

# Robust Denoising Method Based on Tensor Models Decomposition for Hyperspectral Imagery

SALAH BOURENNANE  
École Centrale de Marseille,  
Institut Fresnel-CNRS UMR 7249  
Aix-Marseille Université  
FRANCE

salah.bourennane@centrale-marseille.fr

CAROLINE FOSSATI  
École Centrale de Marseille,  
Institut Fresnel-CNRS UMR 7249  
Aix-Marseille Université  
FRANCE

*Abstract:* In the hyperspectral images (HSI) acquired by the new-generation hyperspectral sensors the signal dependent noise is an important limitation to the detection or classification. Therefore, noise reduction is an important preprocessing step to analyze the information in the hyperspectral image (HSI). A signal dependent noise cannot be reduced by conventional linear filtering. Therefore, a new method based on multiple linear regression (MLR) and 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 and classification applications.

*Key-Words:* Hyperspectral Imagery, Denoising, Detection, Classification, Hyperspectral Sensor

## 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 and classification [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 [2–4], the additive SI noise is still an important part of noise [5]. 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 [6] an au-

tomatic 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, [7] 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 [8] 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 is used for the estimation of the SD and SI noise parameters in the second step [9], 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 [10] 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 two-step method to remove these two types of noise respectively. Firstly, by considering the SI noise as the residual image obtained by MLR method [8]. Then, the SD noise could be further reduced by PARAFAC decomposition [11, 12] due to the statistical property of SD noise. So the proposed MLR-PARAFAC hybrid method can denoise HSIs distorted by both SD and colored SI noise. 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 MLR-PARAFAC to other methods, prewhitening-multiway Wiener filter (PMWF) method proposed to reduce colored SI noise [13], PARAFAC, HYNPE method, SSATV method and two well-known 2D denoising methods, minimum noise fraction (MNF) and noise-adjusted principal components analysis (PCA) [14], are used in the experiment section. The experimental results show that the proposed method can estimate the variance of the noise in each band accurately and 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 multiple linear regression and multilinear algebra decomposition for enhancement of SNR. Some denoising and comparative results are contained in Section 5. Finally, section 6 concludes the paper.

## 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  $\mathbf{X} \in$

$\mathbb{R}^{I_1 \times I_2 \times I_3}$  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  $\mathbf{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$  being rank 1 means that it can be written as the outer product [15] of 3 vectors, that is:  $\mathbf{X} = \mathbf{a}^{(1)} \circ \mathbf{a}^{(2)} \circ \mathbf{a}^{(3)}$ . So, each element of  $\mathbf{X}$  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  $\mathbf{X}_n \in \mathbb{R}^{I_n \times M_n}$  ( $n = 1, 2, 3$ ) denotes the  $n$ -mode unfolding matrix of a tensor  $\mathbf{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ , 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  $\mathbf{X}_n$  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  $\mathbf{X} \in \mathbb{R}^{I_1 \times \dots \times I_N}$  and a matrix  $\mathbf{B} \in \mathbb{R}^{J \times I_n}$  in mode  $n$ . It leads to the tensor  $\mathbf{U} = \mathbf{X} \times_n \mathbf{B}$  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 [15]. For instance, tensor  $\mathbf{Y} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$  can be expressed as

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

where  $K_s$  is the rank,  $\hat{\mathbf{Y}}$  is the rank- $K_s$  PARAFAC approximation of  $\mathbf{Y}$ ;  $\mathbf{Y}_k \in \mathbb{R}^{I_1 \times I_2 \times I_3}$  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  $\mathbf{Y}$  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{\mathbf{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 \mathbf{R} \in \mathbb{R}^{I\_1 \times I\_2 \times I\_3} composed of a multidimensional signal \mathbf{X} \in \mathbb{R}^{I\_1 \times I\_2 \times I\_3} impaired by an additive random noise \mathbf{N}(\mathbf{X}) \in \mathbb{R}^{I\_1 \times I\_2 \times I\_3}. The tensor \mathbf{R} can be expressed as [16]:

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

where \mathbf{N}(\mathbf{X}) accounts for both SI and SD noise [17] 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 \sigma\_{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 \sigma\_{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:

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

and Eq.(3) can be correspondingly rewritten as

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

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 \mathbf{N}(\mathbf{X}) in band i\_3 of the HSI could be written as [2, 9]

