Applied Mathematics and Computation 408 (2021) 126342

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

A novel non-convex low-rank tensor approximation model for hyperspectral image restoration

Jie Lin[a], Ting-Zhu Huang[a,*], Xi-Le Zhao[a], Tian-Hui Ma[b], Tai-Xiang Jiang[c], Yu-Bang Zheng[a]

[a] School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, Sichuan, 611731, China
[b] School of Science, Civil Aviation University of China, Tianjin, 300300, China
[c] FinTech Innovation Center, School of Economic Information Engineering, Southwestern University of Finance and Economics, Chengdu, Sichuan, 611130, China

ARTICLE INFO

Article history:
Received 4 November 2019
Revised 1 April 2021
Accepted 2 May 2021

Keywords:
Low-Rank tensor approximation
Mixed noise
Non-convex optimization
Restoration
Hyperspectral image

ABSTRACT

Remote sensing hyperspectral images (HSIs) are inevitably corrupted by several types of noise in the process of acquisition and transmission. In this paper, we propose a non-convex low-rank tensor approximation (NonLRTA) model for mixed noise removal, which can estimate the intrinsic structure of the underlying HSI from its noisy observation. The clean HSI component is characterized by the ϵ-norm, which is a non-convex surrogate to Tucker rank. The mixed noise is modeled as the sum of sparse and Gaussian components, which are regularized by the l_1-norm and the Frobenius norm, respectively. An efficient augmented Lagrange multiplier (ALM) algorithm is developed to solve the proposed model. Experiments implemented on simulated and real HSIs validate the superiority of the proposed method, as compared to the state-of-the-art matrix-based and tensor-based methods.

1. Introduction

Hyperspectral images (HSIs) are third-order tensors, containing abundant spatial and spectral information and have a wide range of applications, including agriculture production, environmental monitoring, and urban planning [1–6]. Unfortunately, the acquisition process of HSIs is inevitably perturbed by various factors including photon effects, sensor saturation, and transmission errors, leading to observed images corrupted by a mix of noise, such as Gaussian noise, impulse noise, stripes, and deadlines [7–12]. Noise corruption severely limits the subsequent applications of HSI, including classification, unmixing, and anomaly detection [13–16]. Therefore, HSI restoration, which focuses on restoring the clean HSI from its noise observation, is a vital process for the subsequent application.

Considering the correlation along the spectral dimension, several low-rank matrix approximation-based approaches have been proposed for HSI restoration. The approaches rearrange the third-order HSI tensor to the Casorati matrix and use the rank function to describe the low-rankness of the matrix. As the surrogates to the rank function, convex and non-

* Corresponding author. fax: +86 2861831280.
E-mail addresses: jielin96@126.com (J. Lin), tingzhuhuang@126.com (T.-Z. Huang), xlzhao122003@163.com (X.-L. Zhao), nkmth0307@126.com (T.-H. Ma), taixiangjiang@gmail.com (T.-X. Jiang), zhengyubang@163.com (Y.-B. Zheng).

https://doi.org/10.1016/j.amc.2021.126342
0096-3003/

convex matrix norms derived from singular value decomposition (SVD) are powerful tools to model the low-rankness of the matrix. Inspired by the robust principal component analysis (RPCA) problem [17], Zhang et al. [18] proposed a low-rank matrix recovery (LRMR) method, which uses nuclear norm to promote the low-rankness of the rearranged matrix, opening up a new perspective for mixed noise removal. Regarding the spatial smoothness, He et al. [19] integrated nuclear norm and total variation (TV) regularization to capture the spectral and spatial prior, and suggested a TV-regularized low-rank matrix factorization (LRTV). To jointly model the spectral-spatial prior knowledge, Liu et al. [20] incorporated a local spatial neighborhood weighted spectral-spatial TV into the nuclear norm-based low-rank restoration model. All methods achieve satisfactory performance in HSI restoration. However, the nuclear norm is not a better surrogate to the rank function and the singular values, which have clear physical meanings in image processing, should not be treated equally. To overcome the deficiency of convex relaxation, researchers proposed non-convex low-rank regularizers to treat different singular values adaptively, leading to better low-rankness promotion. Based on weighted nuclear norm [21], Xue et al. [22] assumed that spatial dimensions have similar structures in different locations and took both spectral low-rank prior and nonlocal self-similarity prior [23–25] into consideration. Kang et al. [26] proposed a non-convex rank approximation which provides a closer approximation to the matrix rank than the nuclear norm; Ye et al. [27] utilized general smooth function as an approximate of the rank function and proposed a smooth rank approximation (SRA) model. However, these methods convert the HSI tensor into the matrix by vectorizing each band of the original data, which only exploits the spectral low-rank prior and cannot finely preserve the intrinsic structure of the underlying HSI tensor [28,29].

To explore the global correlation underneath the natural HSI tensor, low-rank tensor approximation-based approaches have been proposed. The approaches use the tensor decomposition or its derived tensor norm to capture the low-rankness of the underlying HSI tensor. Three typical tensor decompositions are the CANDECOMP/PARAFAC (CP) [30], Tucker tensor decompositions [31] and tensor singular value decomposition (t-SVD) [32]. CP decomposition-based restoration methods include the PARAFAC model [33] and the rank-1 tensor decomposition noise reduction model [34]. Tucker decomposition-based restoration methods embody the lower rank-(k_1, k_2, k_3) tensor approximation (LRTA) [35], the genetic kernel Tucker decomposition [36], the multidimensional Wiener filtering [37] and the low-rank tensor decomposition with TV-regularization (LRTDTV) [38]. Under the definition of Tucker decomposition, Liu et al. [39] suggested a convex surrogate for Tucker rank named sum of the nuclear norm (SNN) to restore the missing entries in a low-rank tensor. Based on the tensor nuclear norm (TNN) [40] derived from t-SVD, the low-rank tensor recovery (LRTR) [41] method was proposed and achieved a performance improvement. However, t-SVD describes the spatial correlations by SVDs and captures the spectral correlation by the embedded circular convolution [42], while the correlation along the spectral dimension is more significant than those along the spatial dimension, leading to the fact that TNN cannot fully explore the correlation along each dimension of HSI tensor.

To overcome the above-mentioned limitations, in this paper, we proposed a novel tensor decomposition model for HSI restoration, where an ϵ-norm is employed to characterize the low-Tucker-rank tensor. Our main idea is to utilize an ϵ-norm to approximate the tensor rank, offering a better surrogate than SNN and preserving the intrinsic structure of clean HSI.

The main contributions of this paper can be summarized as follows.

First, we construct a novel low-Tucker-rank regularization ϵ-norm for capturing the low-rank structure underneath a tensor. Compared to traditional SNN, it is a better surrogate to Tucker rank and shrinks the singular values differently, achieving better effect to retain major components and suppress noise.

Second, we propose an ϵ-norm based non-convex low-rank tensor approximation (NonLRTA) model to preserve the intrinsic structure of low-rank tensor and jointly exploit the correlations along the spatial and spectral modes, which can effectively remove mixed noise, including Gaussian noise, impulse noise, stripes, and deadlines.

Third, we develop an efficient algorithm based on the augmented Lagrangian multipliers (ALM) [43] method to solve the proposed non-convex optimization problem. Extensive experiments indicate the proposed method improves the restoration results in terms of the qualitative and quantitative evaluation.

The remainder of this paper is organized as follows. Section 2 proposes the designed ALM algorithm. Section 3 presents simulated and real experiment results. Section 4 concludes this paper.

2. Notations

We denote tensors by capitalized calligraphic letters, e.g., $\mathcal{X}$; matrices are expressed with capitalized letters, e.g., X; vectors are denoted by bold lowercase letters, e.g., $\mathbf{x}$; scalars are represented as lowercase letters, e.g., x. An mth-order tensor is denoted as $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_m}$, whose elements are denoted as $x_{i_1,i_2,\cdots,i_m}$. The unfolding matrix $X_{(k)} = \text{unfold}_k(\mathcal{X}) \in \mathbb{R}^{n_k \times (\prod_{i \neq k} n_i)}$ is composed by taking the k-mode vectors of $\mathcal{X}$ as its columns. Conversely, the unfolding matrices along the k-mode can be transformed back to the tensor by $\mathcal{X} = \text{fold}_k(X_{(k)}) \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_m}$. The rank of unfolding matrix in the kth mode is denoted as $r_k = \text{rank}(X_{(k)})$ and the Tucker rank of mth-order tensor is defined as $\text{rank}(\mathcal{X}) = (r_1, r_2, \cdots, r_m)$ [30]. The l_1-norm of the tensor $\mathcal{X}$ is defined as $\|\mathcal{X}\|_1 = \sum_{i_1,i_2,\cdots,i_m} |x_{i_1,i_2,\cdots,i_m}|$ and the Frobenius norm of mth-order tensor $\mathcal{X}$ is defined as the square root of the sum of the squares of all its elements, i.e., $\|\mathcal{X}\|_F = (\sum_{i_1,i_2,\cdots,i_m} |x_{i_1,i_2,\cdots,i_m}|^2)^{1/2}$.

