MULTIDIMENSIONAL WIENER FILTERING USING FOURTH ORDER STATISTICS OF HYPERSPECTRAL IMAGES

Damien Letexier and Salah Bourennane

Institut Fresnel, GSM Team
D.U. de Saint Jérôme
13397 Marseille Cedex 20, France

ABSTRACT

In this paper we propose a new multidimensional filtering method based on fourth order cumulants to denoise of data tensor impaired by correlated gaussian noise. We overview the multidimensional Wiener filtering that overcomes the well known lower rank-$(K_1, \ldots, K_N)$ tensor approximation. But this method only exploits second order statistics. In some applications, it may be interesting to consider a correlated Gaussian noise. Then, we propose to introduce the fourth order statistics in the denoising algorithm. Indeed, the use of fourth order cumulants enables to remove the Gaussian components of an additive noise. Qualitative results of the improved multidimensional Wiener filtering are shown for the case of noise reduction in hyperspectral imagery.

***Index Terms*—** denoising, tensor, cumulants, wiener.

1. INTRODUCTION

Some multidimensional denoising methods have been emerging for some years. Actually, the need of such methods is becoming interesting since there are many multisensor applications now. For instance hyperspectral imagery lead to a collection of images in several hundreds of contiguous spectral bands. This paper proposes an extension of the classical multidimensional filtering, namely Higher Order Singular Value Decomposition (HOSVD) [1] or Multidimensional Wiener Filtering [2, 3].

In these studies, the multidimensional data set $\mathcal{R}$ of size $I_1 \times I_2 \times \ldots \times I_N$ are often impaired by additive white Gaussian noise $\mathcal{N}$ statistically independent from the signal $\mathcal{X}$ ($\mathcal{R} = \mathcal{X} + \mathcal{N}$. Then, the estimated signal tensor $\hat{\mathcal{X}}$ can be obtained using a multidimensional filtering that can be written using n-mode products as $\hat{\mathcal{X}} = \mathcal{R} \times_1 \mathbf{H}_1 \times_2 \mathbf{H}_2 \times_3 \ldots \times_N \mathbf{H}_N$ [2, 3]. Each dimension of a N^{th} order tensor is denoted by n-mode. For instance, a hyperspectral image can be modeled as a third order tensor whose three modes are rows, columns (both corresponding to pixel localization) and spectra. Each matrix $\mathbf{H}^{(n)}$ can be called n-mode filter [2, 3]. It is a projector on the K_n dimensional n-mode signal subspace. For HOSVD or Lower Rank Tensor Approximation (LRTA), $\mathbf{H}^{(n)}$ is a simple projector [1] whereas it is obtained from the minimization of the mean squared error in multidimensional Wiener filtering [2, 3]. All these techniques are based on TUCKER3 tensor decomposition [4] that generalizes the matrix SVD. The computation of each processing can only be run thanks to an Alternative Least Square (ALS) loop. That is an iterative algorithm that permits to jointly estimate n-mode filters $\mathbf{H}^{(n)}, n = 1 \ldots N$.

In this paper, we propose to improve the multidimensional denoising methods in the case of an additive correlated Gaussian noise. For that purpose, we use fourth order cumulants to eliminate the Gaussian components of the additive noise.

2. MULTIDIMENSIONAL WIENER FILTERING

2.1. Tensor flattening

A tensor can be turned into a n-mode matrix (Fig. 1). The n-mode flattening matrix $\mathbf{A}_n$ of a tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times \ldots \times I_N}$ is defined as a matrix [1] from $\mathbb{R}^{I_n \times M_n}$ where :

$$M_n = I_1 \ldots I_{n-1} I_{n+1} \ldots I_N.$$

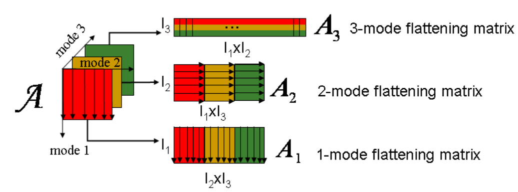

Fig. 1. *Flattening matrices for a third order tensor $\mathcal{A}$.*

Therefore, in our problem, $\mathbf{R}_n, \mathbf{X}_n$, and $\mathbf{N}_n$ are respectively the n-mode flattening matrices of data, signal and noise tensors.

