

Research Article

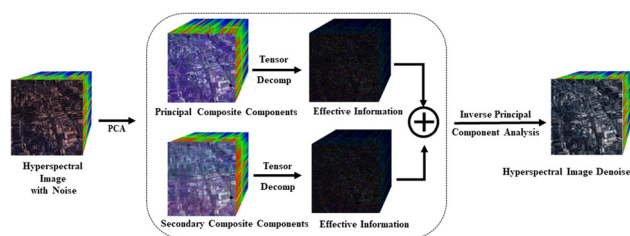
Hao Wu, Ruihan Yue*, Ruixue Gao, Rui Wen, Jun Feng, and Youhua Wei

Hyperspectral denoising based on the principal component low-rank tensor decomposition

<https://doi.org/10.1515/geo-2022-0379>

received October 27, 2021; accepted May 01, 2022

Abstract: Due to the characteristics of hyperspectral images (HSIs), such as their high spectral resolution and multiple continuous narrow bands, HSI technology has become widely used in fields such as target recognition, environmental detection, and agroforestry detection. HSIs are subject, for various reasons, to noise in the processes of data acquisition and transmission. Therefore, the denoising of HSIs is very necessary and important. In this article, according to the characteristics of HSIs, an HSI denoising model combining principal component analysis (PCA) and CANDECOMP/PARAFAC decomposition (CP decomposition) is proposed, which is called PCA-TensorDecomp. First, we use PCA to reduce the dimension of HSI signals by obtaining the first K principal components and get the principal composite components. The low-rank part corresponding to the first K principal components is considered the characteristic signal. Then, low-rank CP decomposition is carried out, to denoise the first principal components and the remaining minor components, the secondary composite components, which contain a large amount of noise. Finally, the inverse PCA is then used to restore the HSIs denoised, such that the effect of comprehensive denoising is achieved. To test the effectiveness of the improved algorithm introduced in this article, we compare it with several methods on simulated and real hyperspectral data. The



Graphical abstract

results of the analysis herein indicate that the proposed algorithm possesses a good denoising effect.

Keywords: hyperspectral denoising, principal component, low-rank, tensor decomposition

1 Introduction

In recent years, the rapid development of hyperspectral image (HSI) technology has provided abundant spectral spatial features for a wide range of applications, such as target detection and classification [1]. An HSI is equivalent to the superposition of many grayscale images. Therefore, the HSI can be regarded as a cube structure image, which contains rich information. On one hand, it is equivalent to a two-dimensional spatial image, from which spatial dimension information can be obtained. On the other hand, it also contains the spectral structure of the observed target.

In the process of imaging and transmission, HSIs are inevitably polluted by noise from sensors, the atmosphere, and other sources, such that the quality of HSIs may be seriously reduced, which undoubtedly has adverse effects on the extraction of HSI information. Therefore, as a key preprocessing step to improve the subsequent recognition ability, HSI noise reduction has attracted the attention of many scholars and has become a hot topic in the field of hyperspectral remote sensing in recent years. Compared with panchromatic and multi-spectral imagery, the noise distribution of hyperspectral imagery is more complex. As a result, traditional gray image denoising

* **Corresponding author: Ruihan Yue**, College of Mathematics and Physics, Chengdu University of Technology, Chengdu, 610059, China; Geomathematics Key Laboratory of Sichuan Province, Chengdu University of Technology, Chengdu, 610059, China, e-mail: yueruihan@stu.cdu.edu.cn

Hao Wu: College of Management Science, Chengdu University of Technology, Chengdu, 610059, China; Geomathematics Key Laboratory of Sichuan Province, Chengdu University of Technology, Chengdu, 610059, China

Ruixue Gao, Rui Wen, Jun Feng, Youhua Wei: College of Mathematics and Physics, Chengdu University of Technology, Chengdu, 610059, China; Geomathematics Key Laboratory of Sichuan Province, Chengdu University of Technology, Chengdu, 610059, China

methods, such as K-SVD [2], block matching, and 3D filtering [3], may not always be good at removing noise from HSIs. Therefore, a reasonable and effective denoising method should be designed, according to the characteristics of HSIs.

In recent years, numerous denoising methods for HSIs have been proposed, most of which have focused on enhancing the restoration effect in the spatial and spectral domains. Denoising methods can be roughly divided into two types: The first type refers to two-dimensional image denoising methods, which introduces the spatial spectrum characteristics in different transform domains for HSI denoising, such as wavelet transforms [4], principal component analysis (PCA) [5], and sparse 3D transform-domain collaborative filtering (BM3D) method [3]. However, such methods only consider spatial domain denoising, and they do not use spectral information to attenuate noise. Thus, there generally still exists some noise, which could not be removed by such methods. HSI denoising must consider both spatial and spectral domains. The second type of methods consists of denoising in the spectral and spatial domains in different stages. To better reduce the noise in homogeneous regions while maintaining edge and texture information, Yuan et al. [6] have proposed a spectral adaptive total variation model to smooth the noise and signals. In addition, Sun and Luo [7] have proposed a three-dimensional mixed denoising method in the differential domain of HSIs, in which a spectral dimension transformation is first performed on the HSI, following which the inverse transformation is performed after denoising. Compared with the first type, methods of the second type take both spatial and spectral information into consideration in the denoising algorithm, such that the denoising effect can be significantly improved.

An HSI contains three dimensions: one spectral dimension and two spatial dimensions. It can be modeled as three-order tensor data. Therefore, tensor-based models can improve the denoising results, because they can fully capture the spatial-spectral correlation of the HSI. Recently, with the development of tensor theory, many tensor-based HSI denoising methods have emerged, such as the Tucker decomposition-based HSI denoising algorithm [8,9], but these methods often have the problem of low efficiency and ignore the nonlocal self-similarity prior. Zheng et al. [10] proposed a novel nonlocal patch-based FCTN (NL-FCTN) decomposition for MSI inpainting. At the same time, due to the characteristics of HSIs, the structured high-dimensional information inherent in the original observations may be discarded when the high-dimensional HSI data are converted into two-dimensional data by traditional methods and processed separately. Inspired by the noise estimation method proposed in ref. [11], Meng

et al. [12] proposed an HSI denoising method by jointly exploiting Tucker tensor decomposition and PCA based on the noise power ratio (NPR) analysis. With the NPR, they estimate the Tucker rank. And as this method exploits the correlation effects in all of the dimensions of HSI data, it gets some good results when denoising.

When using Tucker decomposition, the first step is to estimate the rank of Tucker decomposition. However, for a fixed N-rank image, the uniqueness of the Tucker decomposition is not guaranteed. Scholars have made some improvements on this basis, such as the kernel non-negative Tucker decomposition [13], non-negative Tucker decomposition based on non-local self-similarity in the spatial domain [14], and Tucker decomposition based on non-local self-similarity and global correlation across the spectrum [15]. Although these methods can better demonstrate the structural correlation of HSIs and reduce the artifacts of overlapping regions, they cannot solve the problem that the Tucker decomposition is not unique.

CANDECOMP/PARAFAC decomposition (CP decomposition), can be regarded as a special form of Tucker decomposition [16,17]. It implements the tensor approximation by manipulating the CP rank; thus, it has been used as an effective method. Liu et al. [18] proposed an HSI denoising method based on CANDECOMP/PARAFAC decomposition (CP decomposition) and estimated the CP rank depending on the variance of additive noise. However, in practice, the variance is often unknown. Therefore, some low-order models of joint local and nonlocal self-similarity can be used to improve the quality of denoising, such as the non-local image denoising model based on the weighted tensor rank-1 decomposition of Wu et al. and Guo et al. [19,20].

In addition to being widely used in the field of HSI denoising [21,22], low-rank priors can improve HSI super-resolution. Dian and Li [23] proposed a novel subspace-based low tensor multi-rank regularization method for the fusion, which makes full use of spectral correlations and non-local similarities of the HR-HSI. Xue et al. [24] propose a new HSI super-resolution method that fully considers the spatial/spectral subspace low-rank relationships between available high-resolution multispectral image/low-resolution HSI and potential HSI.