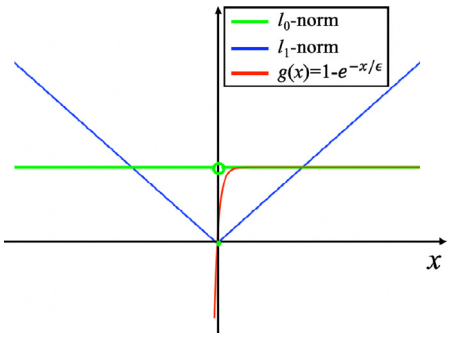

Fig. 1. Comparison of the l_0-norm, l_1-norm, and $g(x)$ for scalars (ϵ=0.05).

3. The proposed model

3.1. ϵ-norm regularization

Recently, Liu et al. [39] firstly proposed the SNN of a mth-order tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_m}$, which is defined as the linear combination of the nuclear norm of all unfolding matrices along each mode, i.e., $\|\mathcal{X}\|_{\text{SNN}} := \sum_{i=1}^{m} w_i \|X_{(i)}\|_*$, where w_i are nonnegative constants satisfying $\sum_{i=1}^{m} w_i = 1$.

Although many SNN based methods have achieved excellent performance in tensor recovery, SNN still has two defects. First, it measures the singular values with l_1-norm, which is not a better surrogate to Tucker rank. Second, it treats each singular value equally and this may result in the loss of main information because the larger singular values usually correspond to the main information, which means it should be shrunk less to preserve more information [21]. To remedy the above two defects, we propose the following ϵ-norm as the non-convex surrogate of the Tucker rank.

Definition 1 (ϵ-norm) The ϵ-norm of a mth-order tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_m}$ is defined as

$$\|\mathcal{X}\|_\epsilon := \sum_{i=1}^{m} w_i \sum_{j=1}^{s} g(\sigma_j(X_{(i)})), \tag{1}$$

where $s = \min\{n_i, \prod_{k \neq i} n_k\}$, w_i's are nonnegative constants satisfying $\sum_{i=1}^{m} w_i = 1$, $\sigma_j(X)$ is the jth singular value of X, and $g(x) = 1 - e^{-\frac{x}{\epsilon}}$ with constant $\epsilon > 0$. Specially, when the mode number m is equal to 2, i.e., $X \in \mathbb{R}^{n_1 \times n_2}$, (1) can be simplified to $\|X\|_\epsilon = \sum_{j=1}^{\min\{n_1, n_2\}} g(\sigma_j(X))$. Note that the ϵ-norm is a pseudonorm.

Compared to SNN, the proposed ϵ-norm has two advantages. First, it approximates the tensor rank better than SNN, which can promote the low-rankness of the solution. Fig. 1 shows that the function g can approximate l_0-norm closer than l_1-norm. Thus the ϵ-norm, the sum of the function g of singular values, is a better surrogate to the Tucker rank than SNN. Second, the ϵ-norm-based weighted singular value thresholding (WSVT) treats singular values differently. It can shrink the larger singular values less to keep major information and shrink the smaller ones more to suppress the noise (see the detailed proof in section 3.2: Remark 1).

3.2. NonLRTA model and solving algorithm

Under the definition of ϵ-norm, we propose a NonLRTA model for HSI restoration, which is formulated as

$$\begin{aligned} \min_{\mathcal{L},\mathcal{S},\mathcal{N}} \quad & \|\mathcal{L}\|_\epsilon + \lambda\|\mathcal{S}\|_1 + \rho\|\mathcal{N}\|_F^2 \\ \text{s.t.} \quad & \mathcal{X} = \mathcal{L} + \mathcal{S} + \mathcal{N}, \end{aligned} \tag{2}$$

where $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is the observed noisy HSI, $\mathcal{L} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ is the underlying clean HSI, the data noise is modeled as a mixture of the sparse component $\mathcal{S} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and the Gaussian component $\mathcal{N} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$. Here, $n_1 \times n_2$ denotes the spatial resolution and n_3 represents the number of spectral bands. $\lambda > 0$ and $\rho > 0$ are the regularization parameters.

We use the augmented Lagrangian multiplier (ALM) method to solve the optimization problem (2). By introducing auxiliary variables $\mathcal{M}_i (i = 1, 2, 3)$ and letting $\mathcal{L} = \mathcal{M}_i$, the problem is separable and can be rewritten as

$$\begin{aligned} \min_{\mathcal{L},\mathcal{S},\mathcal{M}_i,\mathcal{N}} \quad & \sum_{i=1}^{3} w_i \|M_{i(i)}\|_\epsilon + \lambda\|\mathcal{S}\|_1 + \rho\|\mathcal{N}\|_F^2 \\ \text{s.t.} \quad & \mathcal{X} = \mathcal{L} + \mathcal{S} + \mathcal{N}, \mathcal{L} = \mathcal{M}_i, i = 1, 2, 3. \end{aligned} \tag{3}$$

The ALM algorithm seeks the saddle point of the following augmented Lagrangian function:

$$\begin{aligned} \mathrm{L}(\mathcal{L}, \mathcal{S}, \mathcal{N}, \mathcal{M}_i, \mathcal{Y}_i, \mathcal{Z}; \alpha_i, \beta) = & \sum_{i=1}^{3} w_i \|M_{i(i)}\|_\epsilon + \lambda\|\mathcal{S}\|_1 + \rho\|\mathcal{N}\|_F^2 + < \mathcal{Y}_i, \mathcal{L} - \mathcal{M}_i > \\ & + < \mathcal{Z}, \mathcal{X} - \mathcal{L} - \mathcal{S} - \mathcal{N} > + \tfrac{\alpha_i}{2}\|\mathcal{L} - \mathcal{M}_i\|_F^2 + \tfrac{\beta}{2}\|\mathcal{X} - \mathcal{L} - \mathcal{S} - \mathcal{N}\|_F^2, \end{aligned} \tag{4}$$

where $\mathcal{Y}_i \in \mathbb{R}^{n_1\times n_2\times n_3}$ and $\mathcal{Z} \in \mathbb{R}^{n_1\times n_2\times n_3}$ are Lagrangian multipliers, $\lambda > 0$ and $\rho > 0$ are regularization parameters, and α_i and β are penalty parameters. According to the idea of Gauss-Seidel iteration, ALM alternatively optimizes the augmented Lagrangian function (4) with respect to one variable, while fixing the others, i.e.,

$$\begin{cases} \mathcal{M}_i^{k+1} = \arg\min\limits_{\mathcal{M}_i} \mathrm{L}(\mathcal{L}^k, \mathcal{S}^k, \mathcal{N}^k, \mathcal{M}_i, \mathcal{Y}_i^k, \mathcal{Z}^k; \alpha_i, \beta), i = 1, 2, 3, \\ \mathcal{L}^{k+1} = \arg\min\limits_{\mathcal{L}} \mathrm{L}(\mathcal{L}, \mathcal{S}^k, \mathcal{N}^k, \mathcal{M}_i^{k+1}, \mathcal{Y}_i^k, \mathcal{Z}^k; \alpha_i, \beta), \\ \mathcal{S}^{k+1} = \arg\min\limits_{\mathcal{S}} \mathrm{L}(\mathcal{L}^{k+1}, \mathcal{S}, \mathcal{N}^k, \mathcal{M}_i^{k+1}, \mathcal{Y}_i^k, \mathcal{Z}^k; \alpha_i, \beta), \\ \mathcal{N}^{k+1} = \arg\min\limits_{\mathcal{N}} \mathrm{L}(\mathcal{L}^{k+1}, \mathcal{S}^{k+1}, \mathcal{N}, \mathcal{M}_i^{k+1}, \mathcal{Y}_i^k, \mathcal{Z}^k; \alpha_i, \beta), \\ \mathcal{Y}_i^{k+1} = \mathcal{Y}_i^k + \alpha_i(\mathcal{L}^{k+1} - \mathcal{M}_i^{k+1}), i = 1, 2, 3, \\ \mathcal{Z}^{k+1} = \mathcal{Z}^k + \beta(\mathcal{X} - \mathcal{L}^{k+1} - \mathcal{S}^{k+1} - \mathcal{N}^{k+1}). \end{cases} \tag{5}$$

1) Update $\mathcal{M}_i$. The $\mathcal{M}_i$-subproblem is as follows:

$$M_{i(i)}^{k+1} = \arg\min_{M_{i(i)}} \frac{w_i}{\alpha_i}\|M_{i(i)}\|_\epsilon + \frac{1}{2}\|\mathcal{M}_i - (\mathcal{L}^k + \frac{\mathcal{Y}_i^k}{\alpha_i})\|_F^2. \tag{6}$$

According to (1), we can deduce the function g is continuous, differentiable, smooth, concave, and monotonically increasing on $[0, +\infty)$ with respect to the singular values of $M_{i(i)}$. Then, by linearizing it, we consider the following relaxation problem:

$$M_{i(i)}^{k+1} = \arg\min_{M_{i(i)}} \tfrac{w_i}{\alpha_i}\sum_{j=1}^{\min\{m_i,n_i\}} \nabla g(\sigma_j(M_{i(i)}^k))\sigma_j(M_{i(i)}) + \tfrac{1}{2}\|\mathcal{M}_i - (L_{(i)}^k + \tfrac{Y_{i(i)}^k}{\alpha_i})\|_F^2. \tag{7}$$

According to Theorem 2 in [44], $M_{i(i)}^*$ can be solved by generalized WSVT and has the following closed-form solution:

$$M_{i(i)}^* = US_{\frac{w_i}{\alpha_i}}(L_{(i)}^k + \frac{Y_{i(i)}^k}{\alpha_i})V^T, \tag{8}$$

$$S_{\frac{w_i}{\alpha_i}}(L_{(i)}^k + \frac{Y_{i(i)}^k}{\alpha_i}) = \mathrm{Diag}\{(\Sigma_{jj} - \frac{w_i\nabla g(\sigma_j^k)}{\alpha_i})_+\}, \tag{9}$$

where $U\Sigma V^T$ is the SVD of $L_{(i)}^k + \frac{Y_{i(i)}^k}{\alpha_i}$, $\sigma_1^k \geq \sigma_2^k \geq \cdots \geq \sigma_s^k$ are singular values of $M_{i(i)}^k$, and $x_+ = \max\{x, 0\}$. By folding operation, we get $\mathcal{M}_i = \mathrm{fold}_i(M_{i(i)})$.