2.2. Multidimensional Wiener Filtering

We represent a noisy multidimensional data set as a tensor resulting from a multidimensional signal $\mathcal{X}$ impaired by an ad-

1-4244-1484-9/08/$25.00 ©2008 IEEE

ditive white noise $\mathcal{N}$ [5]. Multidimensional Wiener Filtering (MWF) aims at estimating the desired signal $\mathcal{X}$ from data tensor $\mathcal{R}$ using multilinear algebra tools.

The optimal n-mode filter $\mathbf{H}^{(n)}$ is obtained through the minimization of the mean squared (MSE) error: $e(\mathbf{H}^{(1)}, \mathbf{H}^{(2)}, \mathbf{H}^{(3)}) = \mathrm{E}\left[\left\|\mathcal{X} - \hat{\mathcal{X}}\right\|^2\right]$. The computation of n-mode filters $\mathbf{H}^{(n)}$ is performed using an Alternating Least Squares (ALS) algorithm. In this iterative algorithm, the n-mode filters are initialized by corresponding identity matrices. Every m-mode filter $\mathbf{H}^{(m)}$ fixed, $m \neq n$, the expression of the optimal n-mode filter $\mathbf{H}^{(n)}$ is [2] :

$$\mathbf{H}^{(n)} = \mathbf{V}_s^{(n)} \mathbf{\Lambda}^{(n)} \mathbf{V}_s^{(n)T}, \tag{1}$$

with :

$$\mathbf{\Lambda}^{(n)} = diag\left(\frac{\lambda_1^{\gamma} - \sigma_{\gamma}^{(n)^2}}{\lambda_1^{\mathbf{\Gamma}}}, \ldots, \frac{\lambda_{K_n}^{\gamma} - \sigma_{\gamma}^{(n)^2}}{\lambda_{K_n}^{\mathbf{\Gamma}}}\right) \tag{2}$$

where, λ_i^{γ} and $\lambda_i^{\mathbf{\Gamma}}$, $i = 1$ to K_n, are respectively the K_n first eigenvalues of signal and data n-mode covariance matrices $\mathrm{E}[\mathbf{X}_n \hat{\mathbf{X}}_n^T]$ and $\mathrm{E}[\hat{\mathbf{X}}_n \hat{\mathbf{X}}_n^T]$ [2]. The noise power $\sigma_{\gamma}^{(n)^2}$ is estimated by computing the average of the $I_n - K_n$ smallest eigenvalues of $\mathrm{E}[\mathbf{R}_n \mathbf{R}_n^T]$: $\hat{\sigma}_{\gamma}^{(n)^2} = \frac{1}{I_n - K_n} \sum_{i=K_n+1}^{I_n} \lambda_i^{\gamma}$. The computation of n-mode filters (Eqs. (1) and (2)) involves the n-mode rank values K_1, K_2 and K_3.

However, when MWF is applied for images, some artifacts can occur when few rows or columns are redundant. In this case, the dimension of signal subspace may be underestimated. Consequently $\mathbf{H}^{(1)}, \mathbf{H}^{(2)}$ and $\mathbf{H}^{(3)}$ are not correctly estimated.

2.3. n-mode signal subspace dimension

In fact, as MWF is an extension of the singular value decomposition (SVD) truncation, it is obvious that some parts of the reconstructed image are missing after truncation. This is because the image information is spread over a large signal subspace.

Each singular vector $\mathbf{V}_s^{(n)}$, $n = 1, 2, 3$, is estimated by the truncation of n-mode covariance matrices $\mathrm{E}[\mathbf{R}_n \mathbf{R}_n^T]$ SVD. That is, keeping the K_n singular vectors associated with the K_n highest singular values of $\mathrm{E}[\mathbf{R}_n \mathbf{R}_n^T], n = 1, 2, 3$.