Another widely used HSI denoising method is tensor regularization [25,26]. Xue et al. [25] used the local regularization CP low-rank tensor decomposition method for HSI denoising. Zhang et al. [27] approximated the rank by a sum of tensor ranks in different directions, considering the validity of singular value decomposition (SVD) in the literature. Liu et al. [28] suggest a novel HSI restoration

model by introducing a fibered rank constrained tensor restoration framework with an embedded plug-and-play-based regularization. For HIS mixed noise, Zheng et al. [29] proposed a double-factor-regularized LRTF model for HSI mixed noise removal. Zheng et al. [30] proposed a fibered rank minimization model for HSI mixed noise removal, in which the underlying HSI is modeled as a low-fibered-rank component. The key problem in tensor regularization is how to construct a reasonable regularization term to reflect the global correlation and non-local similarity. Tensor regularization can introduce prior information into the HSI denoising and retain the spectral and spatial dimensional information. It can be seen that tensor regularization has a good application prospect for HSI denoising.

To make use of the advantage of tensor decomposition and tensor regularization and take the characteristics of HSIs into account, in this article, we propose an HSI denoising model combining PCA and CP decomposition, which we call PCA TensorDecomp.

In summary, the main contributions of this work are as follows:

- (1) A novel method proposed to solve HSI denoising by combining PCA and CP decomposition named PCA TensorDecomp.
- (2) Inspired by the alternating direction method of multipliers (ADMM) and alternate least squares, a new algorithm has been proposed to effectively solve the method proposed in this article.
- (3) A broad experimental analysis has been provided based on several datasets. The numerical outcomes are also corroborated by a qualitative analysis.

The remainder of this article is arranged as follows: Section 2 briefly introduces the principles of PCA and CP decomposition. In Section 3, the PCA TensorDecomp is proposed. In Section 4, some numerical experiments are provided and Section 5 is the conclusion.

2 Notations and related methods

2.1 Notations

In this article, lowercase or uppercase letters (e.g., i , $I \in \mathbb{R}$) are used to denote scalars. Vectors are represented by boldface lowercase letters, e.g., $\mathbf{x} \in \mathbb{R}^I$. Boldface capital letters (e.g., $\mathbf{X} \in \mathbb{R}^{I \times J}$) are employed to represent matrices. A tensor is a multi-dimensional data array, and an n -order tensor

($n \geq 3$) is represented by a calligraphic letter $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_n}$. $\mathbf{X}_k \in \mathbb{R}^I$ represents the k -th column vector of $\mathbf{X}$. An element value of $\mathcal{X}$ in position $(i_1, i_2, \dots, i_n)$ is represented as $x_{i_1 i_2 \dots i_n}$. The normal mode- k matricization of a tensor is represented as $\mathbf{X}_{(k)} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_n}$. Moreover, the Frobenius norm of a matrix $\mathbf{X}$ is calculated as $\|\mathbf{X}\|_F = \sqrt{\sum_i \sum_j x_{ij}^2}$.

2.2 PCA

HSI data are complex. HSIs consist of dozens or even hundreds of continuous and subdivided spectral bands imaging the target region at the same time, which provides us with abundant information and is conducive to the fine division of ground objects. The increase in band number not only leads to information redundancy but can also lead to a requirement for rather complex data processing, such that it is necessary to reduce the dimension of the HSI data. PCA can separate the principal and secondary components of HSIs well. On one hand, redundant data can be removed; on the other hand, essential information can be retained. It is, therefore, suitable for the complex data processing of HSIs.

The essence of PCA is to transform a group of related variables into a group of unrelated variables through an orthogonal transformation and it has been widely applied in many fields [5,31]. The formula is as follows:

$$\mathbf{Z} = \mathbf{A}\mathbf{X}, \quad (1)$$

where $\mathbf{X}$ is the original data, $\mathbf{Z} = (\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_p)^T$ is the principle components, and $\mathbf{A}$ is the orthogonal transformation to realize this process.

Considering that PCA cannot directly process three-dimensional data, such as HSIs, we carry out a preprocessing to reduce the dimension before conducting the PCA. The specific steps for dimensionality reduction of hyperspectral data are as follows.

The first step is to convert the HSI data $\mathcal{X}$ into two-dimensional data $\mathbf{X}$. Compressing each band into one-dimensional data, the multiple bands then form a data matrix.

The second step is to calculate the covariance matrix $\mathbf{S}$ of the data matrix $\mathbf{X}$. The specific formula is as follows:

$$\mathbf{S} = \frac{1}{n}[\mathbf{X} - \bar{\mathbf{X}}\mathbf{l}][\mathbf{X} - \bar{\mathbf{X}}\mathbf{l}]^T, \quad (2)$$

where $\mathbf{l} = [1, 1, \dots, 1]_{1 \times p}$, $\bar{\mathbf{X}} = [\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, \dots, \bar{\mathbf{x}}_p]_{1 \times p}^T$, and $\bar{\mathbf{x}}_i = \frac{1}{n} \sum_{k=1}^n x_{ik}$.

In the third step, with the covariance matrix $\mathbf{S}$ obtained, and in the second step, solve the eigenequation

$(\lambda \mathbf{I} - \mathbf{S})\mathbf{U} = 0$ and get all the eigenvalues of the matrix $\mathbf{S}$. Then, the corresponding eigenvectors should be found.

Finally, all the eigenvectors are sequentially arranged to form the orthogonal matrix $\mathbf{A}$.

Usually, with PCA of HSIs, the new principal composite components contain less and less information. Therefore, we mainly retain most of the relevant information with the first few principal components and discard the secondary composite components, which are the so-called noise signal components. However, those secondary composite components usually contain some useful information and, so, after discarding the secondary composite components, the reconstruction of HSIs will lead to some loss of the information. Therefore, how to extract information from the secondary composite components is a very critical problem.

2.3 CP decomposition

The essence of CP decomposition is to decompose third-order tensor data into a sum of R rank-1 tensors [16]. Note that HSIs can be treated as cube data (namely, a third-order tensor). Because of this property, CP decomposition can not only solve the problem of low-rank tensor completion [32] but also be used for HIS denoising.

Given a third-order tensor $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$, the CP decomposition formula is as follows:

$$\mathcal{X} \approx \sum_{r=1}^R \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r, \quad (3)$$

where R is an integer and $\mathbf{a}_r \in \mathbb{R}^I$, $\mathbf{b}_r \in \mathbb{R}^J$, and $\mathbf{c}_r \in \mathbb{R}^K$, $r = 1, 2, \dots, R$. According to the definition of the cross product, each element can also be expressed as:

$$x_{ijk} \approx \sum_{r=1}^R a_{ir} b_{jr} c_{kr}, \quad (4)$$

$i = 1, 2, \dots, I; \quad j = 1, 2, \dots, J; \quad k = 1, 2, \dots, K.$

Tensor $\mathcal{X}$ can be unfolded to form matrices. The combination of vectors from the rank-1 components, i.e., $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_R]$ is called the factor matrix. In this way, we can construct $\mathbf{B}$ and $\mathbf{C}$ similarly. Using these definitions, we can equivalently define CP decomposition in the matrix form as:

$$\begin{aligned} \mathbf{X}_{(1)} &\approx \mathbf{A}(\mathbf{C} \circ \mathbf{B})^T, \\ \mathbf{X}_{(2)} &\approx \mathbf{B}(\mathbf{C} \circ \mathbf{A})^T, \\ \mathbf{X}_{(3)} &\approx \mathbf{C}(\mathbf{B} \circ \mathbf{A})^T, \end{aligned} \quad (5)$$

where $\mathbf{X}_{(i)}$ represents mode- n matricization of tensor $\mathcal{X}$ and $\circ$ is the Khatri–Rao product.

In general, the column vectors of the matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are normalized. Let $\lambda \in \mathbb{R}$, then, the CP decomposition can also be expressed as follows:

$$\mathcal{X} = [[\lambda; \mathbf{A}, \mathbf{B}, \mathbf{C}]] \equiv \sum_{r=1}^R \lambda_r \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r, \quad (6)$$

where $\mathbf{a}_r \in \mathbb{R}^I$, $\mathbf{b}_r \in \mathbb{R}^J$, and $\mathbf{c}_r \in \mathbb{R}^K$, $r = 1, 2, \dots, R$.