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

where \mu\_{i\_3} \triangleq 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\_3th band of \mathbf{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 \mathbf{R} \in \mathbb{R}^{I\_1 \times I\_2 \times I\_3} (with M\_3 = I\_1 I\_2) can be expressed as :

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

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

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

with \mathbf{N}\_{SD}(\mathbf{X})\_3 and \mathbf{W}\_3 being the 3-mode unfolding matrices of \mathbf{N}\_{SD}(\mathbf{X}) and \mathbf{W} respectively. Using the mean noise variance of the i\_3th spectral band defined as 1/(I\_1 I\_2) \sum\_{i\_1=1}^{I\_1} \sum\_{i\_2=1}^{I\_2} \sigma\_{\mathbf{N}(\mathbf{X}), 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\_3th band of \mathbf{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}(\mathbf{X})\_3 can be expressed as:

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

## 4 Proposed method : Noise estimation

### 4.1 Signal Independent Noise Estimation Method Based on Multiple Linear Regression

Hyperspectral sensors measure the radiance from the observed scene in many spectral bands very close in wavelengths, thus the signal \mathbf{X} is generally characterized as having strong spectral correlation [14, 18]. By supposing that the noise in HSI is spectrally uncorrelated and the signal has strong spectral correlation, the signal \mathbf{X} could be estimated by resorting to the well-known MLR based approach [8] which exploits the strong spectral correlation of the signal and the weak between band correlation of the SI noise in HSIs. After estimating the pixel values in each band as a linear combination of the pixel values in the remaining I\_3 - 1 bands, the noise is reduced as a result.

For the unfolding matrix \mathbf{R}\_3 \in \mathbb{R}^{I\_3 \times M\_3} of the HSI data tensor \mathbf{R} \in \mathbb{R}^{I\_1 \times I\_2 \times I\_3} (with M\_3 = I\_1 I\_2), the M\_3-dimensional row vector is defined as \mathbf{r}\_{i\_3} = [r\_{i\_3,1}, r\_{i\_3,2}, \dots, r\_{i\_3, M\_3}] with i\_3 = 1, \dots, I\_3. Considering the (I\_3 - 1 \times 1) vector \mathbf{z}\_m^{(i\_3)} = [r\_{1,m}, \dots, r\_{i\_3-1,m}, r\_{i\_3+1,m}, \dots, r\_{I\_3,m}]^T with m = 1, \dots, M\_3 and the (I\_3 - 1 \times M\_3) matrix \mathbf{Z}\_{i\_3} = [\mathbf{z}\_1^{(i\_3)}, \mathbf{z}\_2^{(i\_3)}, \dots, \mathbf{z}\_{M\_3}^{(i\_3)}], the MLR-based approach obtains an estimate \hat{\mathbf{r}}\_{i\_3} of \mathbf{r}\_{i\_3} as a linear function of \mathbf{Z}\_{i\_3}: \hat{\mathbf{r}}\_{i\_3} = \mathbf{c}\_{i\_3}^T \mathbf{Z}\_{i\_3}, where \mathbf{c}\_{i\_3} is the regression vector of size I\_3 - 1 \times 1. The least squares estimator of \mathbf{c}\_{i\_3} is given by \mathbf{c}\_{i\_3} = (\mathbf{Z}\_{i\_3} \mathbf{Z}\_{i\_3}^T)^{-1} \mathbf{Z}\_{i\_3} \mathbf{r}\_{i\_3}^T. Then the signal after noise reduction by multiple linear regression method is estimated by \hat{\mathbf{r}}\_{i\_3} = \mathbf{c}\_{i\_3}^T \mathbf{Z}\_{i\_3}, and the reduced noise can be

calculated by  $\hat{\mathbf{w}}_{i_3} = \mathbf{r}_{i_3} - \hat{\mathbf{r}}_{i_3} = \mathbf{r}_{i_3} - \mathbf{c}_{i_3}^T \mathbf{Z}_{i_3}$ . Thus, the estimated tensor  $\hat{\mathbf{R}}$  could be rebuilt by the estimated unfolding matrix  $\hat{\mathbf{R}}_3$  constructed by  $\hat{\mathbf{r}}_{i_3}$  with  $i_3 = 1, \dots, I_3$ , i.e.,  $\hat{\mathbf{R}}_3 = [\hat{\mathbf{r}}_{i_1}^T, \dots, \hat{\mathbf{r}}_{i_3}^T]^T$ , while a considerable part of the SD noise is still remained in  $\hat{\mathbf{R}}$ . However, an estimate of SI noise is