Next we illustrate that the ϵ-norm based WSVT (8) gives more shrinkage to the smaller singular values and less shrinkage to the larger singular values. According to (9), for $j = 1, \ldots, s$, Σ_{jj} and σ_j^k have the same monotonicity. Thus, we just need to prove that $w_i\nabla g(\sigma_j^k)/\alpha_i$ is monotonically decreasing with respect to σ_j^k. The following Remark gives the proof.

Remark 1 Defining $h(x) = \frac{w_i}{\alpha_i}\nabla g(x)$ with $g(x) = 1 - e^{-\frac{x}{\epsilon}}$, where w_i, α_i, and ϵ are positive constants and $x \in [0, +\infty)$. Then the derivative of $h(x)$ is $h'(x) = -\frac{w_i}{\epsilon^2\alpha_i}e^{-\frac{x}{\epsilon}} < 0$, when $x \in [0, +\infty)$. Thus, $h(x)$ is a monotonically decreasing function.

The computational complexity of this step mainly is in computing the SVD of a matrix with size $n_i \times (\prod_{i\neq j} n_j)(i = 1, 2, 3)$, whose time complexity is $O(\min\{n_1^2(n_2n_3), (n_2n_3)^2n_1\} + \min\{n_2^2(n_3n_1), (n_3n_1)^2n_1\} + \min\{n_3^2(n_1n_2), (n_1n_2)^2n_3\})$.

2) Update $\mathcal{L}$. The $\mathcal{L}$-subproblem is as follows:

$$\mathcal{L}^{k+1} = \arg\min_{\mathcal{L}} \sum_{i=1}^{3}\frac{\alpha_i}{2}\|\mathcal{L} - (\mathcal{M}_i^{k+1} - \frac{\mathcal{Y}_i^k}{\alpha_i})\|_F^2 + \frac{\beta}{2}\|\mathcal{L} - (\mathcal{X} - \mathcal{S}^k - \mathcal{N}^k + \frac{\mathcal{Z}^k}{\beta})\|_F^2. \tag{10}$$

Taking the derivative with respect to $\mathcal{L}$, we get the exact solution of $\mathcal{L}$:

$$\mathcal{L}^{k+1} = \frac{\sum_{i=1}^3(\alpha_i\mathcal{M}_i^{k+1} - \mathcal{Y}_i^k) + \beta(\mathcal{X} - \mathcal{S}^k - \mathcal{N}^k) + \mathcal{Z}^k}{\sum_{i=1}^3\alpha_i + \beta}. \tag{11}$$

The computational complexity of this step is $O(n_1n_2n_3)$.

3) Update $\mathcal{S}$. The $\mathcal{S}$-subproblem is as follows:

$$\mathcal{S}^{k+1} = \arg\min_{\mathcal{S}} \lambda\|\mathcal{S}\|_1 + \frac{\beta}{2}\|\mathcal{S} - (\mathcal{X} - \mathcal{L}^{k+1} - \mathcal{N}^k + \frac{\mathcal{Z}^k}{\beta})\|_F^2. \tag{12}$$

The solution of (12) is given by

$$\mathcal{S}^{k+1} = \mathrm{soft}_{\frac{\lambda}{\beta}}(\mathcal{X} - \mathcal{L}^{k+1} - \mathcal{N}^k + \frac{\mathcal{Z}^k}{\beta}), \tag{13}$$

where $\mathrm{soft}_{\frac{\lambda}{\beta}}(\cdot)$ is defined as

$$\mathrm{soft}_{\frac{\lambda}{\beta}}(x) = \begin{cases} x - \frac{\lambda}{\beta}, & \text{if } x > \frac{\lambda}{\beta}, \\ x + \frac{\lambda}{\beta}, & \text{if } x < \frac{\lambda}{\beta}, \\ 0, & \text{otherwise}, \end{cases}$$

The computational complexity of this step is $O(n_1n_2n_3)$.

4) Update $\mathcal{N}$. The $\mathcal{N}$-subproblem is as follows:

$$\mathcal{N}^{k+1} = \arg\min_{\mathcal{N}} \ \rho\|\mathcal{N}\|_F^2 + \frac{\beta}{2}\|\mathcal{N} - (\mathcal{X} - \mathcal{L}^{k+1} - \mathcal{S}^{k+1} + \frac{\mathcal{Z}^k}{\beta})\|_F^2, \tag{14}$$

whose solution can be exactly calculated as

$$\mathcal{N}^{k+1} = \frac{\beta(\mathcal{X} - \mathcal{L}^{k+1} - \mathcal{S}^{k+1}) + \mathcal{Z}^k}{2\rho + \beta}. \tag{15}$$

The computational complexity of this step is $O(n_1n_2n_3)$.

For the penalty parameters, we update them in each iteration to accelerate the convergence of the algorithm [38,43],

$$\begin{cases} \alpha_i = \min\{\alpha_i * r, \alpha_{max}\}, \ i = 1, 2, 3, \\ \beta = \min\{\beta * r, \beta_{max}\}. \end{cases} \tag{16}$$

The updating of $\mathcal{M}_i$, $\mathcal{L}$, $\mathcal{S}$, and $\mathcal{N}$ have closed-form solutions and the overall algorithm for solving NonLRTA model is summarized in Algorithm 1. For an input third-order tensor $\mathcal{X} \in \mathbb{R}^{n_1\times n_2\times n_3}$, it is clear that the major cost in each iteration is updating $\mathcal{M}_i$. Therefore, the computational complexity for each iteration is $O(\min\{n_1^2(n_2n_3), (n_2n_3)^2n_1\} + \min\{n_2^2(n_3n_1), (n_3n_1)^2n_2\} + \min\{n_3^2(n_1n_2), (n_1n_2)^2n_3\})$.

4. Experimental results

We report the experimental results of the proposed method and compare it with state-of-the-art methods: matrix-based method LRMR [15], LRTV [6] and SRA1 [27], and tensor-based methods LRTA [35], LRTR [41] and LRTDTV [38]. In all simulated and real experiments, the parameters mentioned in the competing methods are optimally set or manually adjusted according to the reference papers. The gray values of testing HSIs are scaled to the interval [0, 1] band by band, and they are returned to the original level after restoration. All experiments are performed under Windows 10 and MATLAB R2016a running on a desktop with a Core i5 3.30GHz CPU and 16GB memory.

4.1. Simulated experiment

Washington DC Mall dataset. We utilize the HYDICE image of the Washington DC Mall to test the restoration effect of the proposed method in our simulated experiment. The whole image contains 1208×307 pixels and 191 spectral channels. Due to the page limitation, a sub-image of size $256 \times 256 \times 191$ is utilized in our experiment.

We test five cases of noise degradation, which are listed below.

Case 1: *Gaussian noise* and *impulse noise*. Gaussian noise with zero-mean and variance 0.025 is added to each band; impulse noise with percentage 0.05 is added to 30 bands randomly chosen from all bands.

Case 2: *Gaussian noise* and *impulse noise*. Gaussian noise with zero-mean and variance 0.1 is added to each band; impulse noise with percentage 0.2 is added to random 30 bands.

Case 3: *Gaussian noise, stripes* and *deadlines*. Gaussian noise with zero-mean and variance 0.025 is added to each band; stripes with number from 20 to 40 are added to random 30 bands; deadlines with number from 5 to 15 are added to random 30 bands.

Case 4: *Gaussian noise, impulse noise, stripes* and *deadlines*. Gaussian noise with zero-mean and variance 0.025 is added to each band; impulse noise with percentage 0.05 is added to random 30 bands; stripes with number from 20 to 40 are added to random 30 bands; deadlines with number from 5 to 15 are added to random 30 bands.

Case 5: *Gaussian noise, impulse noise, stripes* and *deadlines*. Gaussian noise with zero-mean and variance 0.1 is added to each band; impulse noise with percentage 0.2 is added to random 30 bands; stripes with number from 20 to 40 are added to random 30 bands; deadlines with number from 5 to 15 are added to random 30 bands.