CP decomposition requires two steps:

- (1) Determine the number of decomposition ranks; that is, the rank of the so-called tensor.

The rank of a tensor is different from that of a matrix, such that there is no simple method to solve the rank of CP decomposition. That is to say, solving the tensor rank is currently an NP-hard problem, and there is no definite method or algorithm, which is accurate for determining the decomposition rank. We can, however, choose several different values to compare their effect, or start from rank-1 and carry out an iterative approach. Of course, the tensor ranks of some dimensions have been obtained by some scholars, which provide us with a good reference in particular cases [33,34].

- (2) Once R , the rank of a tensor, has been determined, decomposition of the tensor can be performed.

When we solve the CP decomposition of a third-order tensor, it is equivalent to using the sum of R rank-1 tensors to approximate the original tensor $\mathcal{X}$. The objective function is as follows:

$$\begin{aligned} \min_{\hat{\mathcal{X}}} \|\mathcal{X} - \hat{\mathcal{X}}\|, \\ \text{s.t. } \hat{\mathcal{X}} = \sum_{r=1}^R \lambda_r \mathbf{a}_r \circ \mathbf{b}_r \circ \mathbf{c}_r = [[\lambda; \mathbf{A}, \mathbf{B}, \mathbf{C}]], \end{aligned} \quad (7)$$

where $\hat{\mathcal{X}}$ is the tensor to be optimized.

Equation (7) can be solved by the ADMM [35]. The optimization principle of the ADMM optimization algorithm is to first fix the other parameters, except for the one to be optimized. In the above formula, the specific process is fixed first and then optimized. The operation is repeated until the iteration has reached a certain number or the error gets to the requirement of the experiment. The specific optimization processes of different parameters are similar, so the following only introduces the optimization process for the parameter $\mathbf{A}$.

As all of the matrices are fixed, except for one, the problem becomes a linear least-squares problem. Here, we optimize for parameter $\mathbf{A}$, with $\mathbf{B}$ and $\mathbf{C}$ fixed. According to equation (5), we can write the above minimization problem, in the matrix form, as:

$$\min_{\hat{\mathbf{A}}} \|\mathbf{X}_{(1)} - \hat{\mathbf{A}}(\mathbf{C} \odot \mathbf{B})^T\|_F. \quad (8)$$

When $\hat{\mathbf{A}} = \mathbf{A} \cdot \text{diag}(\lambda)$, the optimal solution is:

$$\hat{\mathbf{A}} = \mathbf{X}_{(1)}[(\mathbf{C} \odot \mathbf{B})^T]^\dagger. \quad (9)$$

Here $\mathbf{A}^\dagger$ denotes the Moore–Penrose pseudoinverse of $\mathbf{A}$ [36].

Due to the special form of the Khatri–Rao product, the above formula can be written as:

$$\hat{\mathbf{A}} = \mathbf{X}_{(1)}(\mathbf{C} \odot \mathbf{B})(\mathbf{C}^T \mathbf{C} * \mathbf{B}^T \mathbf{B})^\dagger. \quad (10)$$

The advantage of equation (10) is that we only need to calculate the pseudo-inverse of a $R \times R$ matrix, instead of seeking a $JK \times R$ matrix. Due to the potential ill-posed problem of numerical calculation, to solve this problem, we normalized the columns of $\hat{\mathbf{A}}$ to get $\mathbf{A}$. In other words, let $\lambda_r = \|\hat{\mathbf{a}}_r\|$ and $\mathbf{a}_r = \frac{\hat{\mathbf{a}}_r}{\lambda_r}$, for $r = 1, 2, \dots, R$.

Algorithm 1: ADMM-based CP denoising algorithm

Input: noisy image $\mathcal{X}$, k_{mit} , ϵ

Output: denoising image $\hat{\mathcal{X}}$

1. Initialization: $\hat{\mathbf{A}}^0 = \hat{\mathbf{B}}^0 = \hat{\mathbf{C}}^0 = 0$, $k = 1$;
 2. **while** $k < k_{\text{mit}}$ or $\text{relcha} > \epsilon$ **do**
 3. Update $\hat{\mathbf{A}}$ via(10);
 4. Update $\hat{\mathbf{B}}$ via(11);
 5. Update $\hat{\mathbf{C}}$ via(12);
 6. Update $\hat{\mathcal{X}}$ via(7);
 7. **end**
-

For the parameters $\mathbf{B}$ and $\mathbf{C}$, we have the following equations to keep iterating until the experimental requirements are met.

$$\hat{\mathbf{B}} = \mathbf{X}_{(2)}(\mathbf{C} \odot \mathbf{A})(\mathbf{C}^T \mathbf{C} * \mathbf{A}^T \mathbf{A})^\dagger, \quad (11)$$

$$\hat{\mathbf{C}} = \mathbf{X}_{(3)}(\mathbf{A} \odot \mathbf{B})(\mathbf{A}^T \mathbf{A} * \mathbf{B}^T \mathbf{B})^\dagger. \quad (12)$$

The corresponding denoising algorithm is summarized in Algorithm 1, in which, k_{mit} is the maximum iteration of execution. The convergence of the designed scheme is guaranteed [37].

3 PCA tensor decomposition method

HSI data contain multi-dimensional information, such that analysis of the data in each dimension may lose the potential structure of the data. Fortunately, tensor decomposition can effectively mine the potential information among the data. In this section, we propose an HSI denoising method based on the PCA and low-rank CP decomposition (PCA TensorDecomp).

The denoising process of PCA TensorDecomp is shown in Figure 1.

The HSI model is formulated as:

$$\mathcal{X} = \mathcal{A} + \mathcal{N}, \quad (13)$$

where $\mathcal{X}$ is the noisy HSI data, $\mathcal{A}$ is the clean image, and $\mathcal{N}$ is the noise to be removed. Our purpose is to remove the noise data $\mathcal{N}$ in $\mathcal{X}$ through PCA TensorDecomp.

Considering the roles of PCA, on one hand, it can achieve the aim of dimensionality reduction for the HSI and, on the other hand, it can separate the principal composite components from the secondary composite components. With the few main principal composite

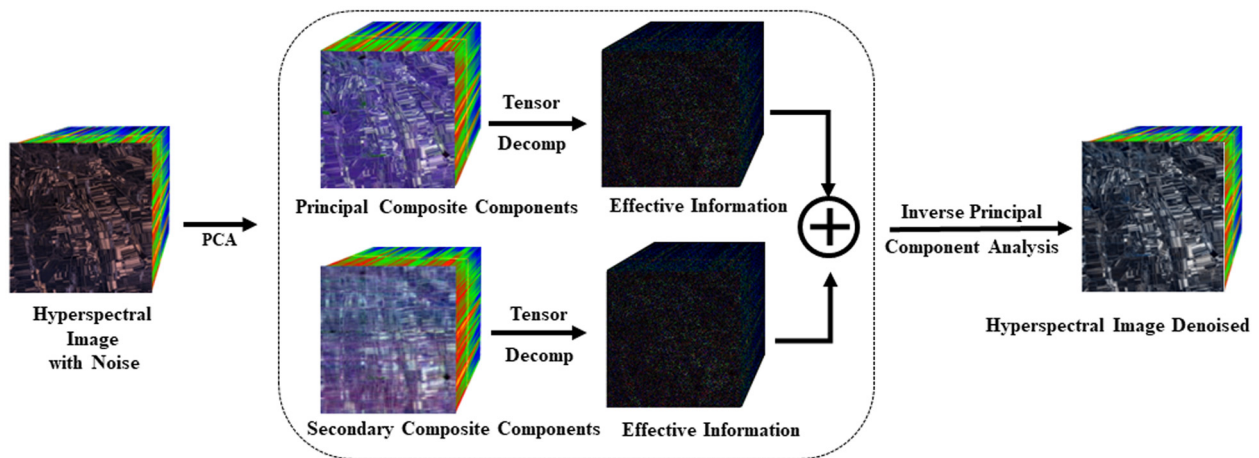


Figure 1: Flowchart of PCA TensorDecomp.

components, most of the relevant information can be retained. As mentioned earlier, those secondary composite components contain most of the noise, but they still have some useful information. Discarding the secondary composite components, this useful information will lose.