$$\hat{\mathbf{W}} = \mathbf{R} - \hat{\mathbf{R}} \quad (11)$$

with

$$\hat{\mathbf{R}} = \mathbf{X} + \mathbf{N}_{SD}(\mathbf{X}) \quad (12)$$

## 4.2 Signal Dependent Noise Estimation Method Based on PARAFAC Decomposition Model

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

$$\mathbf{R}_a = \sum_{k=1}^{K_s} \lambda_k \mathbf{a}_k^{(1)} \circ \mathbf{a}_k^{(2)} \circ \mathbf{a}_k^{(3)} \quad (13)$$

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

$$\begin{aligned} \hat{\mathbf{R}} &= \sum_{k=1}^{K_s} \lambda_k \mathbf{a}_k^{(1)} \circ \mathbf{a}_k^{(2)} \circ \mathbf{a}_k^{(3)} \quad (14) \\ &+ \sum_{k=K_s+1}^K \lambda_k \mathbf{a}_k^{(1)} \circ \mathbf{a}_k^{(2)} \circ \mathbf{a}_k^{(3)} = \mathbf{R}_a + \mathbf{M} \end{aligned}$$

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

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

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

According to the definition in Eq. (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

$$\mathbf{a}_k^{(1)} \circ \mathbf{a}_k^{(2)} \circ \mathbf{a}_k^{(3)} = \quad (15)$$

$$\frac{a_{i,k}^{(1)} a_{j,k}^{(2)} a_{p,k}^{(3)}}{\left( \sum_{i=1}^{I_1} |a_{i,k}^{(1)}|^2 \right)^{1/2} \left( \sum_{j=1}^{I_2} |a_{j,k}^{(2)}|^2 \right)^{1/2} \left( \sum_{p=1}^{I_3} |a_{p,k}^{(3)}|^2 \right)^{1/2}}$$

$$= \frac{a_{i,k}^{(1)} a_{j,k}^{(2)} a_{p,k}^{(3)}}{\left( \sum_{i=1}^{I_1} \sum_{j=1}^{I_2} \sum_{p=1}^{I_3} |a_{i,k}^{(1)} a_{j,k}^{(2)} a_{p,k}^{(3)}|^2 \right)^{1/2}}$$

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

Therefore,  $\min \|\hat{\mathbf{R}} - \mathbf{R}_a\|^2 = \min \sum_{k=K_s+1}^K \lambda_k^2 \implies \{\lambda_k \mid k = K_s + 1, \dots, K\}$  are  $(K - K_s)$  smallest terms among  $\{\lambda_k \mid k = 1, \dots, K\}$ . Therefore, the minimum of the square error  $\|\hat{\mathbf{R}} - \mathbf{R}_a\|^2$  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.,

$$\mathbf{R}_a = \sum_{k=1}^{K_s} \lambda_k \mathbf{a}_k^{(1)} \circ \mathbf{a}_k^{(2)} \circ \mathbf{a}_k^{(3)} \approx \hat{\mathbf{X}} \quad (16)$$

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

$$\hat{\mathbf{R}} - \mathbf{R}_a = \hat{\mathbf{R}} - \hat{\mathbf{X}} = \hat{\mathbf{N}}_{SD}(\mathbf{X}) \quad (17)$$

is an estimate of the noise  $\mathbf{N}_{SD}(\mathbf{X})$ . Finally, by using Eq. (11) and Eq. (17), an estimate of the non-white noise can be expressed as

$$\hat{\mathbf{N}}(\mathbf{X}) = \hat{\mathbf{N}}_{SD}(\mathbf{X}) + \hat{\mathbf{W}} \quad (18)$$

## 4.3 Obtaining the Optimal Rank $K_s$ of PARAFAC Decomposition

In this paper, we assume that the statistical properties of both the signal and the remaining SD noise after the SI noise reduction are relatively constant in the estimated tensor  $\hat{\mathbf{R}}$  and the variance of the SD noise in Eq. (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  $\hat{\mathbf{R}}$ , respectively. Therefore, the covariance matrix  $\mathbf{C}_{\hat{\mathbf{N}}_{SD}(\mathbf{X})}^{(3)}$  of remaining SD noise is still a diagonal matrix. If the squared norm of the covariance  $\|\mathbf{C}_{\hat{\mathbf{N}}_{SD}(\mathbf{X})}^{(3)}\|^2$  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  $\mathbf{C}_{\hat{\mathbf{N}}_{SD}(\mathbf{X})}^{(3)}$  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 SD noise.