Consider two noisy images $\mathcal{R}_a$ and $\mathcal{R}_b$ containing the same features but otherwise disposed, such that $K_n^{(b)} < K_n^{(a)}$, where $K_n^{(l)}$ is the n-mode rank of $\mathcal{R}_l, l \in \{a, b\}$ -for example $\mathcal{R}_a$ contains a diagonal straight line and $\mathcal{R}_b$ a vertical straight line-. Each n-mode flattening matrix SNR can be defined as follows :

$$SNR = \frac{\sum_{i=1}^{K_n^{(l)}} \lambda_i^{(n)}}{\sum_{i=1}^{K_n^{(l)}} \sigma^2} = \frac{\sum_{i=1}^{K_n^{(l)}} \lambda_i^{(n)}}{K_n^{(l)} \cdot \sigma^2}, \tag{3}$$

where $\lambda_i^{(n)}$ is the i^{th} eigenvalue of the n-mode signal flattening matrix $\mathbf{X}_n$ and σ^2 is the additive white noise power. Assuming an equal signal energy in each n-dimension, $\sum_{i=1}^{K_n^{(a)}} \Lambda_i^{(n)} = \sum_{i=1}^{K_n^{(d)}} \Lambda_i^{(n)} = P$, where $\Lambda_i^{(n)} = \lambda_i^{(n)} + \sigma^2$ is the i^{th} eigenvalue of the n-mode data flattening matrix $\mathbf{R}_n$, SNR becomes :

$$SNR(K_n^{(l)}) = \frac{P - K_n^{(l)} \cdot \sigma^2}{K_n^{(l)} \cdot \sigma^2} = \frac{P}{K_n^{(l)} \cdot \sigma^2} - 1 \tag{4}$$

Eq. (4) clearly emphasizes that $SNR(K_n^{(b)}) > SNR(K_n^{(a)})$. Thus, it is interesting to have the smallest n-mode ranks values. The case is very easy when there is only straight lines in images, because rank reduction is more obvious: a rearrangement of data such that a diagonal line is turned into a vertical line. In the case of real-world image, we assume that there are several directions that carry most of information. Along these directions, the n-mode ranks are as small as possible.

2.4. Tensor Main Directions

In [3], the multidimensional Wiener filtering based on rearrangement of data ($MWFR$) has been proposed. It has been dealing with finding the main directions in a tensor thanks to an array processing method. In this paper, we propose to use the classic Hough transform to find the main directions. In this approach, a straight line is mapped to a bright point in the Hough plane (ρ, θ), where ρ is the distance of the desired line from the origin and θ is the orientation of the line. Then, we are looking for the brightest points in the Hough domain.

For hyperspectral images or color images, we have to compute the mean image in the spectral mode before applying the Hough transform. Then we can obtain the brightest points, corresponding to main straight line-like features. Fig. 2 depicts the Hough transform of a hyperspectral image. The brightest points correspond to orientations $0°$ and $34°$.

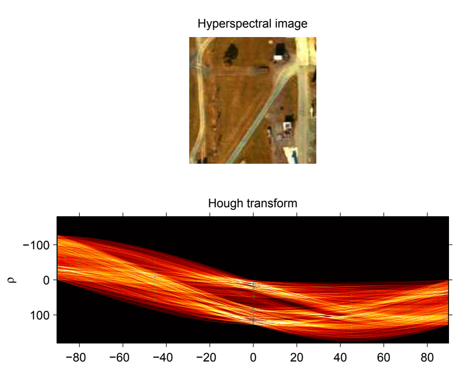

Fig. 2. *Hyperspectral image and its Hough transform. The brightest points correspond to* $0°$ *and* $34°$.

In our application, the offset ρ does not matter. Thus, the computational time of such a simple Hough transform is 0.017 seconds on a 2.66Ghz Dual Core 2 running MATLAB.

2.5. $MWFR$ algorithm

We denote by $MWFR$ the MWF using the retrieval of main directions. It is an Alternating Least Squares algorithm that can be sum up as follows :

