

Improvement of Target Detection Based on Signal Dependent Noise Reduction for Hyperspectral Image

Caroline Fossati and Salah Bourennane

Aix Marseille Univ, CNRS, Centrale Marseille, Institut Fresnel, Marseille, France

Keywords: Detection, multi-linear algebra, reduction, hyperspectral image.

Abstract: In the hyperspectral images (HSI) acquired by the new-generation hyperspectral sensors the signal dependent noise is an important limitation to the detection. Therefore, noise reduction is an important preprocessing step to analyze the information in the hyperspectral image. A signal dependent noise cannot be reduced by conventional linear filtering. Therefore, a new method based on Parallel factor analysis (PARAFAC) decomposition is proposed to estimate the noise of hyperspectral remote sensing image. Then, the estimated noise is used for whitening the colored structural noise. By using this transformation, the data noise from new-generation hyperspectral sensors is diminished, thereby allowing a minimization of negative impacts on hyperspectral detection applications.

1. Introduction

Hyperspectral sensors collect data in hundreds of narrow contiguous spectral bands, providing powerful means to discriminate different materials by their spectral features. Actually, denoising is of great interest for target detection [1] with the underlying principle of targets being distinguishable. As HSIs are normally produced by a series of sensors, the noise mainly comes from two aspects: signal-independent circuitry noise (SI) and signal-dependent photon noise (SD) [2]. The noise level of HSI may vary dramatically from band to band. The noise variance in each band of HSI is not constant, in particular, there exist some bands at which the atmosphere absorbs so much light that the signal received from the surface is unreliable. Although the SD noise has become as dominant as the SI noise in HSI data collected by new-generation hyperspectral sensors due to the improved sensitivity in the electronic components [2, 3, 4, 5, 6, 7]. For this the denoising methods for those two types of noise are not the same. The SI noise term is generally modelled as additive and spatially stationary in each band, but the variance of the noise varies from band to band. That is to say, the level of the noise is dependent on the average amplitude of each band, but spatially stationary in each band. Based on HSI with high spatial resolution often contain a large number of small homogeneous areas, in [9] an automatic algorithm by dividing an image into several small blocks and calculating local means and local standard deviations of these blocks is developed, then estimating the noise by using a histogram statistical algorithm. To approximate better the noise variance of small blocks, [10] utilized data-masking technology by assuming that image textures are generally smoother than noise. However, the aforementioned methods treat separately each band in HSI data and fail to take spectral information into account. To utilize better the information from high spectral resolution of HSIs, a spectral and spatial decorrelation algorithm [11] based on the multiple linear regression (MLR) model was proposed to estimate noise. The residuals of the MLR model are considered to be noise, while the signal of a pixel at a particular band can be described as a linear combination of the neighboring pixels in the same band and the same spatial pixels in immediately adjacent bands. Based on the mixed noise assumption, the splitting of noise and the original signal from a HSI is usually the first step in these algorithms. Then, maximum likelihood estimation (MLE) is used for the estimation of the SD and SI noise parameters in the second step [12], this method is referred to as the hyperspectral noise parameter estimator (HYNPE) method. In [2], the SD and SI noise were estimated from a single scanning window. Because of the high computational complexity, the algorithm is not widely applied. In summary, MLE-based algorithms suffer from two disadvantages: sensitivity in initial value selection and high computational complexity for the optimal solution. Recently, in [13] denoising algorithm employing a spectral-spatial adaptive total variation model (SSATV), in which the spectral noise

differences and spatial information differences are both considered in the process of noise reduction is proposed. In this paper, we propose a novel algorithm that can estimate noise from HSIs with different noise types.

According to the different statistical properties of SI and SD noise, in this paper, we propose a new method based on PARAFAC decomposition [14, 15] to remove these two types of noise respectively. Our proposed method is applied to the simulated HSIs in order to evaluate their performances in a controlled environment. Results obtained on the real-world HYperspectral Digital Imagery Collection Experiment (HYDICE) and airborne visible/infrared imaging spectrometer (AVIRIS) Indian Pines HSIs are also presented and discussed. To compare the denoising performance of PARAFAC to other methods, multiway Wiener filter (MWF), prewhitening-multiway Wiener filter (PMWF) method proposed to reduce colored SI noise [16], HYNPE method, SSATV method and two well-known 2D denoising methods, minimum noise fraction (MNF) and noise-adjusted principal components analysis (PCA) [17], are used in the experiment section. The experimental results show that the proposed method have potential prospective in the reduction of both SI and SD noise in HSIs.

The remainder of the paper is organized as follows : Section 2 reviews some multilinear algebra tools. Section 3 gives the data model of HSI distorted by both SD and SI noise. Section 4 presents the detailed description of our proposed based on multilinear algebra decomposition for enhancement of SNR. Some denoising and comparative results are contained in Section 5. Finally, section 6 concludes the paper.

Within the scope of this paper, scalar is denoted by x , vector by \mathbf{x} , matrix by \mathbf{X} and tensor by \mathcal{X} .

2. Multilinear algebra tools and signal model

In this paper, the order of a tensor is defined as the number of dimensions or modes. Let $I_1, I_2, I_3 \in \mathbb{N}$ denote index upper bound the tensor. A tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ is a real 3-dimensional array, whose element is noted as x_{i_1, i_2, i_3} , where $i_1 = 1, \dots, I_1, i_2 = 1, \dots, I_2$ and $i_3 = 1, \dots, I_3$. In the following, some basic multilinear algebra tools used in tensor decompositions are introduced.