#### 4.4 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. Unfold tensor  $\mathbf{R}$  to matrix  $\mathbf{R}_n = \mathbf{X}_n + \mathbf{N}_n(\mathbf{X})$ ,  $n = 1, 2, 3$ , estimating the SI noise  $\mathbf{W}$  from  $\mathbf{R}$  using MLR, and obtain  $\hat{\mathbf{R}}$
2. Estimate the SD noise  $\mathbf{N}_{SD}(\mathbf{X})$  from  $\hat{\mathbf{R}}$  using PARAFAC:
  - (a) Unfold the estimated noise tensor  $\mathbf{N}_{SD}(\mathbf{X})$  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 SD noise could be obtained from these three steps.
3. Calculate the noise tensor  $\hat{\mathbf{N}}(\mathbf{X}) = \hat{\mathbf{N}}_{SD}(\mathbf{X}) + \hat{\mathbf{W}}$
4. Estimate the signal tensor  $\hat{\mathbf{X}} = \mathbf{R} - \hat{\mathbf{N}}(\mathbf{X})$

## 5 Experimental results

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

$$\text{PSNR} = 10 \log_{10} \frac{(\max\{\mathbf{X}\})^2}{\text{MSE}} \quad (\text{dB}) \quad (19)$$

where

$$\text{MSE} = 1/(I_1 I_2 I_3) \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \sum_{i_3=1}^{I_3} (x_{i_1 i_2 i_3} - \hat{x}_{i_1 i_2 i_3})^2 \quad (20)$$

In addition, the variance of the residual noise,  $\hat{\mathbf{X}} - \mathbf{X}$ , in the simulated HSI and the variance of the removed noise,  $\mathbf{R} - \hat{\mathbf{X}}$ , in the real data are compared in this

paper after denoising by different considered methods: MLR-PARAFAC, PARAFAC, PMWF, HYNPE, 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 Eq. (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  $7 \times 7$  pixels,  $2 \times 2$  pixels and  $1 \times 1$  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 Eq. (7) and added into the signal  $\mathbf{X}$  as Eq. (4) to create the simulated HSI data  $\mathbf{R}$ . 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.

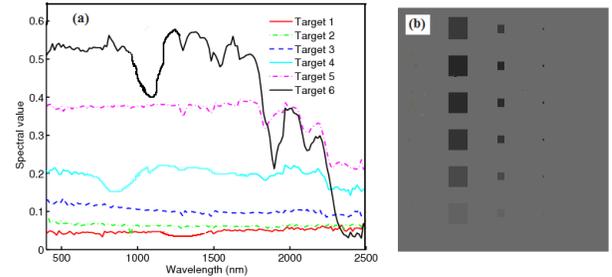


Figure 1: (a) Spectral signatures of the simulated targets and background, (b) Simulated HSI without noise, from top to bottom the index of targets is 1 to 6 respectively

#### 5.1.1 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 ( $\sigma_{\mathbf{N}_{SD}(\mathbf{X})}^2 = \sigma_w^2$ ), thus the noise variance of  $\mathbf{N}(\mathbf{X})$  in this case varies strongly from band to band. The variance of the residual noise,  $\hat{\mathbf{X}} - \mathbf{X}$ , in the denoised HSI is evaluated at each band and some results are shown in Fig. 2 (b) where the noise variance of  $\mathbf{N}(\mathbf{X})$  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 MLR-PARAFAC method contains the least noise when compared against other considered methods.