- Input : Data tensor $\mathcal{R}$
- Initialization $k = 0$:
  $\mathcal{X}^0 = \mathcal{R} \Longleftrightarrow \mathbf{H}^{(n),0} = \mathbf{I}_{I_N}\ \forall n = 1 \dots N$.
- Detection of main directions θ_k using Hough transform and keeping its brightest points corresponding to the direction of smallest n-mode ranks.
- ALS loop:
  while $\|\mathcal{X}^{k+1} - \mathcal{X}^k\| > \varepsilon$
  1. for each main direction θ_k :
     (a) Rotation of the data tensor of angle θ_k
     (b) Estimation of n-mode filters $\mathbf{H}^{(n)}$
        i. $\mathcal{X}_n^k = \mathcal{R}_i \times_1 \mathbf{H}^{(1),k+1} \times_2 \dots \times_{n-1} \mathbf{H}^{(n-1),k+1} \times_{n+1} \mathbf{H}^{(n+1),k} \times_{n+2} \dots \times_N \mathbf{H}^{(N),k}$.
        ii. n-mode flatten $\mathcal{X}_n^k$ into matrix $\mathbf{X}_n^{(n),k} = \mathbf{R}_n(\mathbf{H}_{k+1}^{(1)} \otimes \dots \otimes \mathbf{H}_{k+1}^{(n-1)} \otimes \mathbf{H}_k^{(n)} \otimes \dots \otimes \mathbf{H}_k^{(N)})$
        iii. compute matrices $\mathbf{C}^{(n),k} = \mathbf{X}_n^{(n),k} \mathbf{R}_n^T$ and $\mathbf{D}^{(n),k} = \mathbf{X}_n^{(n),k} \mathbf{X}_n^{(n),k^T}$.
        iv. Process $\mathbf{C}^{(n),k}$ and $\mathbf{D}^{(n),k}$ K_n first eigenvalues λ_i^γ and λ_i^Γ corresponding to the K_n eigenvectors $\mathbf{v}_{i,s}$; $i = 1 \dots K_n$.
        v. Compute $\mathbf{H}^{(n)} = \mathbf{V}_s^{(n)} \mathbf{\Lambda}^{(n)} \mathbf{V}_s^{(n)^T}$.
     (c) Multidimensional Wiener filtering $\mathcal{X}^{k+1} = \mathcal{R}_i \times_1 \mathbf{H}^{(1),k+1} \times_2 \dots \times_{n-1} \mathbf{H}^{(n-1),k+1} \times_{n+1} \mathbf{H}^{(n+1),k+1} \times_{n+2} \dots \times_N \mathbf{H}^{(N),k+1}$.
     (d) $k \leftarrow k + 1$.
     (e) Rotation of the filtered tensor of angle $-\theta_k$
  2. Average of every filtered tensor on the different directions.
- Output : Estimated signal tensor $\hat{\mathcal{X}}$, which is the combination of every $\hat{\mathcal{X}}^{k_s}$. Here, k_s is the convergence iteration index.

2.6. Second order statistics in $MWFR$ algorithm

It is possible to give a statistical sense to matrices $\mathbf{C}^{(n),k}$ and $\mathbf{D}^{(n),k}$ from step 1(b)iii. Let us define $\mathbf{x}_j^{(n),k}$, $j = 1, \dots M_n$, the n-mode vectors of tensor $\mathcal{X}^{(n),k}$, i.e. n-mode flattening matrix $X_n^{(n),k}$ column vectors. Let us define as well $\mathbf{r}_j^{(n)}$, $j = 1, \dots M_n$, the n-mode vectors of tensor $\mathcal{R}$. Matrix $\mathbf{C}^{(n),k} = \mathbf{X}_n^{(n),k} \mathbf{R}_n^T$ can be written as:

$$\mathbf{C}^{(n),k} = [\mathbf{x}_1^{(n),k}, \dots, \mathbf{x}_{M_n}^{(n),k}][\mathbf{r}_1^{(n)}, \dots, \mathbf{r}_{M_n}^{(n)}]^T = \sum_{j=1}^{M_n} \mathbf{b}_j^{(n),k} \mathbf{r}_j^{(n)^T}$$

We have quite the same expression for matrix $\mathbf{D}^{(n),k}$.

As a consequence, up to the multiplicative factor $\frac{1}{M_n}$, matrices $\mathbf{C}^{(n),k}$ and $\mathbf{D}^{(n),k}$ are an estimation of the covariance matrix between data tensor $\mathcal{R}$ ($\mathbf{X}^{(n),k}$ in the case of $\mathbf{D}^{(n),k}$) n-mode vectors and tensor $\mathcal{X}^{(n),k}$ n-mode vectors.