2.1. Rank-one Tensor

An N -mode tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ being rank 1 means that it can be written as the outer product [18] of 3 vectors, that is: $\mathcal{X} = \mathbf{a}^{(1)} \circ \mathbf{a}^{(2)} \circ \mathbf{a}^{(3)}$. So, each element of \mathcal{X} is $x_{i_1, i_2, i_3} = a_{i_1}^{(1)} a_{i_2}^{(2)} a_{i_3}^{(3)}$ for all $1 \leq i_n \leq I_n$ with $n = 1, 2, 3$, where $a_{i_1}^{(1)}, a_{i_2}^{(2)}$ and $a_{i_3}^{(3)}$ are the i_1 th, i_2 th and i_3 th element of $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}$ and $\mathbf{a}^{(3)}$, respectively.

2.2. n -mode Unfolding

The n -mode vectors are the I_n -dimensional vectors obtained from a tensor by varying index i_n while keeping the other indices fixed. The so-called n -mode flattened matrix $\mathcal{X}_n \in \mathbb{R}^{I_n \times M_n}$ ($n = 1, 2, 3$) denotes the n -mode unfolding matrix of a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$, with size $I_n \times M_n$ where $M_n = I_p \times I_q$ with $p \neq q \neq n$ ($p, q = 1, 2, 3$). The columns of \mathcal{X}_n are the I_n -dimensional vectors obtained from \mathbf{X} by varying index i_n while keeping the other indices fixed.

2.3. n -mode Product

The n -mode product " \times_n " is defined as the product between a data tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times \dots \times I_N}$ and a matrix $\mathbf{B} \in \mathbb{R}^{J \times I_n}$ in mode n . It leads to the tensor $\mathcal{U} = \mathcal{X} \times_n \mathbf{B}$ of size $I_1 \times \dots \times I_{n-1} \times J \times I_{n+1} \times \dots \times I_N$, whose entries are given by $u_{i_1, \dots, i_{n-1}, j, i_{n+1}, \dots, i_N} = \sum_{i_n=1}^{I_n} x_{i_1, i_2, \dots, i_N} b_{j, i_n}$ where b_{j, i_n} denotes the (j, i_n) element of matrix \mathbf{B} and $j = 1, \dots, J$.

2.4. PARAFAC Decomposition Model

PARAFAC model factorizes a tensor into a sum of rank-1 tensors [18]. For instance, tensor $\mathcal{Y} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ can be expressed as

$$\mathcal{Y} \approx \hat{\mathcal{Y}} = \sum_{k=1}^{K_s} \mathbf{Y}_k = \sum_{k=1}^{K_s} \lambda_k \mathbf{a}_k^{(1)} \circ \mathbf{a}_k^{(2)} \circ \mathbf{a}_k^{(3)} \quad (1)$$

where K_s is the rank, $\hat{\mathcal{Y}}$ is the rank- K_s PARAFAC approximation of \mathcal{Y} ; $\mathbf{Y}_k \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ is rank-1 tensor; $\mathbf{a}_k^{(1)}, \mathbf{a}_k^{(2)}, \mathbf{a}_k^{(3)} \in \mathbb{R}^{I_n}$ are normalized vectors of the n -mode space of \mathcal{Y} normalized by $\mathbf{a}_k^{(n)} = \mathbf{a}_k^{(n)} / \|\mathbf{a}_k^{(n)}\|$, $n = 1, 2, 3$; and $\lambda_k = \|\mathbf{a}_k^{(1)}\| \|\mathbf{a}_k^{(2)}\| \|\mathbf{a}_k^{(3)}\|$, with $k = 1, 2, \dots, K_s$. From Eq. (1) the (i_1, i_2, i_3) th entry of the tensor $\hat{\mathcal{Y}}$ can be expressed as

$$\hat{y}_{i_1, i_2, i_3} = \sum_{k=1}^{K_s} \lambda_k a_{i_1, k}^{(1)} a_{i_2, k}^{(2)} a_{i_3, k}^{(3)} \quad (2)$$

with $i_1 = 1, \dots, I_1, i_2 = 1, \dots, I_2, i_3 = 1, \dots, I_3$.

3. Data model

A noisy HSI can be expressed as a third order tensor $\mathcal{R} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ composed of a multidimensional signal $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ impaired by an additive random noise $\mathcal{N} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$. The tensor \mathbf{R} can be expressed as [19]:

$$\mathcal{R} = \mathcal{X} + \mathcal{N} \quad (3)$$

where \mathcal{N} accounts for both SI and SD noise [20] and its variance depends on the pixel x_{i_1, i_2, i_3} in the useful signal \mathbf{X} . Element-wise, the data model is:

$$r_{i_1, i_2, i_3} = x_{i_1, i_2, i_3} + (x_{i_1, i_2, i_3})^{1/2} \cdot u_{i_1, i_2, i_3} + w_{i_1, i_2, i_3} \quad (4)$$

