le (\dot{\varepsilon}_i \cup 1) \cap (\ddot{\varepsilon}_i \cup 1) \\ &= (\dot{\varepsilon}_i \cap \ddot{\varepsilon}_i) \cup (\dot{\varepsilon}_i \cap 1) \cup (\ddot{\varepsilon}_i \cap 1) \cup (1 \cap 1) \\ &= 1 \quad i = p+1, p+2, \ldots, k1 \end{aligned} \quad (25b)$$

Since R_1 and R_2 given in (16) are correlated with each other, i.e. $R_{12} = R_{21} \neq 0$, the intersection of two signal energies $(\dot{\lambda}_i \cap \ddot{\lambda}_i)$ in (25a) should be greater than zero. This observation implies that $\gamma_{i,j}^2 > 1, i = 1,\ldots,p$. Therefore, the statistical threshold for the new hypothesis test H_{p+1} is $\gamma_{i,j}^2 \le 1$. Restated, the number of endmembers can be determined by simply counting the number of diagonal elements whose values are greater than unity in (25).

5. Experimental Results

The analysis performed herein includes one data cubes acquired over the Cuprite, Nevada, in 1992 by the NASA/JPL AVIRIS instrument. Four methods are evaluated in the experiments: (a) conventional principal components analysis (*PCA*) using all bands, (b) *NAPCA* transform

using all bands described in Section *III-1*, (c) the singular value decomposition (*SVD*) to estimate the signal subspace of R_{12}, and (d) the *PNAPCA* procedure.

The data set used in the experiment is a subsection of the Cuprite image, which is a 200x200 pixel scene. Figure 1 depicts its image of the 0.752 *um* band. Since bands corresponding to the water absorption regions and the low *SNR* have no useful energy, they are removed before processing which leaves 192 bands in this study. This area has been extensively studied using field measurements [10], where the "alphabet" symbols denote the regions of the pure materials which have been found containing six significant materials: playa, kaolinite, alunite, silica, buddingtonite, and varnished tuff.

In the real detection, these four techniques were directly applied to the subsection of the Cuprite image. Figure 2 displays their results and Table 1 lists the details of the first ten eigenvalues. In Fig. 2(a), only the first two eigenvalues produced by *PCA* can be clearly separated while the gap among other consecutive signal and noise eigenvalues is inadequately large. Figure 2(b) and (c) summarizes the results from *NAPCA* and *SVD*, respectively, where the number of endmembers is underestimated to be five only. Finally, results obtained from *PNAPCA* with *UIMT* in Fig. 2(d) indicate that the number of endmembers is six. Therefore, the *PNAPCA* with *UIMT* is more advantageous in terms of resolving the inherent dimensionality problem.

6.Conclusion

Knowledge of the number of endmembers is a prerequisite for accurate spectral mixture modeling and endmember abundance estimation. To cope with the inaccuracy noise estimation in *NAPCA*, this work performs Partitioned version Noise-Adjusted Principal Components Analysis (*PNAPCA*) to solve the problem of the intrinsic dimensionality. The notion behind *PNAPCA* is the partitioning of the original data space into two distinct subspaces by a simultaneous transform. In addition, applying a simple hypothesis test, the *UIMT* allows us to accurately estimate the number of endmembers. Experimental results demonstrate that *PNAPCA* not only performs better than other conventional methods in computer simulations, but also functions well in realistic AVIRIS data.

7.References

[1] T. W. Anderson, "*An introduction to multivariate statistical analysis*," John Wiley and sons, New York, 1984.

[2] M. Wax and T. Kailaith, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoustics, Speech, Signal Processing*, vol. 33, no. 2, pp. 387-392, April 1985.

[3] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Automatic Control*, vol. 19, no. 6, pp716-723, 1974.

[4] J. Rissanen, "Modelling by shortest data description," *Automatica*, vol. 14, pp. 465-471, 1978.

[5] A. A. Green, M. Berman, P. Switzer and M. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Trans. Geoscience and Remote Sensing*, vol. 26, no. 1, pp. 65-74, Jan. 1988.

[6] J. B. Lee, A. S. Woodyatt and M. Berman, "Enhancement of high spectral resolution remote sensing data by a noise-adjusted principal components transform," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 28, pp. 295-304, May 1990.

[7] R. E. Roger, "A fast way to compute the noise-adjusted principal components transform matrix," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 32, pp. 1194-1196, 1990.

[8] B. Adams, and M. O. Smith, "Spectral mixture modeling: a new analysis of rock and soil types at the Viking Lander 1 site," *J. Geophys. Res.*, vol. 91, pp. 8098-8112, July 1986.

[9] K. Fukanaga, *Introduction to Statistical Pattern Recognition*, 2nd Edition, Academic Press, New York, 1990.

[10] G. A. Swayze, R.N. Clark, S. Sutley, and A Gallagher, "Ground-Truthing AVIRIS Mineral Mapping at Cuprite, Nevada," *Summaries of the Third Annual JPL Airborne Geosciences Workshop, Volume 1: AVIRIS Workshop*, JPL Publication 92-14, 1992.

Figure 1: A subsection of the Cuprite scene. The "alphabet" symbols denote the position of the pure materials found. Letter "A" is playa, "B" is kaolinite, "C" is alunite, "D" is silica, "E" is buddingtonite, and "F" is varnished tuff.

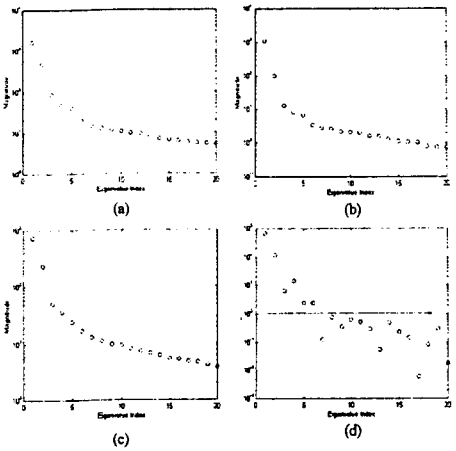

Figure 2: Results of real detection for the four techniques: (a) *PCA*, (b) *NAPCA*, (c) *SVD*, and (d) *PNAPCA*. Only the first fifty eigenvalues are plotted for clarity.

λ_i	*PCA*	*NAPCA*	*SVD*	*PNAPCA*
1	1.88e+6	1621.00	719.63	4.22e+005
2	1.94e+4	469.29	166.61	1.31e+004
3	2.50e+3	84.56	18.9	4.16
4	889.14	45.87	15.8	188.25
5	507.79	40.01	14.0	5.84
6	345.82	20.16	5.24	5.51
7	236.17	14.86	3.92	0.015
8	171.44	12.90	3.52	0.51
9	111.32	11.53	3.36	0.12
10	99.13	10.69	2.97	0.36

Table 1: the details of the first ten eigenvalues for four methods.