Considering, in the previous expression of matrices $\mathbf{C}^{(n),k}$ and $\mathbf{D}^{(n),k}$, that $\{\mathbf{r}_j^{(n)},\ j = 1, \dots M_n\}$, and $\{\mathbf{x}_j^{(n),k},\ j = 1, \dots M_n\}$ are the M_n realizations of two random vectors $\mathbf{r}^{(n)}$ and $\mathbf{b}^{(n),k}$ associated respectively with the n-mode vectors of data tensors $\mathcal{R}$ and $\mathcal{X}^{(n),k}$, matrices $\mathbf{C}^{(n),k}$ and $\mathbf{D}^{(n),k}$ can be written as a second order moment: $\mathbf{C}^{(n),k} = \mathrm{E}[\mathbf{x}^{(n),k} \mathbf{r}^{(n)^T}]$ and $\mathbf{D}^{(n),k} = \mathrm{E}[\mathbf{x}^{(n),k} \mathbf{x}^{(n)^T}]$, where $\mathrm{E}[\cdot]$ denotes the expectation operator.

3. IMPROVEMENT THROUGH FOURTH ORDER CUMULANT

In practice, the noise whiteness and gaussianity are not always fulfilled conditions. The use of higher order statistics consists to eliminate the noise Gaussian components [6]. As remarked in previous section, n-mode covariance matrices $\mathbf{C}^{(n),k}$ and $\mathbf{D}^{(n),k}$ are defined as second order moments. They can be replaced by fourth order cumulants [6]: $\mathbf{O}^{(n),k} = \mathrm{Cum}(\mathbf{x}^{(n),k}, \mathbf{x}^{(n),k^T}, \mathbf{r}^{(n)}, \mathbf{r}^{(n)^T})$ and $\mathbf{Q}^{(n),k} = \mathrm{Cum}(\mathbf{x}^{(n),k}, \mathbf{x}^{(n),k^T}, \mathbf{x}^{(n)}, \mathbf{x}^{(n)^T})$. In practice, in order to reduce the computational load, a cumulant slice matrix of $\mathbf{C}^{(n),k}$ can be computed. The cumulant slice matrix associated with the first component of vector $\mathbf{x}^{(n),k}$, is given by the following $(I_n \times I_n)$-hermitian matrix [6, 7]: $\mathbf{O}_\mathbf{1}^{(n),k} = \mathrm{Cum}(\mathbf{x}_1^{(n),k}, \mathbf{x}_1^{(n),k^T}, \mathbf{r}^{(n)}, \mathbf{r}^{(n)^T})$.

The generic (i,j)-term of cumulant slice $\mathbf{C}_\mathbf{1}^{(n),k}$ expressed with the expectation operator is: $\mathbf{O}_{\mathbf{1}ij}^{(n),k} = \mathrm{E}[x_1^{(n),k^2} r_i^{(n)} r_j^{(n)}] - 2\mathrm{E}[x_1^{(n),k} r_i^{(n)}]\mathrm{E}[x_1^{(n),k} r_j^{(n)}]$.

The practical estimation of $\mathbf{O}_\mathbf{1}^{(n),k}$ uses the M_n realizations of random vectors $\mathbf{r}^{(n)}$ and $\mathbf{x}^{(n),k}$. Defining by $b_{ij}^{(n),k}$ and $r_{ij}^{(n)}$ the (i,j)-term of $\mathbf{X}_n^{(n),k}$ and $\mathbf{R}_n$ n-mode unfolding matrices, the estimation of $\mathbf{O}_{\mathbf{1}ij}^{(n),k}$ term is given by:

$$\begin{aligned} \mathbf{C}_{\mathbf{1}ij}^{(n),k} = & \tfrac{1}{M_n}(\textstyle\sum_{p=1}^{M_n} b_{1p}^{(n),k^2} r_{ip}^{(n)} r_{jp}^{(n)}) \\ & -\tfrac{2}{M_n^2}(\textstyle\sum_{p=1}^{M_n} b_{1p}^{(n),k} r_{ip}^{(n)})(\textstyle\sum_{p=1}^{M_n} b_{1p}^{(n),k} r_{jp}^{(n)}) \end{aligned} \tag{5}$$

The cumulant slice matrix permits to reduce the computational load involved by the whole fourth order cumulants but give similar results. In the following we use this expression instead of the second order moments in the $MWFR$ algorithm.