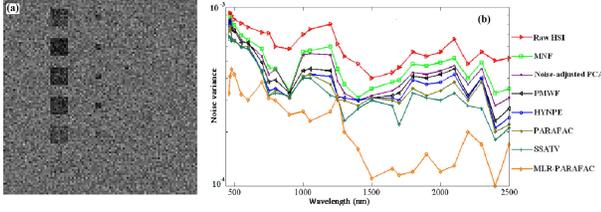


Figure 2: (a) Noisy image with SNR =30 dB, (b) Noise variance of raw simulated HSI and residual noise variances of denoised HSI

### 5.1.2 Influence of SD and SI Noise Variance Values on the Performance of Denoising Methods

The PSNR index is used to give a quantitative assessment of the results obtained by applying different denoising methods to the simulated HSI. In this experiment, the random noise is generated with different variance values depending on the mean value of the useful signal according to Eq. (7) and added into the signal  $\mathbf{X}$  as Eq. (4) to create the simulated HSI data  $\mathbf{R}$ . To set the values of the two noise variances  $\sigma_{N_{SD}(\mathbf{X})}^2 = 1/I_3 \sum_{i_3=1}^{I_3} \sigma_{u,i_3}^2 \cdot \mu_{i_3}$  and  $\sigma_{w,i_3}^2$ , we define some quantities: The signal to noise ratio (SNR) of the synthetic noisy HSI

$$\text{SNR} = 10 \log_{10} \frac{\|\mathbf{X}\|^2}{\|\mathbf{N}(\mathbf{X})\|^2} \quad (\text{dB}) \quad (21)$$

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

$$P = 1/(I_1 I_2 I_3) \sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \sum_{i_3=1}^{I_3} x_{i_1, i_2, i_3}^2. \quad (22)$$

Assuming

$$\delta = \sigma_{N_{SD}(\mathbf{X})}^2 / \sigma_w^2, \quad (23)$$

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

$$\begin{aligned} \sigma_{N_{SD}(\mathbf{X})}^2 &= \frac{\delta \cdot \sigma_{\mathbf{N}(\mathbf{X})}^2}{1 + \delta} \\ \sigma_w^2 &= \frac{\sigma_{\mathbf{N}(\mathbf{X})}^2}{1 + \delta} \end{aligned} \quad (24)$$

With the definitions above, it is clear that for  $\delta = 1$ ,  $\sigma_{N_{SD}(\mathbf{X})}^2 = \sigma_w^2$ , 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  $\delta < 1$ ,  $\sigma_w^2$  is higher than  $\sigma_{N_{SD}(\mathbf{X})}^2$  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. The effectiveness of the proposed MLR-PARAFAC hyperspectral image denoising algorithm can be illustrated with quantitative assessment results shown in Table 1, it can be clearly seen that the the PSNR values using the MLR-PARAFAC model are higher than the PSNR using the other methods for all tested values of  $\delta$ .

Table 1: PSNR (DB) OBTAINED BY APPLYING DIFFERENT DENOISING METHODS TO SIMULATED HSIS WITH SNR=30DB

$\delta$	0.1	0.5	1.0	1.5	2.0
MNF	31.02	31.04	31.05	31.06	31.06
Noise-adj. PCA	31.12	31.14	31.14	31.15	31.16
HYNPE	36.64	36.82	36.90	37.10	37.10
PMWF	31.95	31.99	31.99	32.08	32.09
PARAFAC	38.71	38.84	38.85	38.88	38.88
SSATV	39.22	39.28	39.29	39.33	39.35
MLR-PARAFAC	43.05	43.09	43.11	43.25	43.29

### 5.1.3 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) [19] which is largely applied to HSI data. The results of ACE target detection of both simulated and real data denoised by our proposed MLR-PARAFAC method and other considered methods are shown and discussed. For the HSI described in Eq. (3), the ACE detector can be expressed as:

$$\text{ACE} = \frac{(\mathbf{s}^T \Gamma^{-1} \mathbf{r}_j)^2}{(\mathbf{s}^T \Gamma^{-1} \mathbf{s})(\mathbf{r}_j^T \Gamma^{-1} \mathbf{r}_j)} \quad (25)$$

where  $\mathbf{r}_j$  is the vector in the unfolding matrix  $\mathbf{R}_3$  of tensor  $\mathbf{R}$  with  $j = 1, \dots, I_1 I_2$ ,  $\mathbf{s}$  is the target spectrum template  $\Gamma$  is the covariance matrix of  $\mathbf{R}_3$ . To assess

the performance of detection, the probability of detection (PD) is defined as:

$$PD = \frac{\sum_i^{n_s} N_i^{rd}}{\sum_i^{n_s} N_i}, \quad (26)$$

and the probability of false alarm (PFA) is defined as:  $PFA = \frac{\sum_i^{n_s} N_i^{fd}}{\sum_i^{n_s} (I_1 \times I_2 - N_i)}$ , 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. 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 MLR-PARAFAC method outperforms other methods. According to subsection 4.2, the SD noise is removed by throwing away other smallest terms from  $K_s + 1$  to  $K$ , MLR removes SI noise, thus the quality of the denoised HSI by MLR-PARAFAC is ameliorated so much that the ACE target detection of the denoised HSI is improved greatly.