where u_{i_1, i_2, i_3} is a stationary, zero-mean uncorrelated random process independent of x_{i_1, i_2, i_3} with variance σ_{u, i_3}^2 and w_{i_1, i_2, i_3} is electronics noise which is zero-mean white Gaussian noise in each band with variance σ_{w, i_3}^2 . The additive term $x^{1/2}u$ is the generalized SD noise and denoted as photon noise, w is the SI noise component and is generally assumed to be Gaussian distribution in each band. Then, we can define:

$$\mathcal{N} = \mathcal{N}_{SD} + \mathcal{N}_{SI} = \mathcal{N}_{SD} + \mathcal{W}, \quad (5)$$

and Equation (3) can be correspondingly rewritten as

$$\mathcal{R} = \mathcal{X} + \mathcal{N}_{SD} + \mathcal{W}. \quad (6)$$

With the assumption that x , u and w are independent and both u and w are zero mean and are stationary, the variance of noise \mathcal{N} in band i_3 of the HSI could be written as [2, 12]

$$\sigma_{\mathcal{N}, i_3}^2 = \sigma_{u, i_3}^2 \cdot \mu_{i_3} + \sigma_{w, i_3}^2 \quad (7)$$

where $\mu_{i_3} \triangleq \mathbb{E}[\mathbf{X}] = 1/(I_1 I_2) \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} x_{i_1, i_2, i_3}$ is the mean of all x_{i_1, i_2, i_3} in the i_3 th band of \mathcal{X} with $i_3 = 1, \dots, I_3$. The unfolding matrix $\mathbf{R}_3 \in \mathbb{R}^{I_3 \times M_3}$ of the HSI data tensor $\mathcal{R} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ (with $M_3 = I_1 I_2$) can be expressed as :

$$\mathbf{R}_3 = \mathbf{X}_3 + \mathbf{N}_3 \quad (8)$$

where \mathbf{X}_3 is the 3-mode unfolding matrix of the multidimensional signal tensor \mathcal{X} and

$$\mathbf{N}_3 = \mathbf{N}_{SD,3} + \mathbf{W}_3 \quad (9)$$

with $\mathbf{N}_{SD,3}$ and \mathbf{W}_3 being the 3-mode unfolding matrices of \mathcal{N}_{SD} and \mathcal{W} respectively. Using the mean noise variance of the i_3 th spectral band defined as $1/(I_1 I_2) \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \sigma_{\mathcal{N}, i_1, i_2, i_3}^2 = \mu_{i_3} \sigma_{u, i_3}^2 + \sigma_{w, i_3}^2$ where $\mu_{i_3} = 1/(I_1 I_2) \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} x_{i_1, i_2, i_3}$ is the mean of all x_{i_1, i_2, i_3} in the i_3 th band of \mathcal{X} with $i_3 = 1, \dots, I_3$, and the assumption of the independence of x and u , where u is zero-mean and independent between spectral bands, the covariance matrix of the 3-mode unfolding matrix $\mathbf{N}_{SD,3}$ can be expressed as:

$$\mathbf{C}_{\mathbf{N}_{SD}}^{(3)} = \text{diag}(\mu_1 \sigma_{u,1}^2, \mu_2 \sigma_{u,2}^2, \dots, \mu_{I_3} \sigma_{u, I_3}^2) \quad (10)$$

4. Proposed denoising method based on PARAFAC decomposition model

Our aim is to estimate the signal tensor \mathcal{X} from tensor $\hat{\mathcal{R}}$, in the sense of minimum mean square error. From Equation (1) the rank- K_s PARAFAC approximation of noisy $\hat{\mathcal{R}}$ is:

$$\mathcal{R}_a = \sum_{k=1}^{K_s} \lambda_k \mathbf{a}_k^{(1)} \circ \mathbf{a}_k^{(2)} \circ \mathbf{a}_k^{(3)} \quad (11)$$

With the assumption that the noisy tensor $\hat{\mathcal{R}}$ can be exactly expressed by sum of K ($K > K_s$) rank-1 tensors, then:

$$\hat{\mathcal{R}} = \sum_{k=1}^{K_s} \lambda_k \mathbf{a}_k^{(1)} \circ \mathbf{a}_k^{(2)} \circ \mathbf{a}_k^{(3)} + \sum_{k=K_s+1}^K \lambda_k \mathbf{a}_k^{(1)} \circ \mathbf{a}_k^{(2)} \circ \mathbf{a}_k^{(3)} = \mathcal{R}_a + \mathcal{M} \quad (12)$$

where $\mathcal{M} = \sum_{k=K_s+1}^K \lambda_k \mathbf{a}_k^{(1)} \circ \mathbf{a}_k^{(2)} \circ \mathbf{a}_k^{(3)}$ is a residual tensor.