4. EXPERIMENTAL RESULTS

In the following simulations, the proposed method using fourth order cumulants in the improved $MWFR$ algorithm is applied on hyperspectral images. In the following experiments, the signal-to-noise ratio is defined as $SNR = 10 \cdot \log \frac{\|\mathcal{X}\|^2}{\|\mathcal{N}\|^2}$. Before presenting some results, we introduce a quality criterion to quantify *a posteriori* the quality of the estimation : $QC(\hat{\mathcal{X}}) = 10 \cdot \log \left(\frac{\|\mathcal{X}\|^2}{\|\hat{\mathcal{X}} - \mathcal{X}\|^2} \right)$.

We model the additive *correlated* Gaussian noise as : $\mathcal{N} = \mathcal{N}^w \times_1 W^{(1)} \times_2 W^{(2)} \times_3 W^{(3)}$, where $\mathcal{N}^w$ is a *white* Gaussian noise, statistically independent from the signal, and $\forall n = 1, 2, 3$, $W^{(n)}$ are weighting matrices which make the correlation between the n-modes. We consider real-world data hyperspectral images, obtained by HYDICE [8]. HYDICE is an airborne sensor which collects post-processed data in 210 wavelengths: 0.4 - 2.5 μm. Spatial resolution is 1.5 m and spectral resolution is 10 nm. We propose to depict the denoising on a hyperspectral image in Fig. 3.

We consider the image impaired by an additive correlated Gaussian noise as described earlier. We show that the filtering taking into account fourth order cumulants permit to improve significantly the result, compared to the $MWFR$ based on second order moments.

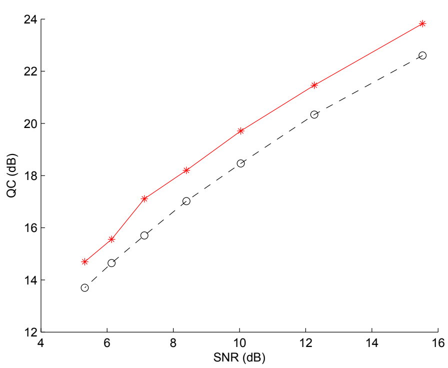

Fig. 3. *Interest of fourth order cumulants (red stars) compared to second order moments (black circles).*

An illustration of the algorithm output is given in Fig. 4. Showing that the denoising is improved when fourth order cumulants are taken into account.

5. CONCLUSION

In this paper, we have extended the multidimensional Wiener filtering so that it is possible to remove additive correlated Gaussian noise from multidimensional data. For that purpose we introduce the fourth order cumulant slice matrix in the $MWFR$ algorithm. This algorithm has been developed to take care of the data set specificities. Actually, it considers that multidimensional images contain several directions that focused most of information. This permits to reduce the rank of each flattening matrix. Some encouraging simulations are given in the last part of the paper for hyperspectral images.

(a) (b) (c)

Fig. 4. *(a) Additive correlated Gaussian noised image SNR=8 dB, (b) $MWFR$ $QC = 10.2dB$, (c) fourth order cumulants in $MWFR$ $QC = 12.4dB$.*

6. REFERENCES

[1] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM Jour. on Matrix An. and Applic.*, vol. 21, pp. 1253–78, 2000.

[2] D. Muti and S. Bourennane, "Survey on tensor signal algebraic filtering," *Signal Processing*, , no. 87, pp. 237–249, 2007.

[3] D. Letexier, S. Bourennane, and J. Blanc-Talon, "Nonorthogonal tensor matricization for hyperspectral image filtering," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 1, 2008.

[4] L. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, pp. 279–311, 1966.

[5] L. Parra, C. Spence, P. Sajda, A. Ziehe, and K. uller, "Unmixing hyperspectral data," *Advances in Neural Information Processing Systems*, vol. 12, pp. 942–948, 2000.

[6] J.M. Mendel, "Tutorial on higher order statistics (spectra) in signal processing and system theory: theoretical results and some applications," in *Proc. of the IEEE*, 1991, vol. 79, pp. 278–305.

[7] N. Yuen and B. Friedlander, "DOA in multipath: an approach using fourth order cumulant," *IEEE Trans. on Signal Process.*, vol. 45, no. 5, pp. 1253–63, 1997.

[8] L. J. Rickard, R. W. Basedow, E. F. Zalewski, P. R. Silverglate, and M. Landers, "HYDICE: an airborne system for hyperspectral imaging," in *Proc. SPIE, Imaging Spect. of the Terrestrial Environment*, G. Vane, Ed., 1993, vol. 1937, pp. 173–179.