Evaluation measures. We select three picture quality indexes to measure the results quantitatively, including the peak signal-to-noise ratio (PSNR), the structural similarity (SSIM) [45], and the feature similarity (FSIM) [46], which are defined as follows:

$$\mathrm{PSNR}_i = 10 \times \log \frac{n_1n_2}{\|X_i - \hat{X}_i\|_F^2},$$

$$\mathrm{SSIM}_i = \frac{(2\mu_{X_i}\mu_{\hat{X}_i} + c_1)(2\sigma_{X_i\hat{X}_i})}{(\mu_{X_i}^2 + \mu_{\hat{X}_i}^2 + c_1)(\sigma_{X_i}^2 + \sigma_{\hat{X}_i}^2 + c_2)},$$

$$\mathrm{FSIM}_i = \frac{\Sigma_{X_i\in\Omega}[S_{PC}(X_i)][S_G(X_i)][PC_m(X_i)]}{\Sigma_{X_i\in\Omega}[PC_m(X_i)]},$$

where n_1 and n_2 denote the size of each band, X_i and $\hat{X}_i$ denote the ith band of original image and restored image; μ_{X_i} and $\mu_{\hat{X}_i}$ denote the average values of image X_i and $\hat{X}_i$, $\sigma_{X_i}^2$ and $\sigma_{\hat{X}_i}^2$ stand for the variances, and $\sigma_{X_i\hat{X}_i}$ is the covariance between

Table 1
Quantitative evaluation and running time (in minutes) of different restoration methods in simulated experiment. For each quality index, the second-best results are underlined and the best results are marked in bold.

		Noisy	LRMR	LRTV	SRA1	LRTA	LRTR	LRTDTV	NonLRTA
Case 1	MPSNR(dB)	29.887	37.392	39.632	39.790	35.806	37.262	41.028	**42.901**
	MSSIM	0.7791	0.9331	0.9763	0.9752	0.8680	0.9639	0.9823	**0.9894**
	MFSIM	0.9014	0.9661	0.9829	0.9851	0.9432	0.9786	0.9912	**0.9943**
	time		3.345	11.556	18.788	0.567	7.926	5.563	4.504
Case 2	MPSNR(dB)	18.714	30.486	32.265	29.037	29.340	30.169	34.138	**34.409**
	MSSIM.	0.3235	0.8153	0.8730	0.8301	0.7473	0.8591	**0.9243**	0.9114
	MFSIM	0.6392	0.9065	0.9043	0.9005	0.8601	0.9214	0.9548	**0.9567**
	time		3.167	11.742	12.746	1.723	10.264	5.607	3.936
Case 3	MPSNR(dB)	29.540	37.600	39.187	38.792	34.804	33.485	39.469	**42.662**
	MSSIM	0.7991	0.9532	0.9723	0.9741	0.8863	0.9066	0.9762	**0.9891**
	MFSIM	0.9057	0.9706	0.9802	0.9825	0.9438	0.9502	0.9871	**0.9940**
	time		18.272	7.211	10.269	0.196	2.826	5.289	5.644
Case 4	MPSNR(dB)	27.450	34.8361	39.150	37.470	31.482	33.786	39.584	**42.574**
	MSSIM	0.7238	0.9123	0.9716	0.9643	0.7915	0.9085	0.9774	**0.9887**
	MFSIM	0.8746	0.9529	0.9800	0.9788	0.9026	0.9508	0.9881	**0.9939**
	time		4.563	10.906	16.080	0.369	6.628	5.680	4.638
Case 5	MPSNR(dB)	18.064	29.025	31.986	27.800	26.660	28.621	33.866	**34.063**
	MSSIM	0.3010	0.7841	0.8690	0.8035	0.6524	0.8110	0.8998	**0.9064**
	MFSIM	0.6245	0.8917	0.9011	0.8862	0.8112	0.9007	0.9522	**0.9548**
	time		5.0700	11.168	11.824	2.209	6.952	5.800	3.940

X_i and $\hat{X}_i$, c_1 and c_2 are default constants; Ω means the whole image spatial domain, S_{PC} and S_G estimate phase congruency and similarity of gradient magnitude, respectively. The mean values among all bands are called the mean PSNR (MPSNR), mean SSIM (MSSIM), and mean FSIM (MFSIM). In general, the higher these three picture quality indexes are, the better the restoration results are.

Parameter selection. There are six parameters in our model. The weight $\mathbf{w}$, the ϵ-norm's parameter ϵ, and the regularization parameters λ, ρ, α_i, and β. The λ is empirically selected from a candidate set $\{0.1, 0.15, 0.2, 0.25, 0.3\}$ with the highest MPSNR and the ρ is set as $\rho = 10 * \lambda$. We initialize both α_i and δ to 0.1 and fixe the remaining parameters as follows: $\mathbf{w} = (0.05, 0.05, 0.90)$ and $\epsilon = 0.02$. A detailed discussion of parameter selection is presented in Section 4.3.

Performance evaluation. Table 1 gives the MPSNR, MSSIM, MFSIM, and running time of different methods in each noise setting. As observed, the proposed method significantly outperforms the competing methods in terms of all evaluation measures. Fig. 2 shows the PSNR, SSIM and FSIM measurements across all brands of the HSI in five cases with different noise. In all cases, NonLRTA achieves the highest PSNR, SSIM, and FSIM values across almost all bands, which demonstrates the superiority of the proposed method.

In terms of visual quality, Figs. 3–5 show three representative bands of restoration results. Band 119 shown in Fig. 3 is polluted by both impulse noise and Gaussian noise in Case 2. It can be observed LRTA cannot remove this complex mixed noise from the noisy images. LRMR, SRA1, and LRTR perform better in noise removal, but the restored HSI loses many details compared with the original image. LRTDTV and NonLRTA can remove all kinds of noise efficiently; NonLRTA preserves the edge and detail information finely, and the intensity of the restoration result is closer to the original image. Band 69 shown in Fig. 4 is polluted by the Gaussian noise, stripes, and deadlines in Case 3. LRTA and LRTR fail to remove stripes and deadlines; LRMR removes part of the Gaussian noise and LRTV, SRA1, LRTDTV, and NonLRTA can remove mixed noise efficiently. Band 79 shown in Fig. 5 is polluted by the impulse noise, Gaussian noise, stripes and deadlines in Case 5. LRTV and LRTA perform poorly in dealing with this high noise level. The result of LRTR remains deadlines. While LRMR and SRA1 can remove most noise, some local details are lost. LRTDTV remove all kind of noise with the loss of some details. In aspect of noise removal and detail preservation, NonLRTA has better performance in visual effects and evaluation indicators.

We display the spectral signatures of original data and restoration results to further compare the performances of the competing methods. Figs. 6 and 7 show the spectral signatures of the pixel (80,80) in case 3 and the pixel (90,90) in case 4. It is clear that the spectral signatures restored by the proposed method are closer to the original one compared with all competing methods.

Regarding the running time in Table 1, LRTA and LRMR are the fastest among all methods, and our method is faster than LRTV, SRA1, and LRTDTV. Considering the restoration performance, our method achieves the best performance/efficiency trade-off.

4.2. Real experiments

We choose the Urban dataset of HYDICE image and the Australia dataset of the EO-1 Hyperion image to test the restoration effect of the proposed method. Specific experimental designs and restoration results are given below.

Urban dataset. The Urban dataset contains 210 spectral bands, and each band contains 307×307 pixels. Spectral bands 104–108, 139–151 and 207–210 are usually deleted before restoration because of serious noise pollution. In this experiment,