Since the rank-1 tensors are orthogonal, then the square error of PARAFAC decomposition

$\|\hat{\mathcal{R}} - \mathcal{R}_a\|^2 = \sum_{k=K_s+1}^K \lambda_k^2 \|\mathbf{a}_k^{(1)} \circ \mathbf{a}_k^{(2)} \circ \mathbf{a}_k^{(3)}\|^2$. According to the definition in Equation (2), $\mathbf{a}_k^{(n)} = \mathbf{a}_k^{(n)} / \|\mathbf{a}_k^{(n)}\|$, $n = 1, 2, 3$, element-wise, $a_{i,k}^{(n)} = a_{i,k}^{(n)} / \left(\sum_{i=1}^{I_n} |a_{i,k}^{(n)}|^2 \right)^{1/2}$, then

$$\begin{aligned} \mathbf{a}_k^{(1)} \circ \mathbf{a}_k^{(2)} \circ \mathbf{a}_k^{(3)} &= \frac{a_{i,k}^{(1)} a_{j,k}^{(2)} a_{p,k}^{(3)}}{\left(\sum_{i=1}^{I_1} |a_{i,k}^{(1)}|^2 \right)^{1/2} \left(\sum_{j=1}^{I_2} |a_{j,k}^{(2)}|^2 \right)^{1/2} \left(\sum_{p=1}^{I_3} |a_{p,k}^{(3)}|^2 \right)^{1/2}} \\ &= \frac{a_{i,k}^{(1)} a_{j,k}^{(2)} a_{p,k}^{(3)}}{\left(\sum_{i=1}^{I_1} \sum_{j=1}^{I_2} \sum_{p=1}^{I_3} |a_{i,k}^{(1)} a_{j,k}^{(2)} a_{p,k}^{(3)}|^2 \right)^{1/2}} \end{aligned} \quad (13)$$

so, $\|\mathbf{a}_k^{(1)} \circ \mathbf{a}_k^{(2)} \circ \mathbf{a}_k^{(3)}\|^2 = 1$ and $\|\hat{\mathcal{R}} - \mathcal{R}_a\|^2 = \sum_{k=K_s+1}^K \lambda_k^2$.

Therefore, $\min \|\hat{\mathcal{R}} - \mathcal{R}_a\|^2 = \min \sum_{k=K_s+1}^K \lambda_k^2 \implies \{\lambda_k \mid k = K_s+1, \dots, K\}$ are $(K - K_s)$ smallest terms among $\{\lambda_k \mid k = 1, \dots, K\}$. Therefore, the minimum of the square error $\|\hat{\mathcal{R}} - \mathcal{R}_a\|^2$ corresponds to throw away other smaller terms from $K_s + 1$ to K of PARAFAC decomposition. The signal components in the smallest terms from $K_s + 1$ to K are the smallest ones among all the signal components from 1 to K of PARAFAC decomposition. Therefore, for fixed K_s , PARAFAC decomposition can reduce the noise, that is to say: the rank- K_s PARAFAC approximation of a noisy tensor results in an estimation of the signal, i.e.,

$$\mathcal{R}_a = \sum_{k=1}^{K_s} \lambda_k \mathbf{a}_k^{(1)} \circ \mathbf{a}_k^{(2)} \circ \mathbf{a}_k^{(3)} \approx \hat{\mathcal{X}}. \quad (14)$$

Since the denoising by PARAFAC decomposition is based on skipping smaller terms from $K_s + 1$ to K where noise components exist, so PARAFAC decomposition has the effect of the reduction of noise, that is to say the residual parts of PARAFAC decomposition, i.e.,

$$\hat{\mathcal{R}} - \mathcal{R}_a = \hat{\mathcal{R}} - \hat{\mathcal{X}} = \hat{\mathcal{N}} \quad (15)$$

is an estimate of the noise \mathcal{N} .

4.1. Estimation of the Optimal Rank K_s of PARAFAC Decomposition

In this paper, we assume that the statistical properties of both the signal and the noises are relatively constant in the tensor \mathcal{R} and the variance of the SD noise in Equation (7) can still be expressed by $\hat{\sigma}_{u,i_3}^2 \cdot \hat{\mu}_{i_3}$, where $\hat{\sigma}_{u,i_3}^2$ and $\hat{\mu}_{i_3}$ are the variance of random process u_{i_1, i_2, i_3} and the mean of all pixels in the i_3 th band of \mathcal{R} , respectively. Therefore, the covariance matrix $\mathbf{C}_N^{(3)}$ of noise is a diagonal matrix. If the squared norm of the covariance $\|\mathbf{C}_N^{(3)}\|^2$ is quite close to the sum of the squared diagonal elements $\sum_{i_3=1}^{I_3} c_{i_3, i_3}^2$, then this $\mathbf{C}_N^{(3)}$ can be considered approaching a diagonal matrix. This criterion is used to estimate the optimal rank K_s of PARAFAC decomposition for the reduction of noise.

4.2. Summary of the Proposed Method for Signal-to-Noise Enhancement

The complete proposed method to reduce both SD and SI noise in HSIs can be summarized as follows:

1. Estimate the noise \mathcal{N} from \mathcal{R} using PARAFAC:
 - (a) Unfold the estimated noise tensor \mathcal{N} to n -mode unfolding matrix with $n = 1, 2, 3$,
 - (b) Calculate the covariance matrix of the n -mode unfolding matrix of the estimated noise tensor,
 - (c) Use the criterion above to assess the result matrix obtained from step b) diagonal matrix. Then, the rank K_s of PARAFAC decomposition for the reduction of noise could be obtained from these three steps.
2. Estimate the signal tensor $\hat{\mathcal{X}} = \mathcal{R} - \hat{\mathcal{N}}$

5. Experimental results

To demonstrate the efficacy of different methods for noise reduction, the peak signal to noise ratio (PSNR) of the denoised HSI will be compared. The PSNR can be calculated by

$$\text{PSNR} = 10 \log_{10} \frac{(\max\{\mathcal{X}\})^2}{\text{MSE}} \quad (\text{dB}) \quad (16)$$

where

$$\text{MSE} = 1/(I_1 I_2 I_3) \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \sum_{i_3=1}^{I_3} (x_{i_1 i_2 i_3} - \hat{x}_{i_1 i_2 i_3})^2 \quad (17)$$

In addition, the variance of the residual noise, $\hat{\mathcal{X}} - \mathcal{X}$, in the simulated HSI and the variance of the removed noise, $\mathcal{R} - \hat{\mathcal{X}}$, in the real data are compared in this paper after denoising by different considered methods: PARAFAC, PMWF, HYNPE, MWF, SSATV, and two well-known 2D denoising methods, MNF and noise-adjusted PCA.