As the HSIs are tensor-type data, low-rank CP decomposition can be used for denoising. The use of a low-rank tensor decomposition can effectively remove the noise contained in the tensor.

Thus, to combine the advantages of PCA and CP decomposition, we propose a denoising algorithm based on the PCA and low-rank tensor CP decomposition named PCA TensorDecomp.

The steps of the algorithm are as follows:

Step 1: Input the noisy image $\mathcal{X} = \mathcal{A} + \mathcal{N}$.

Step 2: Preprocess the hyperspectral noise image $\mathcal{X}$ into the 2D data $\mathbf{E}$ and use PCA to reduce the dimension of the processed data into the data $\mathbf{Y}$; that is ($\mathbf{D}$ is the transformation matrix):

$$\mathbf{Y} = \mathbf{D}\mathbf{E}. \quad (14)$$

Step 3: Separate the principal and minor components and keep the previous principal composite components unchanged, denoted as $\mathbf{X}_1$. Then, treat the remaining principal composite components, which we denoted as $\mathbf{X}_2$.

Step 4: Denoise the previous principal composite components $\mathbf{X}_1$ and the remaining minor components $\mathbf{X}_2$ with the low-rank CP decomposition algorithm to obtain the denoised primary component $\widehat{\mathbf{X}}_1$ and minor component $\widehat{\mathbf{X}}_2$. The optimization function is:

$$\begin{aligned} & \min_{\mathbf{X}_1} \|\mathbf{X}_1 - \widehat{\mathbf{X}}_1\|, \min_{\mathbf{X}_2} \|\mathbf{X}_2 - \widehat{\mathbf{X}}_2\|, \\ & \text{s.t. } \widehat{\mathbf{X}}_1 = \sum_{R_1=1}^{r=1} \lambda_1 \mathbf{a}_{1r} \circ \mathbf{b}_{1r} \circ \mathbf{c}_{1r} = [[\lambda_1; \mathbf{A}_1, \mathbf{B}_1, \mathbf{C}_1]], \\ & \widehat{\mathbf{X}}_2 = \sum_{R_2=1}^{r=1} \lambda_2 \mathbf{a}_{2r} \circ \mathbf{b}_{2r} \circ \mathbf{c}_{2r} = [[\lambda_2; \mathbf{A}_2, \mathbf{B}_2, \mathbf{C}_2]], \end{aligned} \quad (15)$$

where the symbol definition above is the same as in Section 2.2.

Step 5: Finally, combine the denoising former principal composite component $\widehat{\mathbf{X}}_1$ and secondary composite component $\widehat{\mathbf{X}}_2$ together to form $\widehat{\mathbf{Y}}$. Inversely transform the composite components $\widehat{\mathbf{Y}}$ to obtain the final data $\widehat{\mathbf{E}}$. That is:

$$\widehat{\mathbf{Y}} = \widehat{\mathbf{X}}_1 + \widehat{\mathbf{X}}_2, \quad (16)$$

$$\widehat{\mathbf{E}} = \mathbf{D}^{-1}\widehat{\mathbf{Y}}, \quad (17)$$

where $\mathbf{D}^{-1}$ is the invertible matrix of the previous transformation matrix.

Step 6: Finally, the inversely converted data $\widehat{\mathbf{E}}$ are restored to an HSI; that is, the denoised image $\mathcal{A}$.

The flow chart of the proposed algorithm is shown below.

Algorithm 2: PCA TensorDecomp denoising algorithm

Input: noisy image $\mathcal{X}$

Output: denoising image denoising image $\mathcal{A}$

1. Preprocess $\mathcal{X}$ into the 2D data $\mathbf{E}$, and use PCA to reduce the dimension via(14);
 2. Separate the principal components $\mathbf{X}_1$ and minor components $\mathbf{X}_2$;
 3. $\widehat{\mathbf{X}}_1$ and $\widehat{\mathbf{X}}_2$ are obtained by denoising via(15);
 4. Combine $\widehat{\mathbf{X}}_1$ and $\widehat{\mathbf{X}}_2$ together to form $\widehat{\mathbf{Y}}$ via(16);
 5. Inversely transform $\widehat{\mathbf{Y}}$ to obtain the final data $\widehat{\mathbf{E}}$ via(17);
 6. Restore $\widehat{\mathbf{E}}$ to $\mathcal{A}$.
-

4 Experimental results and analysis

To test the effectiveness of the proposed denoising algorithm, in this section, we mainly report the results of experiments on simulated and real HSI data. Five traditional HSI denoising algorithms, Tucker decomposition [38], Adaptive Non-Local Means 3D (ANLM3D) [39], lower rank tensor approximation (LRTA) [40], truncated singular value decomposition (tSVD) [41], and PARAFAC [18], were selected for comparison. We evaluated the denoising results both subjectively and objectively. That is, we discuss the merits and demerits of five denoising algorithms from the perspective of both visual qualitative analysis and quantitative analysis. Peak signal-to-noise ratio (PSNR), structural similarity (SSIM), feature similarity (FSIM) [42], and erreur relative globale adimensionnelle de synthese (ERGAS) [43] were used for the quantitative analysis of simulated HSIs and real hyperspectral data.¹ The values of SSIM and FSIM are in the range [0,1] while the values of PSNR approach infinity as a limit. The greater the value, the less image distortion. Different from the former three measures, the smaller ERGAS is, the better does the target HSI estimate the reference one. Meanwhile, we then examine the horizontal average fluctuation of a given band. Notice that the parameter setting for PCA was three in all the following experiments.

¹ <https://rslab.ut.ac.ir/data>.

4.1 Simulated HSI denoising experiments

We selected the Washington DC Mall set from the Hyperspectral Digital Imagery Collection Experiment sensor for the simulation experiments. The original image is of size $1,208 \times 307 \times 191$. We selected 256×256 pixel and 191 sub-images of bands for the experiment. The shape of a sub-image (the bands 30, 40, and 50 are taken as R, G, and B components, respectively) is shown in Figure 2a. We added Gaussian noise with the noise intensity of 0.1 and 0.2 to this image and then used the denoising algorithm mentioned above to denoise the image and compare the denoising effects.

During the experiment, the number of ranks for the first principal components for CP decomposition was 240/22 (denoised with noise intensity 0.1/denoised with noise intensity 0.2) and the number of ranks for the second principal components for CP decomposition was 180/4.

Comparing Figures 2 and 3, it can be seen that the tSVD algorithm had the worst effect in both cases because it used matrix factorization technology to perform the denoising processing. The image restored by ANLM3D is blurred and has artifacts. Although Tucker obtained acceptable results at low noise levels, it could not remove noise at high noise levels. LRTA and PARAFAC effectively

eliminated noise, but image detail was lost. These visual conditions can be observed in the enlarged box. In comparison, the PCA TensorDecomp algorithms achieved superior results competitive with the other methods.

Table 1 presents the quantitative results from comparison of the denoising approaches with the Washington DC Mall set at two different noise levels. We highlight the top two algorithms with two different representations of the data. Between the two, the best performance for each quality index is highlighted by bolding the data, and the second-best performance is highlighted by highlighting the data. It can also be seen, from the comparison results in Table 1 that the average PSNR of the denoising method proposed in this article was the highest, indicating that it achieves a better improvement of the denoising results, compared with the other denoising algorithms. At the same time, the structural similarity of the denoising algorithm mentioned in this article was also significantly improved, compared with that of the other five algorithms, which indicates that the denoising algorithm proposed in this article also significantly improved the authenticity of image restoration.