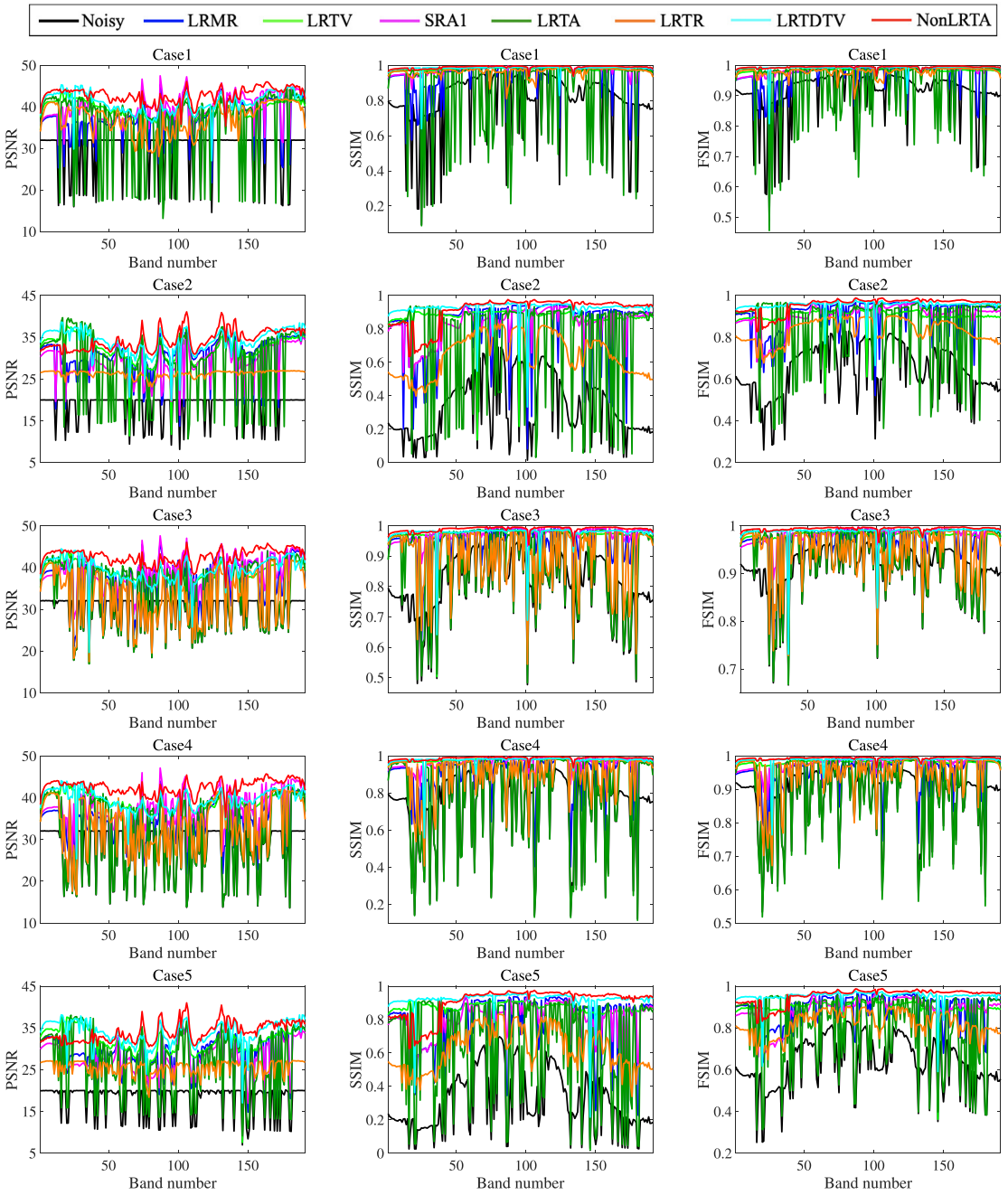

Fig. 2. PSNR, SSIM and FSIM values of each band in simulated experiment. The first to the fifth rows correspond to simulated experiments cases 1 to 5, respectively.

all bands are considered for restoration. The parameters of NonLRTA mothod are set as follows: $\mathbf{w} = (0.05, 0.05, 0.90)$, $\epsilon = 0.02$, $\lambda = 0.1$, and $\rho = 1$.

Performance evaluation. Figs. 8 and 9 show the original and restoration results of band 207 and band 104, respectively. From Fig. 8–9(a), it can be observed that the original image is contaminated by complex structural noise. Figs. 8 and 9 show that LRMR, LRTV, LRTA, LRTR, and LRTDTV remain at least one types of noise. SRA1 can remove all kinds of noise, but the edge exits shadow-like artifacts. The proposed method achieves a better performance than SRA1 with fewer artifacts in comparison.

Figs. 10 and 11 show the horizontal and vertical mean profiles of band 210 in the Urban dataset. From Fig. 10–11(a), we can see that the original image is polluted by mixed noise, especially stripes and impulse noise. The noise is less or more

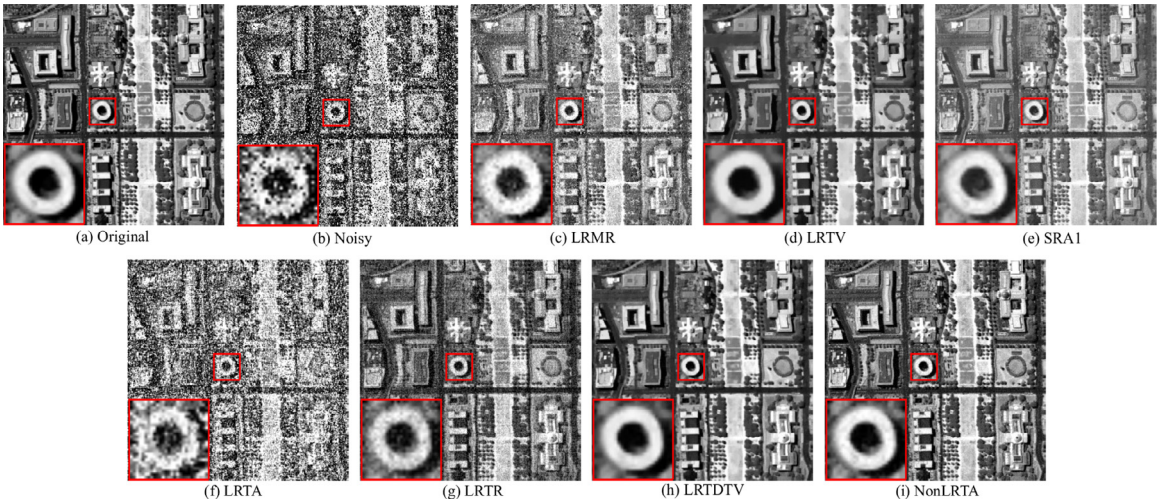

Fig. 3. (a) The original band 119 of Washington DC Mall dataset; (b) The noisy band of case 2; (c)-(i) The restored images of all compared methods. To facilitate the observation of details, the demarcated area in each image is enlarged.

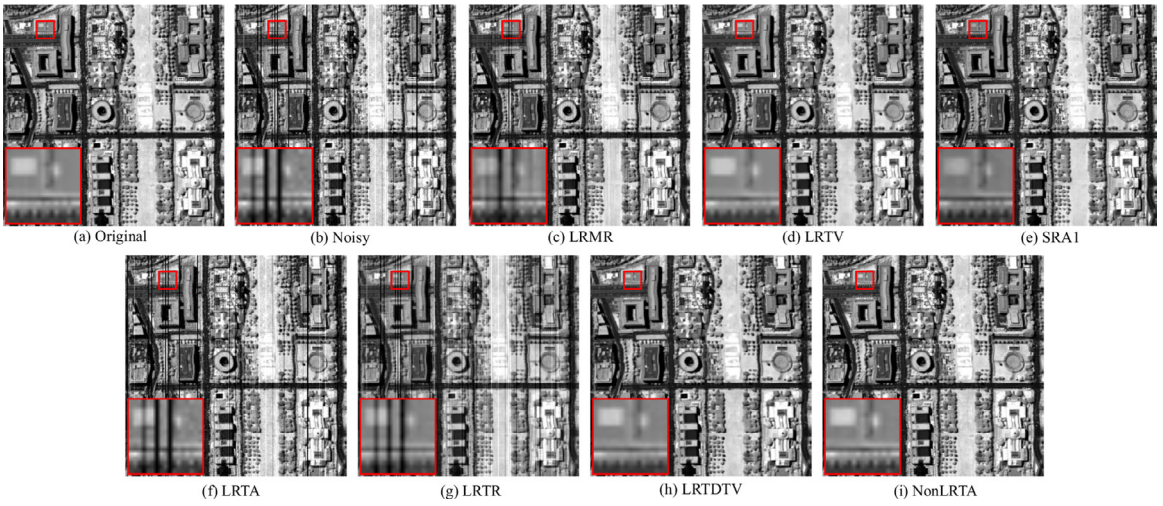

Fig. 4. (a) The original band 69 of Washington DC Mall dataset; (b) The noisy band of case 3; (c)-(i) The restored images of all compared methods. To facilitate the observation of details, the demarcated area in each image is enlarged.

suppressed by all competing methods after restoration. It is clear that the proposed method provides smoother curves, indicating better restoration performance.

Australia dataset. The raw Australia dataset contains 242 spectral bands and 3858×256 pixels per band. We extract a sub-image of size 399×225 for our experiment. The parameters settings of the NonLRTA method are the same as the Urban dataset.

Performance evaluation. Figs. 12–14 show the original and restoration results of band 49, band 108 and band 86. From Fig. 12, we can see that the original image is contaminated by complex structural noise. LRTV and LRTA cannot remove the noise; LRMR and LRTR remain some noise. LRTDTV satisfactorily removes the noise with edge information loss. SRA1 and NonLRTA achieve the best restoration with detail preserving. From Figs. 13 and 14, the original image is contaminated by deadlines and strips. LRTDTV, SRA1, LRTV, and LRMR remove strips and deadlines effectively, which can not be achieved by LRTA and LRTR. NonLRTA approximates the true low-rank tensor more precisely and obtains a better restoration result in stripes and deadline removal.

Fig. 15 shows the vertical mean profiles of band 123 in the Australia dataset. From Fig. 15(a), due to the mixed noise corruption, we can observe the rapid fluctuations in the curve. Compared with other restoration results, the result of NonLRTA offers a smoother curve, suggesting that the mixed noise has been removed finely.