5.1. Simulated Data Experiments

To verify the performance of the proposed method a synthetic HSI is generated according to the data model in Equation (3), with the spectral signatures presented in Fig. 1 (a), having size $150 \times 150 \times 148$. There are six target types and three different spatial sizes 7×7 pixels, 2×2 pixels and 1×1 pixel of each type, which are shown in Fig. 1(b). These targets are mixed to the background by using the linear mixing model with target abundance being 85% (mixing ratio). The random noise is generated with a variance depending on the value of the useful signal according to Equation (7) and added into the signal \mathcal{X} as Eq. (4) to create the simulated HSI data \mathcal{R} . In the following sections, the variance of the residual noise in the simulated HSIs after denoising, the peak signal to noise ratio (PSNR) and the ACE target detection results of simulated HSI denoised by the considered methods will be illustrated and discussed.

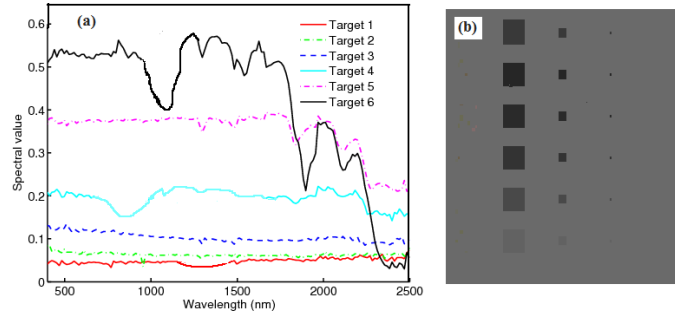


Figure 1: (a) Spectral signatures of the simulated targets and background, (b) Simulated HSI without noise, from top to bottom the index of targets is 1 to 6 respectively

5.2. Variance of Residual Noise

Fig. 2(a) shows the noisy simulated image with SNR= 30dB, from which one can see that the small targets are almost disappeared in the noise. The variances of SD and SI noise are chosen equal ($\sigma_{\mathcal{N}_{SD}}^2 = \sigma_w^2$), thus the noise variance of \mathcal{N} in this case varies strongly from band to band. The variance of the residual noise, $\hat{\mathcal{X}} - \mathcal{X}$, in the denoised HSI is evaluated at each band and some results are shown in Fig. 2 (b) where the noise variance of \mathcal{N} in the raw simulated HSI is also illustrated as a comparison. From Fig. 2 (b) it can be seen that all the considered methods can effectively remove the noise in the simulated HSI since the residual noise variance are much lower than that in raw HSI. But, Fig. 2 (b) demonstrates that the denoised HSI by PARAFAC method contains the least noise when compared against other considered methods.

5.3. Target Detection Performance: Probability of Detection

The main purpose of HSI denoising is to improve the results of detection, classification, etc. In this section, we focus on the improvement of target detection using the adaptive coherence/cosine estimator (ACE) [22] which is largely applied to HSI data. The results of ACE target detection of both simulated and real data denoised by

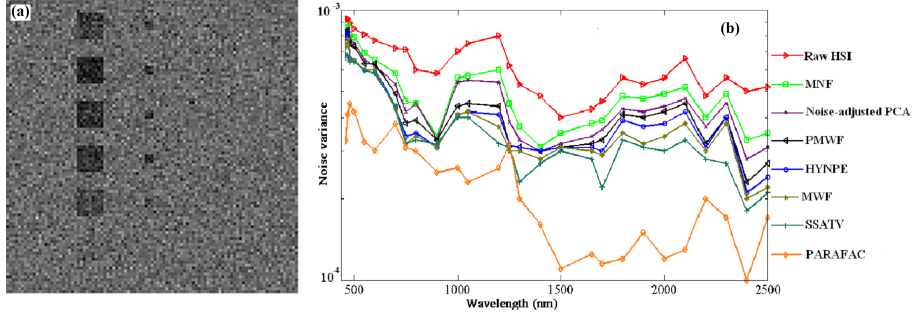


Figure 2: (a) Noisy image with SNR = 30 dB, (b) Noise variance of raw simulated HSI and residual noise variances of denoised HSI

our proposed PARAFAC method and other considered methods are shown and discussed. For the HSI described in Equation (3), the ACE detector can be expressed as:

$$\text{ACE} = \frac{(\mathbf{s}^T \Gamma^{-1} \mathbf{r}_j)^2}{(\mathbf{s}^T \Gamma^{-1} \mathbf{s})(\mathbf{r}_j^T \Gamma^{-1} \mathbf{r}_j)} \quad (18)$$