Figure 4 shows the PSNR and SSIM values of each band on the Washington DC Mall set at two different noise levels. From the experimental results in Figure 4,

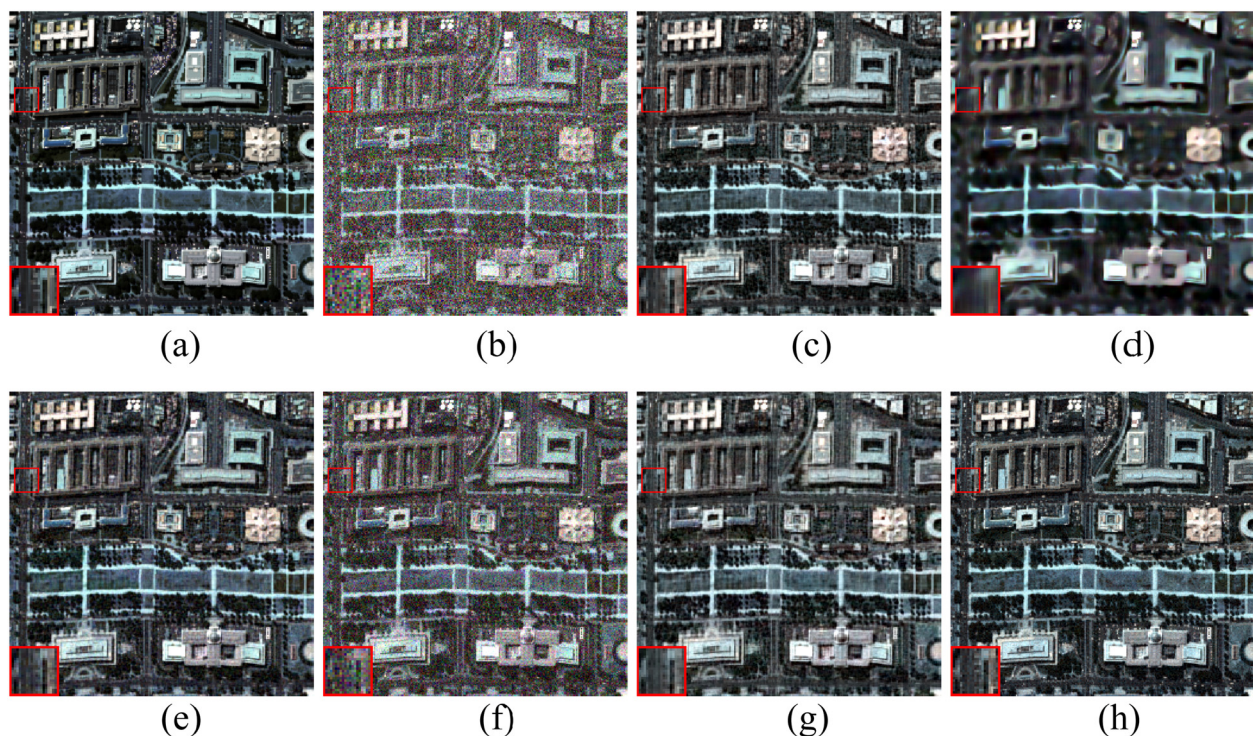


Figure 2: Image schematic diagram after denoising the simulated data with a noise intensity of 0.1. (a) Original image, (b) noisy image (0.1), (c) Tucker, (d) ANLM3D, (e) LRTA, (f) tSVD, (g) PARAFAC, and (h) PCA TensorDecomp.

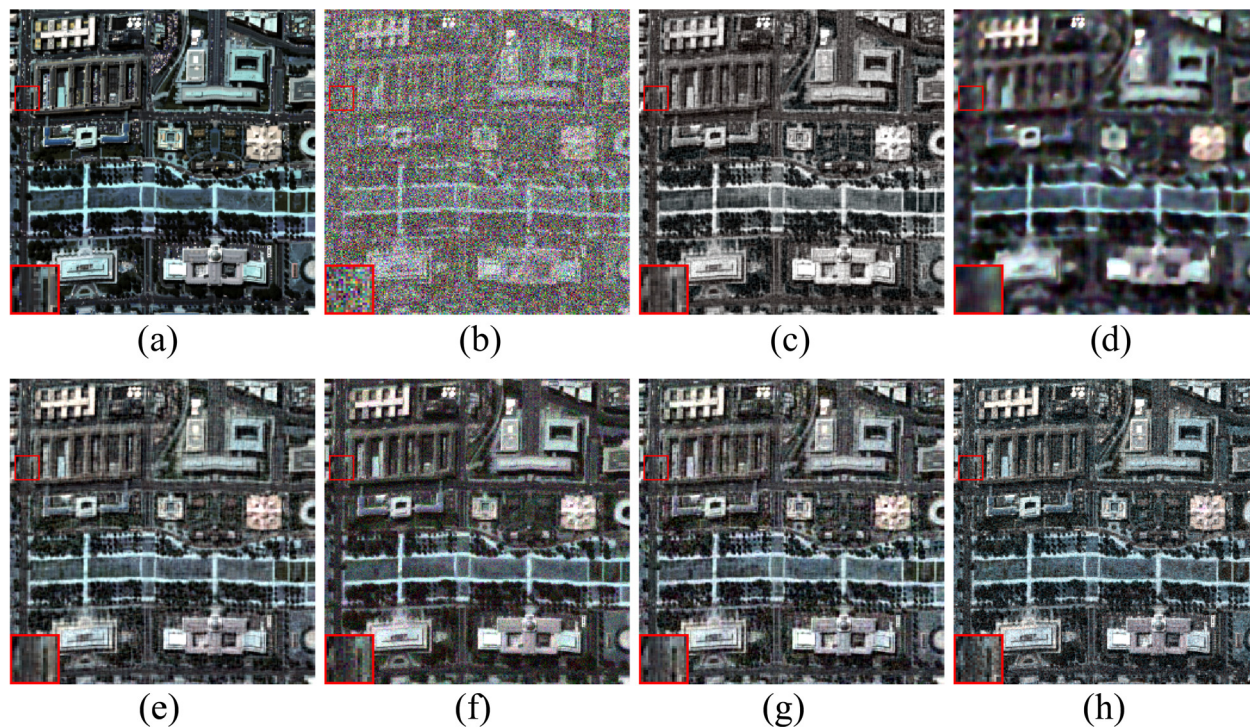


Figure 3: Image schematic diagram after denoising the simulated data with a noise intensity of 0.2. (a) Original image, (b) noisy image (0.2), (c) Tucker, (d) ANLM3D, (e) LRTA, (f) tSVD, (g) PARAFAC, and (h) PCA TensorDecomp.

Table 1: The image quality evaluation index of all bands of each denoising algorithm (bold: best; underline: second best)

Noise intensity	Method	MPSNR	MSSIM	MFSIM	MERGAS
0.1	Noisy image	20.00	0.432	0.732	422.499
	Tucker	32.83	0.920	0.953	93.569
	ANLM3D	27.95	0.740	0.838	162.521
	LRTA	<u>33.62</u>	<u>0.929</u>	<u>0.959</u>	<u>85.635</u>
	tSVD	26.03	0.690	0.867	209.735
	PARAFAC	31.59	0.902	0.942	108.098
	PCA TensorDecomp	35.82	0.953	0.975	67.887
0.2	Noisy image	13.98	0.191	0.560	844.942
	Tucker	29.47	0.877	0.929	139.550
	ANLM3D	25.78	0.630	0.781	208.236
	LRTA	30.05	0.858	0.919	128.480
	tSVD	21.36	0.461	0.764	359.067
	PARAFAC	<u>30.43</u>	<u>0.875</u>	<u>0.929</u>	<u>123.041</u>
	PCA TensorDecomp	32.14	0.879	0.940	115.865

we can see that, with different denoising algorithms, the PSNRs and SSIMs of different sub-images from the simulation experiment were all greater, compared to the noisy image. On the other hand, for different noise intensities and sub-images, the PCA TensorDecomp denoising algorithm was more effective than most of the other denoising methods because it obtained the best results in most bands. This indicates that the PCA TensorDecomp algorithm

proposed in this article can outperform the traditional denoising algorithm.