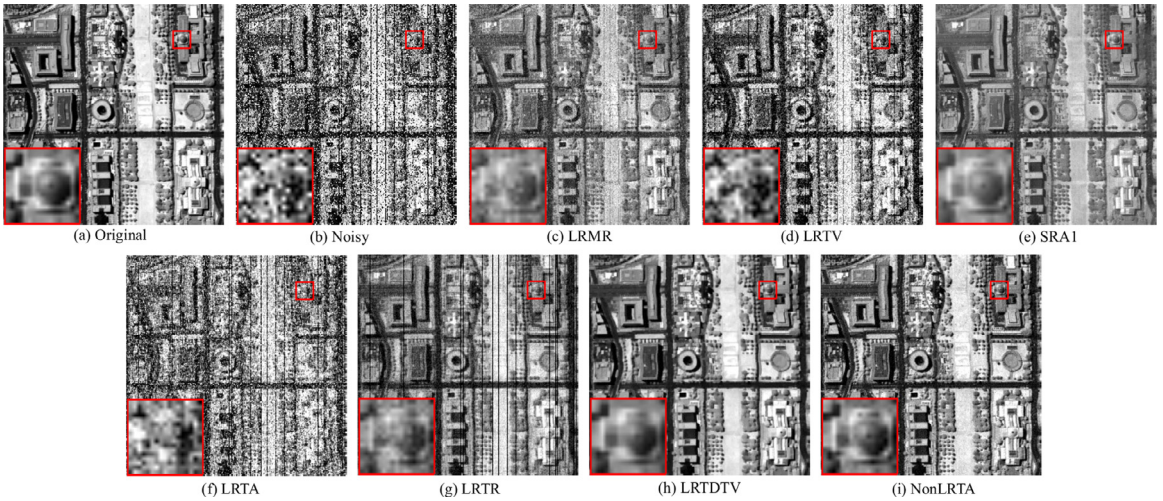

Fig. 5. (a) The original band 79 of Washington DC Mall dataset; (b) The noisy band of case 5; (c)-(i) The restored images of all compared methods. To facilitate the observation of details, the demarcated area in each image is enlarged.

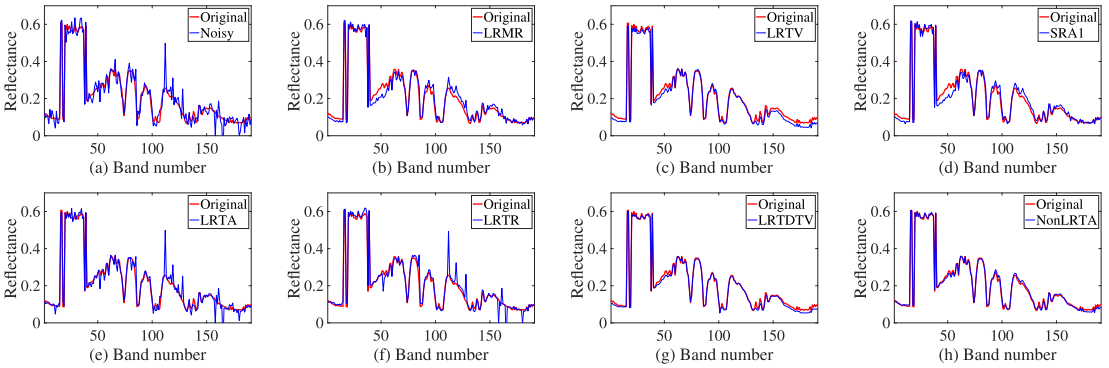

Fig. 6. Spectral signature curves of the (80,80) pixel restored by all the compared methods in Washington DC Mall dataset (case 3).

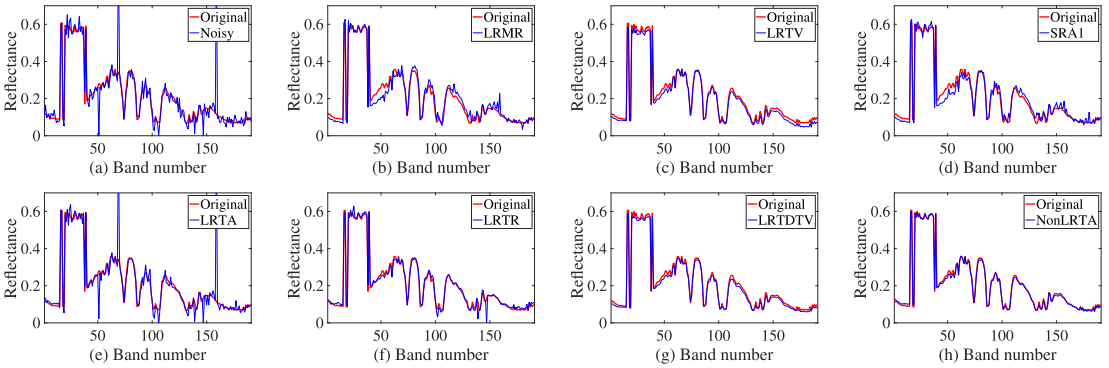

Fig. 7. Spectral signature curves of the (90,90) pixel restored by all the compared methods in Washington DC Mall dataset (case 4).

4.3. Discussion

Parameters study. We test the effects of the parameters on the restoration performance. 1) Parameter ϵ: Fig. 16(a) shows the MPSNR in simulated experiment by our method using difference values of ϵ from 0.005 to 0.06 with an increment 0.005. The MPSNR is relatively stable with the ϵ value is changed from 0.02 to 0.035. Thus we fix the ϵ to 0.02 in all experiments. 2) Weight $\mathbf{w}$: For HSIs, their spectral mode correlation is stronger than spatial mode correlation, which indicates that the rank of an HSI along its spectral mode is lower than that along its spatial modes. We set $\mathbf{w} = ((1-w_3)/2, (1-w_3)/2, w_3)$

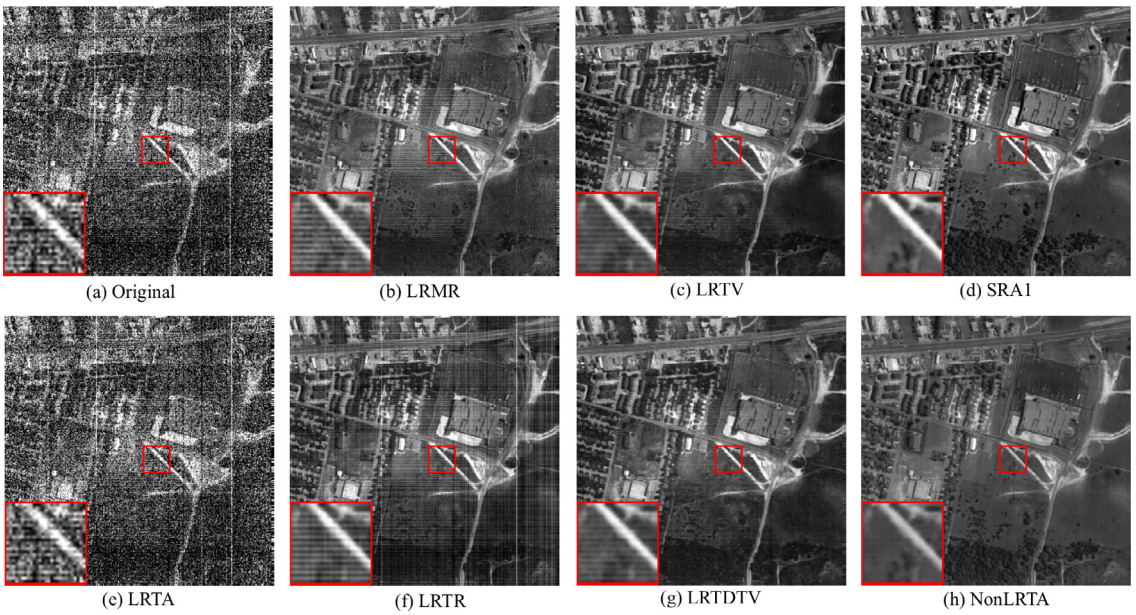

Fig. 8. (a) The original band 207 of Urban dataset polluted by mixed noise; (b)-(h) The restored images of all compared methods. To facilitate the observation of details, the demarcated area in each image is enlarged.

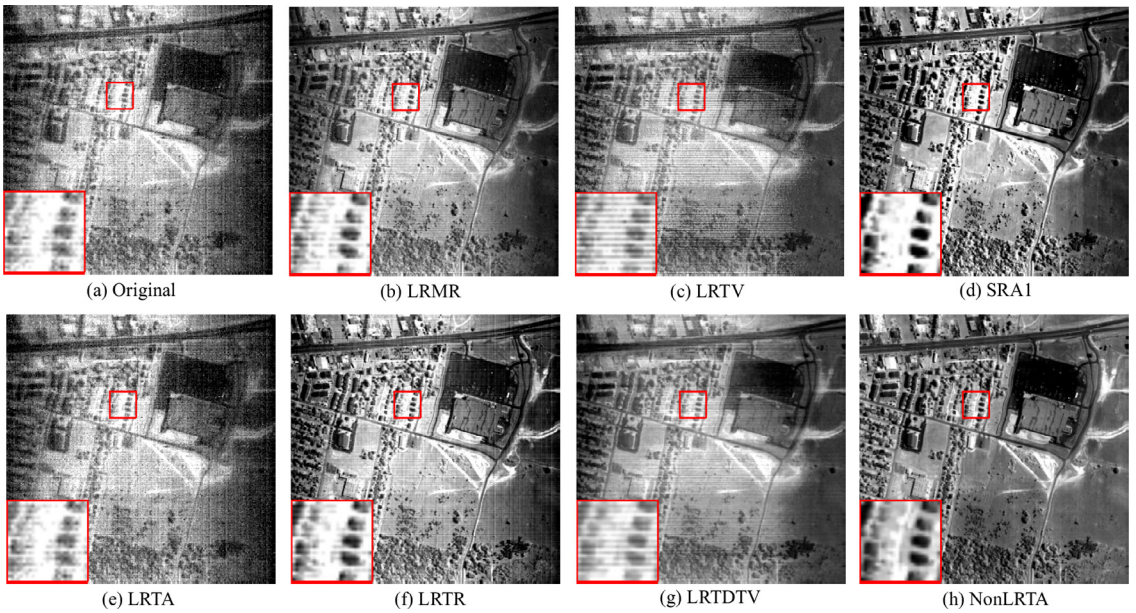

Fig. 9. (a) The original band 104 of Urban dataset polluted by mixed noise; (b)-(h) The restored images of all compared methods. To facilitate the observation of details, the demarcated area in each image is enlarged.