where \mathbf{r}_j is the vector in the unfolding matrix \mathbf{R}_3 of tensor \mathcal{R} with $j = 1, \dots, I_1 I_2$, \mathbf{s} is the target spectrum template Γ is the covariance matrix of \mathbf{R}_3 . To assess the performance of detection, the probability of detection (PD) is defined as:

$$\text{PD} = \frac{\sum_i^{n_s} N_i^{\text{rd}}}{\sum_i^{n_s} N_i}, \quad (19)$$

and the probability of false alarm (PFA) is defined as: $\text{PFA} = \frac{\sum_i^{n_s} N_i^{\text{fd}}}{\sum_i^{n_s} (I_1 \times I_2 - N_i)}$, where n_s is the number of spectral signatures, N_i the number of pixels with spectral signature i , N_i^{rd} the number of rightly detected pixels, and N_i^{fd} the number of falsely detected pixels. To set the values of the two noise variances $\sigma_{\mathcal{N}_{SD}}^2 = 1/I_3 \sum_{i_3=1}^{I_3} \sigma_{u,i_3}^2 \cdot \mu_{i_3}$ and σ_{w,i_3}^2 , we define some quantities: The signal to noise ratio (SNR) of the synthetic noisy HSI

$$\text{SNR} = 10 \log_{10} \frac{\|\mathcal{X}\|^2}{\|\mathcal{N}\|^2} \quad (\text{dB}) \quad (20)$$

then the noise variance of \mathcal{N} is $\sigma_{\mathcal{N}}^2 = P \cdot 10^{-\frac{\text{SNR}}{10}}$ with

$$P = 1/(I_1 I_2 I_3) \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \sum_{i_3=1}^{I_3} x_{i_1, i_2, i_3}^2. \quad (21)$$

Assuming

$$\delta = \sigma_{\mathcal{N}}^2 / \sigma_w^2, \quad (22)$$

for the given values of SNR the values of the noise model parameters can be obtained:

$$\begin{aligned} \sigma_{\mathcal{N}_{SD}}^2 &= \frac{\delta \cdot \sigma_{\mathcal{N}}^2}{1 + \delta} \\ \sigma_w^2 &= \frac{\sigma_{\mathcal{N}}^2}{1 + \delta} \end{aligned} \quad (23)$$

With the definitions above, it is clear that for $\delta = 1$, $\sigma_{\mathcal{N}_{SD}}^2 = \sigma_w^2$, i.e., SD noise source contributes similarly as white SI noise source to the simulated HSI, and there is not a dominant noise source in the simulated HSI, for $\delta < 1$, σ_w^2 is higher than $\sigma_{\mathcal{N}_{SD}}^2$ and the simulated HSI is distorted mainly by white SI noise, otherwise for $\delta > 1$ the SD noise source is dominant. So, in this paper, we consider the cases of $\delta = [0.1, 0.3, 0.5, 1.0, 1.5, 2.0, 2.5]$ with constant SNR = 30dB which can help to evaluate the influence of SD and SI noise variance on the denoising performance of different methods. Fig. 3 shows the PD of ACE target detection results under the condition PFA=10⁻⁴ for simulated HSI (Fig. 2 (a)) denoised by the considered methods in this paper. In Fig. 3 it is clear that the probability of detection of denoised HSI by PARAFAC method for different values of

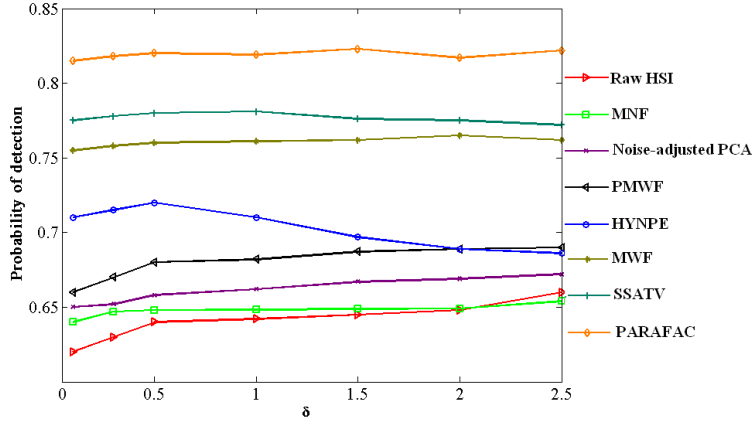


Figure 3: Probability of detection of simulated HSI denoised by different methods with SNR= 30dB and for PFA= 10^{-4}

δ outperforms other methods. The SD and SI noise are removed by throwing away other smallest terms from $K_s + 1$ to K , thus the quality of the denoised HSI by PARAFAC is ameliorated so much that the ACE target detection of the denoised HSI is improved greatly.

For the real-world HSIs, the denoising performance of the proposed method is also verified and discussed in the next section.

5.4. Results on Real-World Data

As the previous test is not realistic in the sense that true HSI is simulated according to the data model in Equation (3)-(7), a real-world image is considered in this section. Referred to as HYDICE HSI, was acquired by HYperspectral Digital Imagery Collection Experiment (HYDICE).