4.2 Real HSI denoising experiments

In this experiment, two groups of real HSI data were selected, namely Indian Pines and Gaofen-5 data.

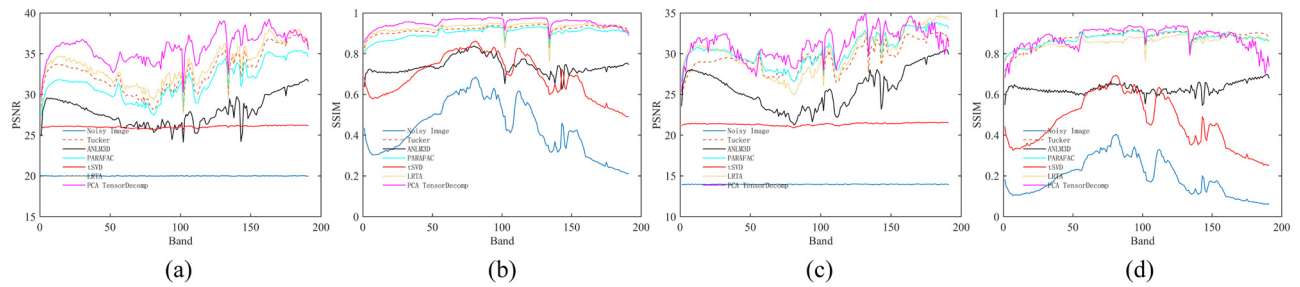


Figure 4: PSNR and SSIM of the different band from the simulated data Washington DC Mall analog. (a) PSNR (0.1), (b) SSIM (0.1), (c) PSNR (0.2), and (d) SSIM (0.2).

4.2.1 Indian Pines data

The real data were collected using the AVIRIS sensor at the Indian Pines Test site in northwestern Indiana. The size of the HSI data is 145×145 pixels and the original data are composed of 220 bands with a wavelength range of 0.4–2.5. The shape of a sub-image in this Indian Pines data (bands 3, 2, and 1 are taken as R, G, and B components, respectively) is shown in Figure 5a.

The HSI was denoised and the three denoising algorithms (Tucker Decomposition, PARAFAC, ANLM3D) mentioned in the previous section were compared. In the experiment, the rank of CP decomposition was 20/20. In the consideration of analyzing the experimental results qualitatively, row signature curves of the first

band estimated by all the compared methods were compared in Figure 6.

Comparing Figure 5, it can be seen that the denoising effect of PARAFAC was not very good for real data denoising. Tucker can get clear edges that ANLM3D cannot get, but it has artifacts in some areas.

Figure 6 shows the horizontal mean profiles of band 1 before and after denoising. As shown in Figure 6a, due to the existence of noise, there are rapid fluctuations in the curve before denoising. After processing by several denoising algorithms, the fluctuations are more or less suppressed. Here, we can see that the curve of the proposed PCA TensorDecomp method is more stable, which is in accordance with the visual results presented in Figure 5.

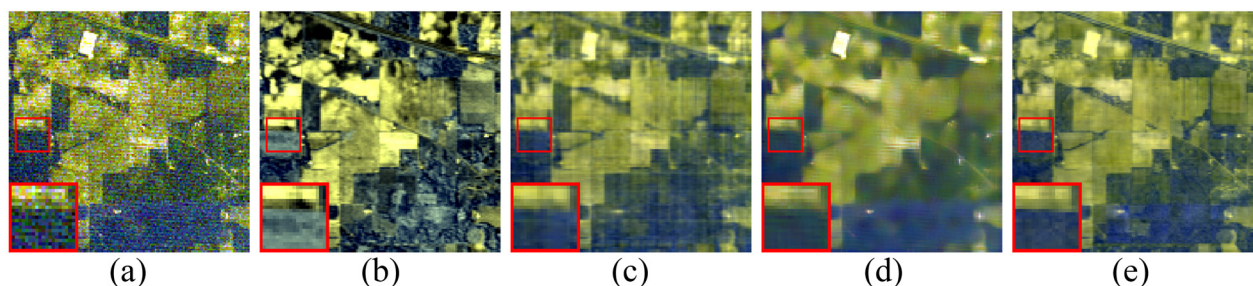


Figure 5: Image schematic diagram after denoising of the real data. (a) Original image, (b) Tucker, (c) PARAFAC, (d) ANLM3D, and (e) PCA TensorDecomp.

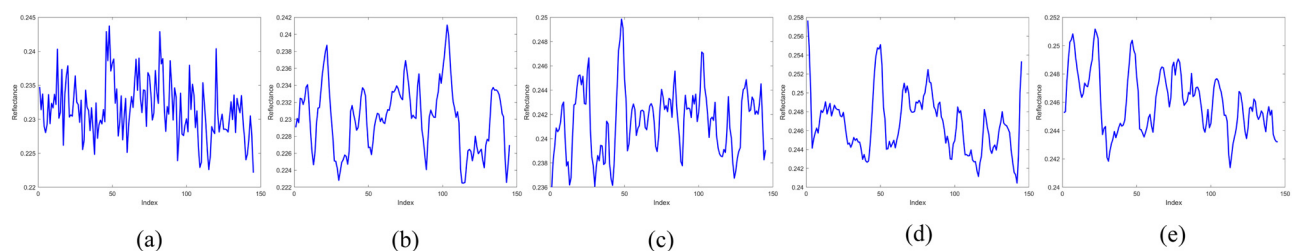


Figure 6: Row signature curves estimated by all the compared methods. (a) Original image, (b) Tucker, (c) PARAFAC, (d) ANLM3D, and (e) PCA TensorDecomp.

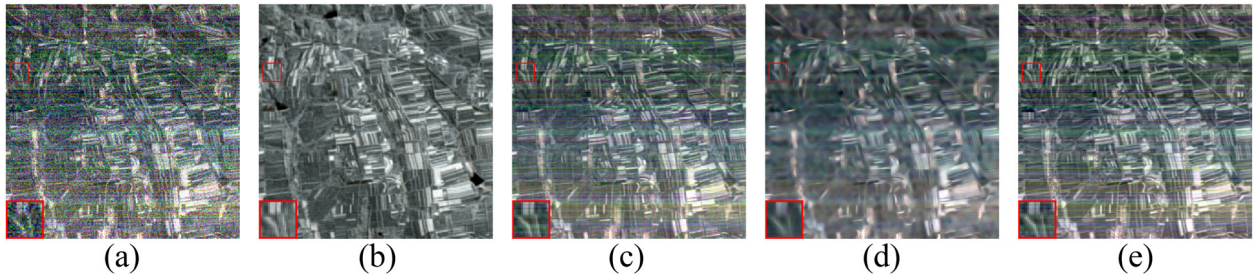


Figure 7: Image schematic diagram after denoising of the real data. (a) Original image, (b) Tucker, (c) PARAFAC, (d) ANLM3D, and (e) PCA TensorDecomp.

4.2.2 Gaofen-5 data

To test the effectiveness of the denoising algorithm proposed in this article, we adopted another data set Gaofen-5 data. The original data set consists of 256×256 pixels and 283 bands after cutting. In the experiment, the rank of CP decomposition was 260/20. The real sub-image of the data (bands 3, 2, and 1 are taken as R, G, and B components, respectively) is shown in Figure 7a.

As above, different denoising algorithms were adopted to denoise the Gaofen-5 data. Results are shown in Figure 7. As you can see from the figure, Tucker gets a distorted image and ANLM3D gets a blurred image. Figure 8 shows the horizontal mean profiles of the third band before and after denoising. As shown in Figure 8a, we get the same conclusion as the Indian Pines data set. Here, we can see that compared with other denoising algorithms, PCA TensorDecomp proposed in this article shows more superior.