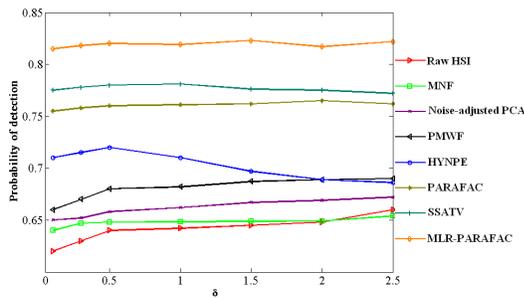


Figure 3: Probability of detection of simulated HSI denoised by different methods with SNR= 30dB and for PFA=  $10^{-4}$

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

## 5.2 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 Eq. (3)-(7), two real-world images are considered in this section. The first one, referred to as HYDICE HSI, was acquired by HYperspectral Digital Imagery Collection Experiment (HYDICE). The second one, referred to as AVIRIS HSI was collected by the airborne visible/infrared imaging spectrometer (AVIRIS) from a mixed forest/agricultural site at the Indian Pines test site in north-west Indiana.

### 5.2.1 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  $\mathbf{R} \in \mathbb{R}^{150 \times 140 \times 148}$ . 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  $5 \times 5$  pixels along the first column,  $3 \times 3$  pixels along the second one and  $1 \times 1$  pixel along the last one. To assess the denoising results obtained by the proposed method, the removed noise  $(\mathbf{R} - \hat{\mathbf{X}})$  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.

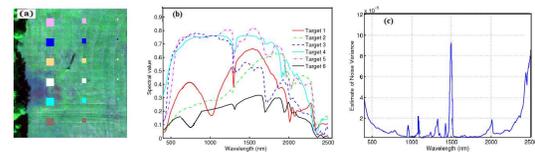


Figure 4: (a) HYDICE HSI used to compare different denoising methods, from top to bottom the index of targets is 1 to 6 respectively. (b) Spectral signatures (spectral reflectances) of the simulated targets. (c) Noise variance of raw HYDICE HSI

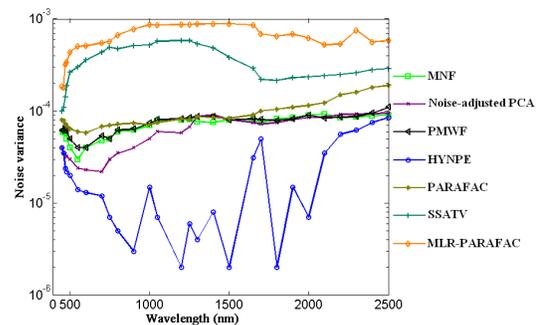


Figure 5: Variance of removed noise from HYDICE HSI with different methods

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.

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 MLR-PARAFAC method are improved significantly when compared against

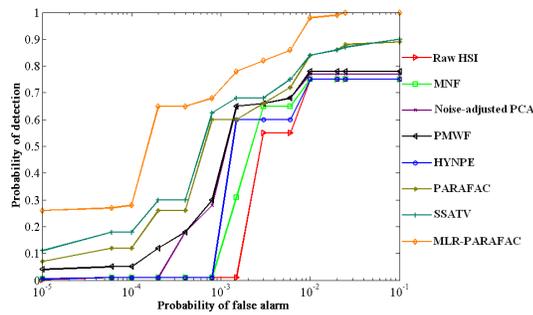


Figure 6: ROC curves of the denoised HYDICE HSIs obtained by ACE detector

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. But the improvement of target detection of the HYDICE HSI denoised by PARAFAC method is inferior to that by SSATV. Since the SSATV 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, 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.

In the next section, the classification is used to evaluate the performance of the different denoising methods.