5.5. Removed Noise Variance and Detection Results

The real-world HYDICE HSI shown in Fig. 4(a) has 150 rows and 140 columns and 148 spectral channels out of 210 with 0.75 m spatial and 10 nm spectral resolution. It can be represented as a 3D data cube, denoted by $\mathcal{R} \in \mathbb{R}^{150 \times 140 \times 148}$. Six targets are added into the HSI and each row of targets in Fig. 4(a) has the same target spectral signature (spectral reflectance) illustrated in Fig. 4(b), which is taken from the image itself. The target size is 5×5 pixels along the first column, 3×3 pixels along the second one and 1×1 pixel along the last one. To assess the denoising results obtained by the proposed method, the removed noise ($\mathcal{R} - \hat{\mathcal{X}}$) variance is calculated at each band and plotted in Fig. 5 and the receiver operating characteristic (ROC) curves of ACE target detection is presented in Fig. 6.

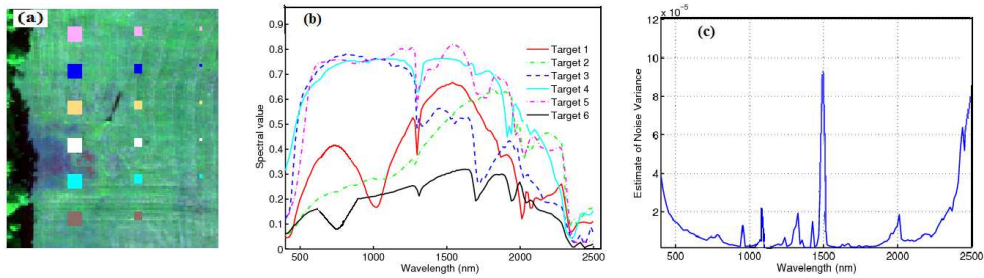


Figure 4: (a) HYDICE HSI used to compare different denoising methods, from top to bottom the index of targets is 1 to 6 respectively. (b) Spectral signatures (spectral reflectances) of the simulated targets. (c) Noise variance of raw HYDICE HSI

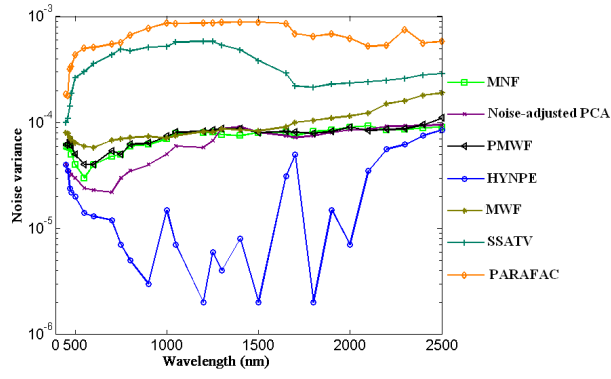


Figure 5: Variance of removed noise from HYDICE HSI with different methods

Fig. 5 further demonstrates the effectiveness of noise reduction of the image by the proposed method, which indicates that SD noise really exists in this HYDICE HSI image.

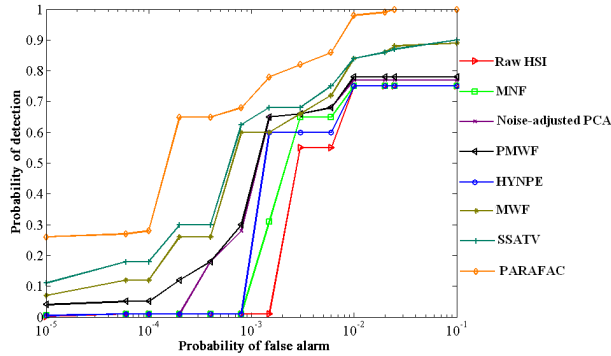


Figure 6: ROC curves of the denoised HYDICE HSIs obtained by ACE detector

Fig. 6 shows the comparison of ROC curves obtained by ACE detector from different methods. It is obvious that the PD values of ACE target detection of denoised HYDICE HSI by PARAFAC method are improved significantly when compared against raw HYDICE HSI counterpart and it is indicated that they are superior to the other considered methods in the reduction of both SD and SI noise. The improvement of target detection of the HYDICE HSI denoised by PARAFAC method is superior to that by SSATV. Since the PARAFAC method has not limitation in denoising colored noise, it can be concluded that the SD noise is at least as dominant as the SI noise in this HSI. The PMWF, MWF, noise-adjusted PCA and MNF methods are not designed for SD noise reduction, so their denoising performance are not ideal, which is reflected indirectly by the target detection results in these ROC curves in Fig. 6. HYNPE method has limitation in removing all SD noise components as shown in Fig. 5, the removed noise by this method is least and correspondingly its denoising performance confines its contribution in the improvement of target detection of this HYDICE HSI.

6. Conclusion

In this paper, we developed a new multidimensional denoising method based on multiple linear regression and multilinear algebra tools to enhance SNR of HSI data collected by new-generation hyperspectral sensors, distorted by both SI colored and SD noise. To reduce both SD and SI noise, we propose a tensor-based method. PARAFAC decomposition is applied to the noisy HSI to estimate the signal tensor. PARAFAC decomposition must be conducted at the appropriate rank which can be estimated according to the statistical properties of noise. The performance of the proposed PARAFAC method are validated on the simulated HSIs distorted by both SD and SI noise and on the real-world HYDICE. The HYDICE dataset is used to evaluate the denoising and detection results. The experimental results show the efficiency of the proposed denoising algorithm to improve the SNR in hyperspectral images and the target detection.