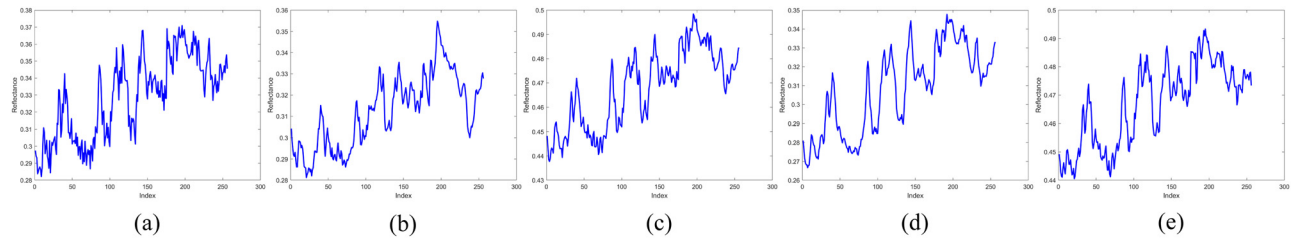


Figure 8: Row signature curves estimated by all the compared methods. (a) Original image, (b) Tucker, (c) PARAFAC, (d) ANLM3D, and (e) PCA TensorDecomp.

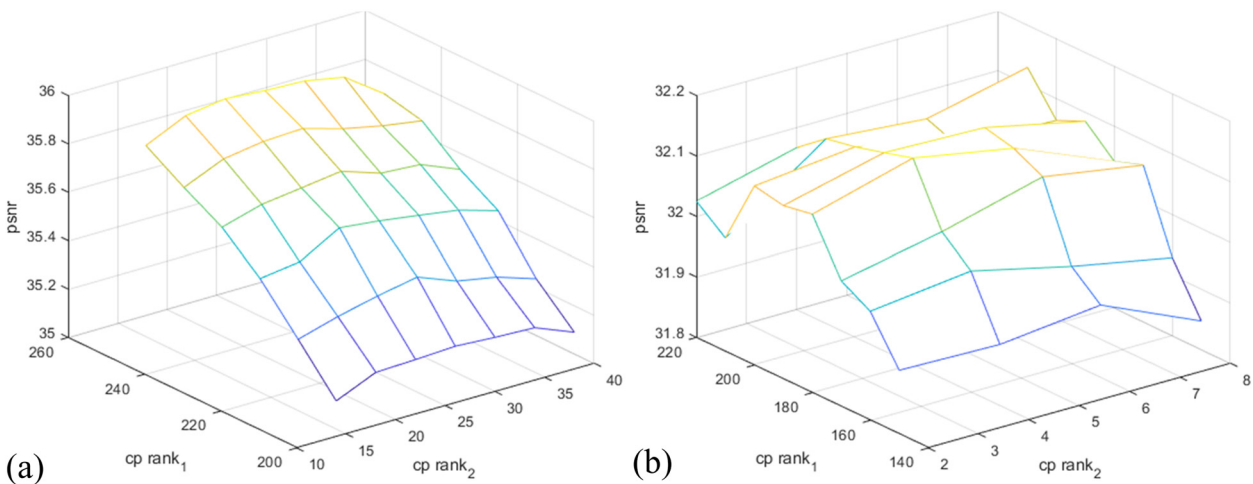


Figure 9: Relationship of CP rank with PSNR values at different noise levels in WDC datasets. (a) noise intensity = 0.1 and (b) noise intensity = 0.2.

4.3 Analysis of CP rank

In the PCA TensorDecomp model, when executing the algorithm, we need to determine the size of the CP rank to get the best denoising effect. Figure 9 shows the changes in PSNR corresponding to different CP ranks in WDC datasets under two noise intensities. It can be seen from the figures that when the noise intensity is 0.1, the greater the CP rank of the first principal component and the CP rank of the second principal component is in the interval of [22,30], the better the denoising effect of the proposed algorithm. When the noise intensity is 0.2, the CP rank of the first principal component is in the interval of [170, 190], and the smaller the CP rank of the second principal component, the better the denoising effect of the proposed algorithm. Therefore, the parameter size set in this article is reasonable and benign.

5 Conclusion

A new method for HSI noise reduction using PCA and a low-rank CP decomposition model (PCA TensorDecomp) was presented in this article. This method takes advantage of the fact that PCA can reduce the complexity of the processing data, as well as take into account the characteristics of HSIs and the correlations between various bands. To verify the validity of this algorithm, we selected simulated HSI data and two sets of real SIs to carry out comparative experiments, through both qualitative (i.e., assessing the visual effect) and quantitative analyses of the proposed method and five traditional methods. The results demonstrated that, compared with these methods, the proposed PCA TensorDecomp method can effectively reduce the noise of HSIs while retaining their fine structure. There are some aspects to be improved; for example, in the process of PCA, the principal component selection is carried out manually according to the degree of contribution. Meanwhile, a better algorithm should be sought, to get the number of CP decomposition ranks.

Acknowledgments: This work was supported in part by the Sichuan Science and Technology Program under Grant 2021YJ0351 and Grant 2021YFG0319, in part by the Opening Fund of Geomathematics Key Laboratory of Sichuan Province under Grant scsxdz2019zd03 and Grant scsxdz2021yb02, and in part by the National Key Research and Development Program of China under Grant 2017YFC0601505.

Author contributions: Hao Wu, Youhua Wei, Ruihan Yue, and Jun Feng proposed the model. Ruixue Gao and Rui Wen designed the experiments and carried them out. Ruixue Gao prepared the manuscript with contributions from all coauthors. All authors revised the manuscript.

Conflict of interest: Authors state no conflict of interest.

References

- [1] Bioucas-Dias JM, Plaza A, Dobigeon N, Parente M, Du Q, Gader P, et al. Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *IEEE J Sel Top Appl Earth Obs Remote Sens.* 2012;5(2):354–79. doi: 10.1109/JSTARS.2012.2194696.
- [2] Aharon M, Elad M, Bruckstein A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans Signal Process.* 2006;54(11):4311–22. doi: 10.1109/TSP.2006.881199.
- [3] Dabov K, Foi A, Katkovnik V, Egiazarian K. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Trans Image Process.* 2007;16(8):2080–95. doi: 10.1109/TIP.2007.901238.
- [4] Mallat S. A wavelet tour of signal processing. 3rd edn. Boston: Academic Press; 2009.
- [5] Huo LG, Feng XC. Denoising of hyperspectral remote sensing image based on principal component analysis and dictionary learning. *Dianzi Yu Xinxu Xuebao/J Electron Inf Technol.* 2014;36:2723–9. doi: 10.3724/SP.J.1146.2013.01840.
- [6] Yuan Q, Zhang L, Shen H. Hyperspectral image denoising employing a spectral-spatial adaptive total variation model. *IEEE Trans Geosci Remote Sens.* 2012;50(10 PART1):3660–77. doi: 10.1109/TGRS.2012.2185054.
- [7] Sun L, Luo JS. Three-dimensional hybrid denoising algorithm in derivative domain for hyperspectral remote sensing imagery. *Guang Pu Xue Yu Guang Pu Fen Xi/Spectroscopy Spectr Anal.* 2009;29(10):2717–20. doi: 10.3964/j.issn.1000-0593(2009)10-2717-04.
- [8] Xie Q, Zhao Q, Meng D, Xu Z. kronecker-basis-representation based tensor sparsity and its applications to tensor recovery. *IEEE Trans Pattern Anal Mach Intell.* 2018;40(8):1888–902. doi: 10.1109/TPAMI.2017.2734888.
- [9] Fan H, Li C, Guo Y, Kuang G, Ma J. Spatial-spectral total variation regularized low-rank tensor decomposition for hyperspectral image denoising. *IEEE Trans Geosci Remote Sens.* 2018;56(10):6196–213. doi: 10.1109/TGRS.2018.2833473.
- [10] Zheng W-J, Zhao X-L, Zheng Y-B, Pang Z-F. Nonlocal patch-based fully-connected tensor network decomposition for multispectral image inpainting. *IEEE Geosci Remote Sens Lett.* 2021;19:1–5. doi: 10.1109/LGRS.2021.3124804.
- [11] Bioucas-Dias JM, Nascimento JMP. Hyperspectral subspace identification. *IEEE Trans Geosci Remote Sens.* 2008;46(8):2435–45. doi: 10.1109/TGRS.2008.918089.
- [12] Meng S, Huang LT, Wang WQ. Tensor decomposition and PCA jointed algorithm for hyperspectral image denoising. *IEEE*