### 5.2.2 Classification Results

To appreciate quantifiable comparisons, we determine the overall accuracy (OA) in percentage exhibited by SVM classifier [20, 21]. For  $P$  classes  $C_i$ ,  $i = 1, \dots, P$ ; if  $a_{ij}$  is the number of testing samples that actually belong to class  $C_i$  and are classified into  $C_j$  for  $i, j = 1, \dots, P$ , then OA is defined as follows:  $OA = \frac{1}{N_{\text{total}}} \sum_{i=1}^P a_{ii}$ , where  $N_{\text{total}}$  is the total number of samples,  $P$  is the number of classes  $C_i$  for  $i = 1, \dots, P$  and  $a_{ii}$  is equal to  $a_{ij}$  for  $i = j$ . That is, in this paper, OA is defined as:

$$OA = \frac{N_{\text{correct}}}{N_{\text{total}}} \times 100\% \quad (27)$$

where  $N_{\text{correct}}$  is the number of testing samples classified correctly into their corresponding classes. The

higher the OA, the better the classification result. For this investigation, AVIRIS HSI (Fig. 7 (a)) is used [8]. The raw image size is  $145 \times 145 \times 220$ . This HSI can be represented as a tensor  $\mathbf{R} \in \mathbb{R}^{145 \times 145 \times 220}$  and its ground truth is shown in Fig. 7 (b), which is supplied with the original data. According to the ground truth, there are 16 land cover classes in this AVIRIS HSI. The numbers of training and testing samples are shown in TABLE 2. Note that the minimum number of training samples is set to 10 for the rare classes such as Grass/pasture-mowed.

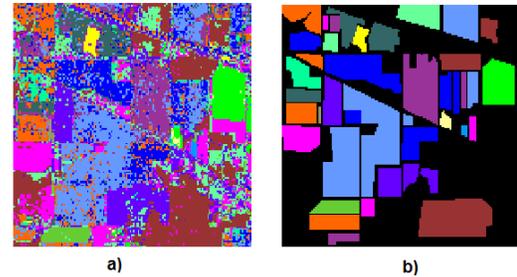


Figure 7: a) Raw AVIRIS HSI. b) Ground truth of the used AVIRIS HSI

Table 2: TRAINING AND TESTING SAMPLES OF CLASSIFICATION

ID	Class	Train. Samples	Test. Samples
1	Corn-min	84	846
2	Hay-windrowed	48	479
3	Stone-st. towers	10	98
4	Woods	126	1264
5	Wheat	22	222
6	Soybean-clean	60	604
7	Oats	10	22
8	Soybean-notill	96	988
9	Corn	23	234
10	Bldg-Tree-Drives	38	372
11	Alfalfa	10	56
12	Corn-notill	142	1421
13	Grass/Trees	73	737
14	Grass/Pasture	53	533
15	Grass/past.-mowed	11	28
16	Soybeans-min	246	2476

To evaluate the performance of the proposed algorithm in improving the classification efficiency, the classification overall accuracy obtained after denoising by different methods is presented in Fig. 8. As demonstrated by the previous experimental results, it is also evident from the results portrayed

in Fig. 8 that the OA of denoised image by the proposed method is much higher than other processes for all considered SNR values. MLR-PARAFAC method reduces both SI and SD noise as shown in the previous section, thus it performs well in the improvement of classification from low to high SNR cases. While PMWF becomes effective when the input SNR is high ( $\text{SNR} > 30$  dB). Because PMWF is the combination of pre-whitening procedure and MWF filter, MWF is based on the estimation of the signal subspace by using the largest singular values of the  $n$ -mode ( $n = 1, 2, 3$ ) unfolding matrices [16]. Therefore, in the denoising process, MWF can preserve more signal details for high SNR, which is important to distinguish two close classes that which permits to PMWF to exhibit an improved OA value. When  $\text{SNR} < 30$  dB, almost all the OA values of SVM classifier of the denoised HSIs by HYNPE, MNF and Noise-adjusted PCA methods are lower than those by other methods due to the imperfect denoising result of these methods in the reduction of SD noise. Because, in the case of low SNR the spectral correlation of the signal becomes weaker and this correlation is further decreased by the stronger noise in HSI. Since these methods seek to exploit the spectral correlation of the signal, so the effect of noise reduction by these methods is decreased by the weak correlation among bands of the signal and the classification of HSIs denoised by these methods are affected correspondingly. When  $\text{SNR} > 30$  dB, the classification results after denoising by these methods are better.