From the analysis and the comparative study against other similar methods in the experiments, it can be concluded that PARAFAC method can effectively reduce both SD and white or colored SI noise from HSIs. It is also necessary to take into account the noise signal-dependency hypothesis when dealing with HYDICE HSI data.

Acknowledgements

The authors thank the reviewers for their careful reading and helpful comments which improve the quality of the paper.

References

- [1] N. Renard, S. Bourennane, and J. Blanc-Talon. Denoising and dimensionality reduction using multilinear tools for hyperspectral images. *IEEE Geoscience and Remote Sensing Letters* 5 (2000) 138–142.
- [2] M. Uss, B. Vozel, V. Lukin, and K. Chehdi. Local signal-dependent noise variance estimation from hyperspectral textural images. *IEEE Journal of Selected Topics in Signal Proc.* 5 (2011) 469–486.

- [3] X. Liu, S. Bourennane and C. Fossati. Reduction of noise from hyperspectral images for target detection. *IEEE Trans. Geoscience and Remote Sensing* 9 (2014) 5396–5411.
- [4] S. Bourennane, C. Fossati and A. Cailly. Improvement of target-detection algorithms based on adaptive three-dimensional filtering. *IEEE Trans. Geosci. Remote Sens.* 49(2011) 1383–1395.
- [5] H.K. Aggarwal, A. Majumdar, Exploiting spatio-spectral correlation for impulse denoising in hyperspectral images. *J. Electron. Imaging* 24 (2015) 013027.
- [6] P. Meza, J.E. Pezoa, S.N. Torres, Multidimensional Striping Noise Compensation in Hyperspectral Imaging: Exploiting Hypercubes Spatial, Spectral, and Temporal Redundancy. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 9 (2016) 4428–4441.
- [7] B. Rasti, M.O. Ulfarsson, P. Ghamisi, Automatic Hyperspectral Image Restoration Using Sparse and Low-Rank Modeling. *IEEE Geosci. Remote Sens. Lett.* (2017) 1-5.
- [8] P. Zhong and R. Wang. Multiple-spectral-band CRFs for denoising junk bands of hyperspectral imagery. *IEEE Trans. Geoscience and Remote Sensing* 4 (2013) 2269–2275.
- [9] B. Gao. An operational method for estimating signal to noise ratios from data acquired with imaging spectrometers. *Remote Sens. Environ.* 43 (1993) 23–33.
- [10] B. Corner, R. Narayanan, and S. Reichenbach. Noise estimation in remote sensing imagery using data masking. *International Journal Remote Sensing* 24 (2003) 689–702.
- [11] R. E. Roger and J. F. Arnold. Reliably estimating the noise in AVIRIS hyperspectral images. *International Journal of Remote Sensing* 17 (1996) 1951–1962.
- [12] L. Alparone, M. Selva, B. Aiazzi, S. Baronti, F. Butera, and L. Chiarantini. Signal-dependent noise modelling and estimation of new-generation imaging spectrometers. *First IEEE Workshop on Hyperspectral Image and Signal Proc.: Evolution in remote sensing* (2009) 1–4.
- [13] Q. Yuan, L. Zhang, and H. Shen. Hyperspectral image denoising employing a spectral-spatial adaptive total variation model. *IEEE Transactions on Geoscience and Remote Sensing* 50 (2012) 3660–3677.
- [14] R. A. Harshman and M. E. Lundy. The parafac model for three-way factor analysis and multidimensional scaling. *Research methods for multimode data analysis*. Praeger, (1984) 122–215.
- [15] R. A. Harshman and M. E. Lundy. Parafac: Parallel factor analysis. *Computational statistics and data analysis* 18 (1994) 39–72.
- [16] X. Liu, S. Bourennane, and C. Fossati. Nonwhite noise reduction in hyperspectral images. *IEEE Geoscience and Remote Sensing letters*, 9 (2012) 368–372.
- [17] C.-I. Chang and Q. Du. Interference and noise adjusted principal components analysis. *IEEE Transactions on Geoscience and Remote Sensing* 37 (1999) 2387–2396.
- [18] T. G. Kolda and B. W. Bader. Tensor decompositions and applications. *SIAM Review* 51 (2009) 455–500.
- [19] D. Letexier and S. Bourennane. Noise removal from hyperspectral images by multidimensional filtering. *IEEE Transactions on Geoscience and Remote Sensing* 46 (2008) 2061–2069.
- [20] J. Meola, Michael T. Eismann, Randolph L. Moses, and Joshua N. Ash. Modeling and estimation of signal-dependent noise in hyperspectral imagery. *Appl. Opt.* 50 (2011) 3829–3846.
- [21] R. E. Roger. Principal components transform with simple, automatic noise adjustment. *International Journal of Remote Sensing* 17 (1996) 2719–2727.
- [22] D. Manolakis and G. Shaw. Detection algorithms for hyperspectral imaging applications. *IEEE Signal Processing Magazine* 19 (2002) 29–43.
- [23] C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery* 2 (1998) 121–167.
- [24] V. Lukin, S. Abramov, S. Krivenko, A. Kurekin and O. Progrebnyak. Analysis of classification accuracy for pre-filtred multichannel remote sensing data. *Journal of expert systems with applications* 40 (2013) 6400–6411.