- Geosci Remote Sens Lett. 2016;13(7):897–901. doi: 10.1109/LGRS.2016.2552403.
- [13] Karami A, Yazdi M, Asli AZ. Noise reduction of hyperspectral images using kernel non-negative Tucker decomposition. *IEEE J Sel Top Signal Process.* 2011;5(3):487–93. doi: 10.1109/JSTSP.2011.2132692.
- [14] Bai X, Xu F, Zhou L, Xing Y, Bai L, Zhou J. Nonlocal similarity based nonnegative Tucker decomposition for hyperspectral image denoising. *IEEE J Sel Top Appl Earth Obs Remote Sens.* 2018;11(3):701–12. doi: 10.1109/JSTARS.2018.2791718.
- [15] Xie Q, Zhao Q, Meng D-Y, Xu Z-B, Gu S-H, Zuo W-M, et al. Multispectral images denoising by intrinsic tensor sparsity regularization. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*; 2016. Vol. 2016-Decem. p. 1692–700. doi: 10.1109/CVPR.2016.187.
- [16] Carroll JD, Chang JJ. Analysis of individual differences in multidimensional scaling via an n-way generalization of ‘Eckart-Young’ decomposition. *Psychometrika.* 1970;35(3):283–319. doi: 10.1007/BF02310791.
- [17] Harshman RA. Foundations of the PARAFAC procedure: Models and conditions for an ‘explanatory’ multimodal factor analysis. Vol. 16. University Microfilms, Ann Arbor, Michigan: UCLA Working Papers in Phonetics, No. 10,085; 1970. p. 1–84.
- [18] Liu X, Bourennane S, Fossati C. Denoising of hyperspectral images using the parafac model and statistical performance analysis. *IEEE Trans Geosci Remote Sens.* 2012;50(10):3717–24. doi: 10.1109/TGRS.2012.2187063.
- [19] Wu Y, Fang L, Li S. Weighted tensor rank-1 decomposition for nonlocal image denoising. *IEEE Trans Image Process.* 2019;28(6):2719–30. doi: 10.1109/TIP.2018.2889914.
- [20] Guo X, Huang X, Zhang L, Zhang L. Hyperspectral image noise reduction based on rank-1 tensor decomposition. *ISPRS J Photogramm Remote Sens.* 2013;83:50–63. doi: 10.1016/j.isprsjprs.2013.06.001.
- [21] Zeng H, Xie X, Cui H, Yin H, Ning J. Hyperspectral image restoration via global L 1–2 spatial–spectral total variation regularized local low-rank tensor recovery. *IEEE Trans Geosci Remote Sens.* 2020;59(4):3309–25.
- [22] Zeng H, Xie X, Cui H, Zhao Y, Ning J. Hyperspectral image restoration via cnn denoiser prior regularized low-rank tensor recovery. *Computer Vis Image Underst.* 2020;197:103004.
- [23] Dian R, Li S. Hyperspectral image super-resolution via subspace-based low tensor multi-rank regularization. *IEEE Trans Image Process.* 2019;28(10):5135–46.
- [24] Xue J, Zhao YQ, Bu Y, Liao W, Chan JC, Philips W. Spatial-spectral structured sparse low-rank representation for hyperspectral image super-resolution. *IEEE Trans Image Process.* 2021;30:3084–97.
- [25] Xue J, Zhao Y, Liao W, Chan JCW. Nonlocal low-rank regularized tensor decomposition for hyperspectral image denoising. *IEEE Trans Geosci Remote Sens.* 2019;57(7):5174–89. doi: 10.1109/TGRS.2019.2897316.
- [26] Kervrann C, Boulanger J. Optimal spatial adaptation for patch-based image denoising. *IEEE Trans Image Process.* 2006;15(10):2866–78. doi: 10.1109/TIP.2006.877529.
- [27] Zhang Z, Ely G, Aeron S, Hao N, Kilmer M. Novel methods for multilinear data completion and de-noising based on tensor-SVD. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*; 2014. p. 3842–9. doi: 10.1109/CVPR.2014.485.
- [28] Liu Y-Y, Zhao X-L, Zheng Y-B, Ma T-H, Hongyan Z. Hyperspectral image restoration by tensor fibered rank constrained optimization and plug-and-play regularization. *IEEE Trans Geosci Remote Sens.* 2021;60:1–17. Accepted. doi: 10.1109/TGRS.2020.3045169.
- [29] Zheng Y-B, Huang T-Z, Zhao X-L, Chen Y, He W. double-factor-regularized low-rank tensor factorization for mixed noise removal in hyperspectral image. *IEEE Trans Geosci Remote Sens.* 2020;58(12):8450–64.
- [30] Zheng Y-B, Huang T-Z, Zhao X-L, Jiang T-X, Ma T-H, Ji T-Y. Mixed noise removal in hyperspectral image via low-fibered-rank regularization. *IEEE Trans Geosci Remote Sens.* 2020;58(1):734–49.
- [31] Zhou B, Duan X, Ye D, Wei W, Woźniak M, Damaševičius R. Heterogeneous image matching via a novel feature describing model. *Appl Sci.* 2019;9(22):4792.
- [32] Xue J, Zhao Y, Huang S, Liao W, Chan JCW, Kong SG. Multilayer sparsity-based tensor decomposition for low-rank tensor completion. *IEEE Trans Neural Netw Learn Syst.* 2021;1–15. doi: 10.1109/TNNLS.2021.3083931.
- [33] Hitchcock FL. The expression of a tensor or a polyadic as a sum of products. *J Math Phys.* 1927;6(1–4):164–89. doi: 10.1002/sapm192761164.
- [34] Kruskal JB. Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear Algebra Appl.* 1977;18(2):95–138. doi: 10.1016/0024-3795(77)90069-6.
- [35] Boyd S, Parikh N, Chu E, Peleato B, Eckstein J. Distributed optimization and statistical learning via the alternating direction method of multipliers found. *Trends Mach Learn.* Jan. 2011;3(1):1–122.
- [36] Iserles A, Golub GH, Van Loan CF. *Matrix computations.* Vol. 74. Issue 469. Baltimore, MD: Johns Hopkins University Press; 1990.
- [37] Glowinski R, Tallec PL. *Augmented lagrangian and operator splitting methods in nonlinear mechanics.* Philadelphia, PA, USA: SIAM; 1989.
- [38] Ahmadi-Asl S, Abukhovich S, Asante-Mensah MG, Cichocki A, Phany AH, Tanaka T. Randomized algorithms for computation of Tucker decomposition and higher order SVD (HOSVD). *IEEE Access.* 2021;9:28684–706. doi: 10.1109/access.2021.3058103.
- [39] Manjón JV, Coupé P, Martí-Bonmatí L, Collins DL, Robles M. Adaptive non-local means denoising of MR images with spatially varying noise levels. *J Magn Reson Imaging.* 2010;31(1):192–203. doi: 10.1002/jmri.22003.
- [40] Renard N, Bourennane S, Blanc-Talon J. Denoising and dimensionality reduction using multilinear tools for hyperspectral images. *IEEE Geosci Remote Sens Lett.* 2008;5(2):138–42. doi: 10.1109/LGRS.2008.915736.
- [41] Nishimura Y, Suzuki T, Fukuda K, Fukuta M. Study of image reconstruction by UT probe array using truncated singular value decomposition. *Int J Appl Electromagn Mech.* 2014;45(1–4):21–6. doi: 10.3233/JAE-141808.
- [42] Zhang L, Zhang L, Mou X, Zhang D. FSIM: A feature similarity index for image quality assessment. *IEEE Trans Image Process.* 2011;20(8):2378–86. doi: 10.1109/TIP.2011.2109730.
- [43] Wald L. *Data fusion. Definitions and architectures – Fusion of images of different spatial resolutions.* Vol. 8. Paris, France: Presses de l’Ecole, Ecole des Mines de Paris; 2002.