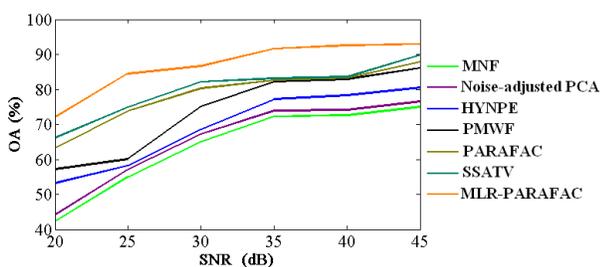


Figure 8: Classification OA for different values of SNR

It can be observed from Fig. 2 (b), Fig. 3, Fig. 5, Fig. 6, Fig. 8 and Table 1 that MLR-PARAFAC method has the potential to denoise real-world HSI images distorted by both SI and SD noise. This pre-processing permits an enhancement of SNR in HSI thereby allowing both an improvement of results of ACE target detector and SVM classifier.

## 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 which consists of two steps. Firstly, MLR algorithm can be applied to remove the SI colored noise. Then, to reduce the residual SD components, PARAFAC decomposition is applied to the denoised HSI by the previous step. PARAFAC decomposition must be conducted at the appropriate rank which can be estimated according to the statistical properties of SD noise. The performance of the proposed MLR-PARAFAC method are validated on the simulated HSIs distorted by both SD and SI noise and on the real-world HYDICE and AVIRIS Indian Pines HSIs. The HYDICE dataset is used to evaluate the denoising and detection results when the AVIRIS dataset of Indian Pines is used to assess classification results. The experimental results show the efficiency of the proposed denoising algorithm to improve the SNR in hyperspectral images, the detection and the classification results.

From the analysis and the comparative study against other similar methods in the experiments, it can be concluded that MLR-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 or AVIRIS HSI data.

### 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, pp. 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, pp. 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, pp. 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, pp. 1383–1395.
- [5] P. Zhong and R. Wang. Multiple-spectral-band CRFs for denoising junk bands of hyperspectral imagery. *IEEE Trans. Geoscience and Remote Sensing*, 4, 2013, pp. 2269–2275.
- [6] B. Gao. An operational method for estimating signal to noise ratios from data acquired with imaging spectrometers. *Remote Sens. Environ.*, 43, 1993, pp. 23–33.
- [7] B. Corner, R. Narayanan, and S. Reichenbach. Noise estimation in remote sensing imagery using data masking. *International Journal Remote Sensing*, 24, 2003, pp. 689–702.
- [8] R. E. Roger and J. F. Arnold. Reliably estimating the noise in AVIRIS hyperspectral images. *International Journal of Remote Sensing*, 17, 1996, pp. 1951–1962.
- [9] 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, pp. 1–4.
- [10] 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, pp. 3660–3677.
- [11] R. A. Harshman and M. E. Lundy. The parafac model for three-way factor analysis and multidimensional scaling. *Research methods for multi-mode data analysis*. Praeger, 1984, pp. 122–215.
- [12] R. A. Harshman and M. E. Lundy. Parafac: Parallel factor analysis. *Computational statistics and data analysis*, 18, 1994, pp. 39–72.
- [13] X. Liu, S. Bourennane, and C. Fossati. Nonwhite noise reduction in hyperspectral images. *IEEE Geoscience and Remote Sensing letters*, 9, 2012, pp. 368–372.
- [14] C.-I. Chang and Q. Du. Interference and noise adjusted principal components analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 37, 1999, pp. 2387–2396.
- [15] T. G. Kolda and B. W. Bader. Tensor decompositions and applications. *SIAM Review*, 51, 2009, pp. 455–500.
- [16] D. Letexier and S. Bourennane. Noise removal from hyperspectral images by multidimensional filtering. *IEEE Transactions on Geoscience and Remote Sensing*, 46, 2008, pp. 2061–2069.
- [17] 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, pp. 3829–3846.
- [18] R. E. Roger. Principal components transform with simple, automatic noise adjustment. *International Journal of Remote Sensing*, 17, 1996, pp. 2719–2727.
- [19] D. Manolakis and G. Shaw. Detection algorithms for hyperspectral imaging applications. *IEEE Signal Processing Magazine*, 19, 2002, pp. 29–43.
- [20] C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2, 1998, pp. 121–167.
- [21] V. Lukin, S. Abramov, S. Krivenko, A. Kurekin and O. Progrebnyak. Analysis of classification accuracy for pre-filtered multichannel remote sensing data. *Journal of expert systems with applications*, 40, 2013, pp. 6400–6411.