and Fig. 16(b) shows the MPSNR in simulated experiment by the proposed method using difference values of w_3 from 0 to 1 with an increment 0.1. The relatively satisfying MPSNR can be obtained with the w_1 is changed from 0.8 to 0.9. We suggest set $\mathbf{w} = (0.05, 0.05, 0.90)$ in all experiments. 3) Regularization parameter λ and ρ: In (4), they balance the regularization terms of $\mathcal{L}$, $\mathcal{S}$ and $\mathcal{N}$. Fig. 17 shows the MPSNR in simulated experiment with respect to the λ selected from the set of $\{0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4\}$ and $\rho = 10 * \lambda$. The proposed method maintains stable when λ is changed from 0.1 to 0.3 and ρ is changed from 1 to 3. Therefore, we suggest the λ is selected from the set of $\{0.1, 0.15, 0.2, 0.25, 0.3\}$ with the highest MPSNR in simulated experiment and $\rho = 10 * \lambda$.

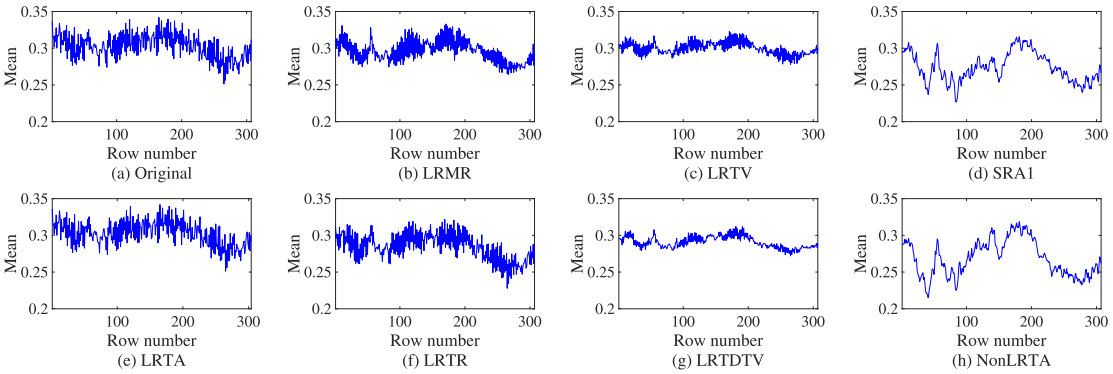

Fig. 10. (a) The horizontal mean profiles of band 210 of original Urban dataset; (b)-(h) The horizontal mean profiles of restored results of all compared methods.

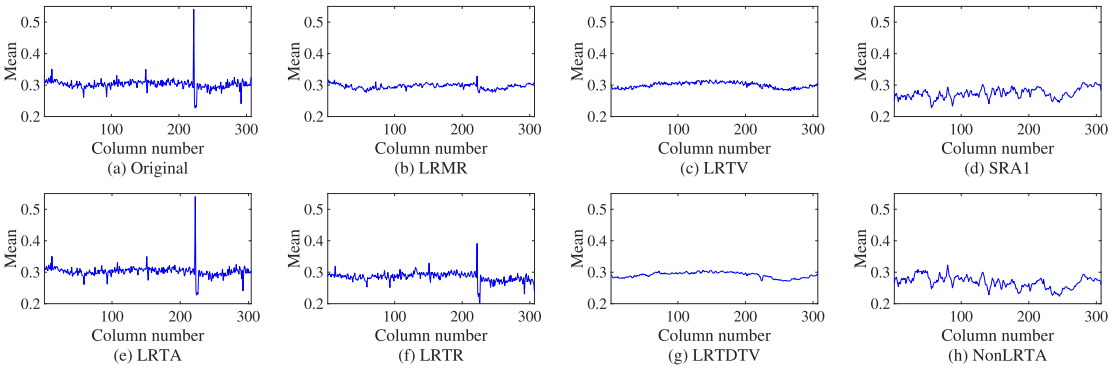

Fig. 11. (a) The vertical mean profiles of band 210 of original Urban dataset; (b)-(h) The vertical mean profiles of restored results of all compared methods.

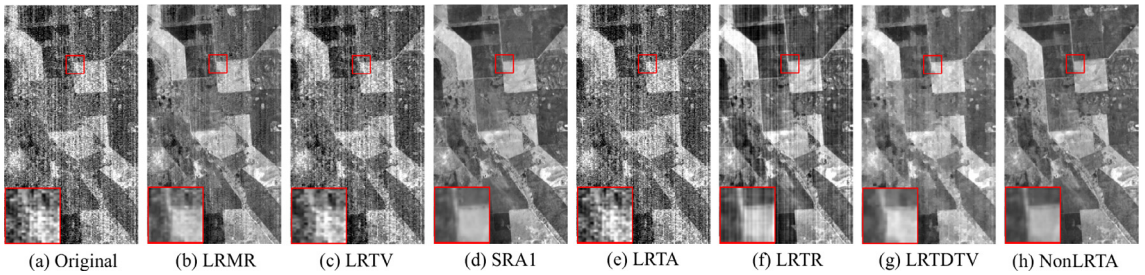

Fig. 12. (a) The original band 49 of Australia dataset polluted by mixed noise; (b)-(h) The restored images of all compared methods. To facilitate the observation of details, the demarcated area in each image is enlarged.

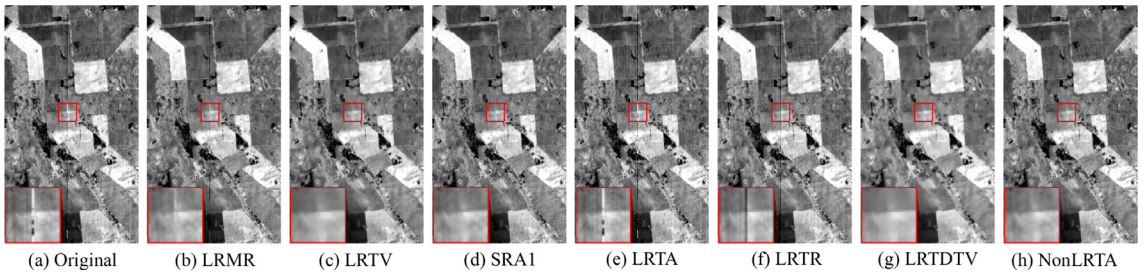

Fig. 13. (a) The original band 109 of Australia dataset mainly polluted by stripe; (b)-(h) The restored images of all compared methods. To facilitate the observation of details, the demarcated area in each image is enlarged.

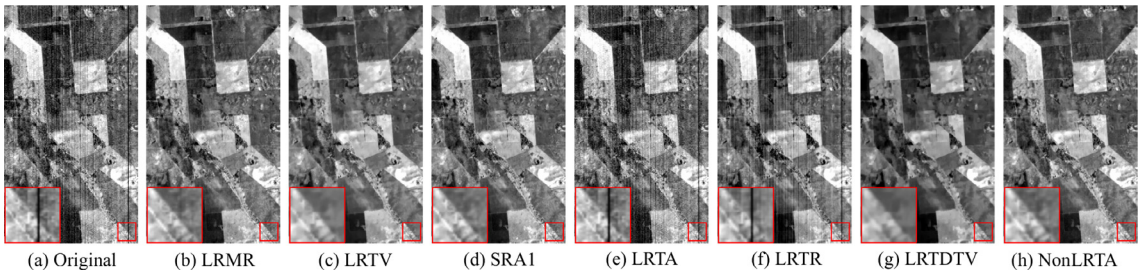

Fig. 14. (a) The original band 86 of Australia dataset mainly polluted by deadline; (b)-(h) The restored images of all compared methods. To facilitate the observation of details, the demarcated area in each image is enlarged.

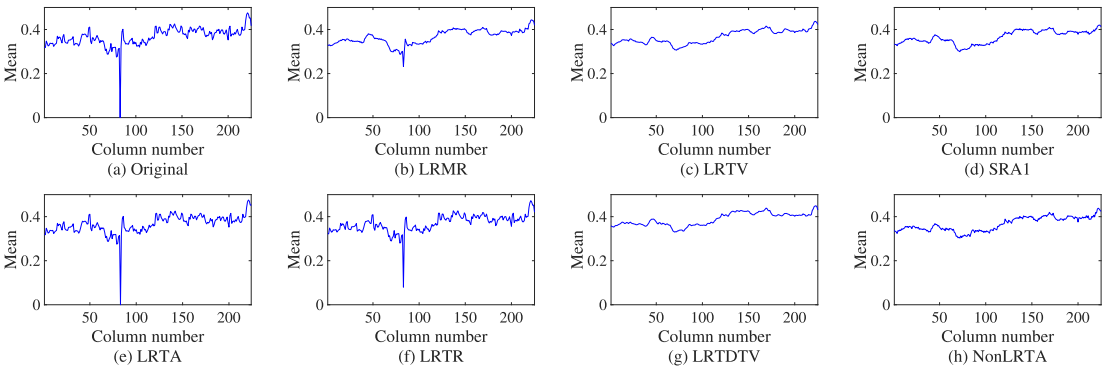

Fig. 15. (a) The vertical mean profiles of band 123 of original Australia dataset; (b)-(h) The vertical mean profiles of restored results of all compared methods.

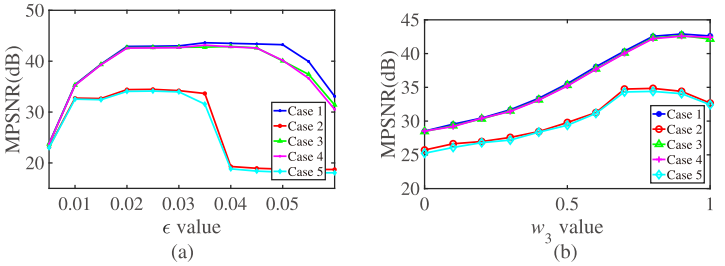

Fig. 16. (a) MPSNR values of the different ϵ values in cases 1–5 of the simulated experiment; (b) MPSNR values of different w_3 values in cases 1–5 of the simulated experiment..

Algorithm 1 Solve the proposed NonLRTA model by ALM.

Input: The noisy HSI $\mathcal{X}$, the weight $\mathbf{w} = (w_1, w_2, w_3)$, and the regularization parameters λ, ρ, α_i and β.

Initialize: $\mathcal{M}_i = \mathcal{L} = \mathcal{S} = \mathcal{N} = \mathcal{O}$, $\mathcal{Y}_i = \mathcal{Z} = \mathcal{O}$, $\delta = 10^{-4}$, $r = 1.1$, and $\alpha_{\max} = \beta_{\max} = 10^6$.
while not converged **do**
  Update $\mathcal{M}_i^{k+1} (i = 1, 2, 3)$ by (8);
  Update $\mathcal{L}^{k+1}$ by (11);
  Update $\mathcal{S}^{k+1}$ by (13);
  Update $\mathcal{N}^{k+1}$ by (15);
  Update $\mathcal{Y}_i^{k+1} (i = 1, 2, 3)$ and $\mathcal{Z}^{k+1}$ by (5);
  Update $\alpha_i (i = 1, 2, 3)$ and β by (16);
  Check the convergence condition: $\|\mathcal{X} - \mathcal{L} - \mathcal{S} - \mathcal{N}\|_F^2 / \|\mathcal{X}\|_F^2 \leq \delta$.
end while

Output: The restored HSI $\mathcal{L}$.

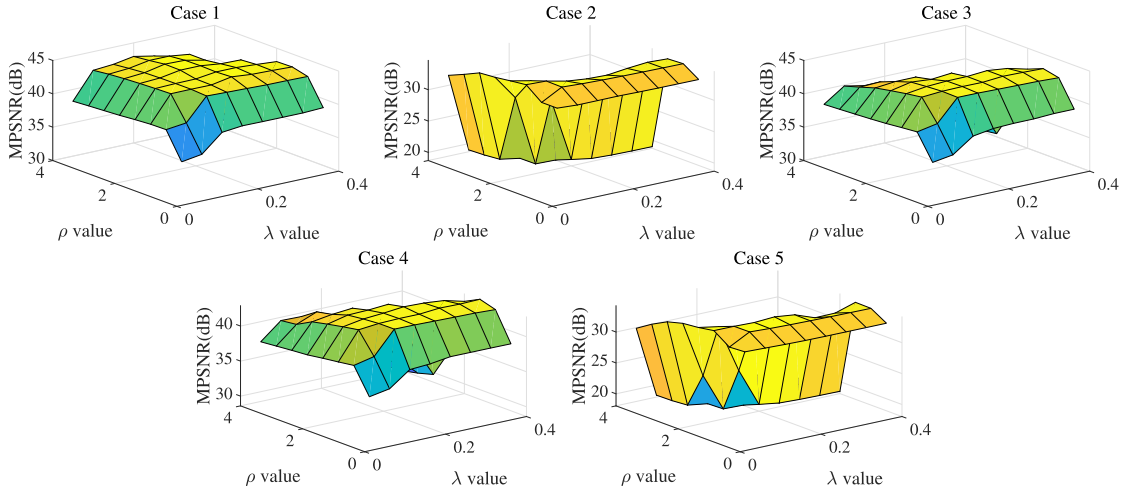

Fig. 17. MPSNR values of different parameters λ and ρ in cases 1–5 of the simulated experiment..

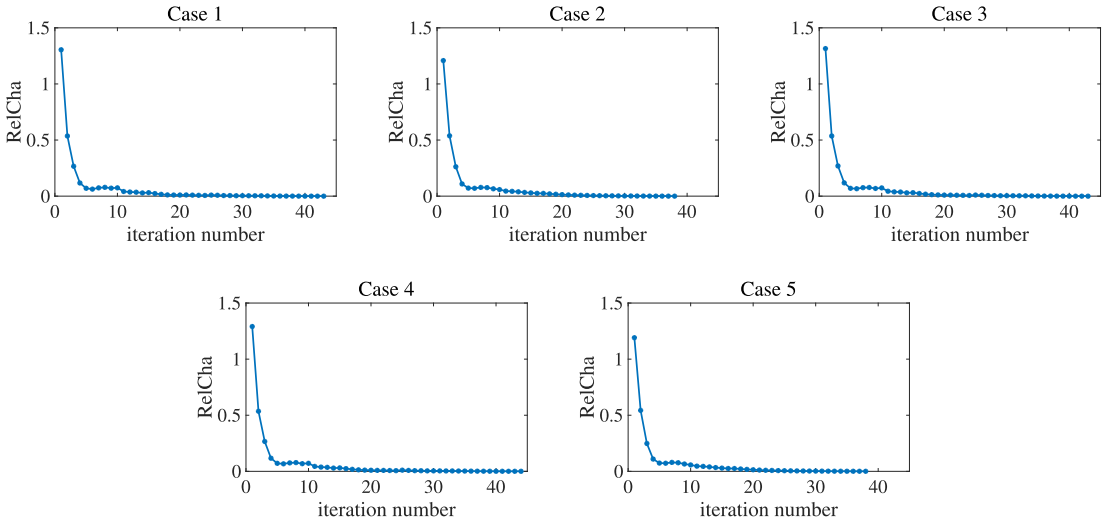

Fig. 18. The RelCha values in each iteration in cases 1–5 of the simulated experiment..

Convergence behavior. We present an analysis of the convergence behavior of the ALM algorithm for NonLRTA. The relative change (RelCha) of the restored tensor is defined as

$$\text{RelCha} = \frac{\|\mathcal{L}^{k+1} - \mathcal{L}^{k}\|_F}{\|\mathcal{L}^{k}\|_F}$$

Fig. 18 shows the RelCha gains versus the iteration number of the proposed algorithm in cases 1–5 of the simulated experiments. It is clear that the RelCha rapidly tends to zero within 45 iterations. This implies the empirical convergence behavior of the proposed method.

5. Conclusions

In this paper, we have proposed a NonLRTA model for HSI restoration. We model the clean HSI as a low-rank tensor and decompose the mixed noise into a sparse component and a Gaussian component. The proposed model is efficiently solved by the ALM algorithm. Simulated and real experiments indicate that our method performs better than the state-of-the-art methods and achieves the trade-off between detail preservation and noise removal.

Acknowledgment

The research is supported by the National Natural Science Foundation of China (61772003, 61876203, 11901450, 12001446), the Key Project of Applied Basic Research in Sichuan Province (2020YJ0216), the Applied Basic Research Project of Sichuan Province (2021YJ0107), the National Key Research and Development Program of China (2020YFA0714001), and the Fundamental Research Funds for the Central Universities (JBK2102001).

References

[1] Y. Chen, W. He, N. Yokoya, T.-Z. Huang, Nonlocal tensor-ring decomposition for hyperspectral image denoising, IEEE Trans. Geosci. Remote Sens. 58 (2) (2020) 1348–1362.
[2] H. Zhang, L. Liu, W. He, L. Zhang, Hyperspectral image denoising with total variation regularization and nonlocal low-rank tensor decomposition, IEEE Trans. Geosci. Remote Sens. 58 (5) (2020) 3071–3084.
[3] Y. Chang, L. Yan, H. Fang, S. Zhong, W. Liao, HSI-Denet: hyperspectral image restoration via convolutional neural network, IEEE Trans. Geosci. Remote Sens. 57 (2) (2018) 667–682.
[4] H. Zhang, J. Cai, W. He, H. Shen, L. Zhang, Double low-rank matrix decomposition for hyperspectral image denoising and destriping, IEEE Trans. Geosci. Remote Sens., doi:10.1109/TGRS.2021.3061148.
[5] J.-L. Wang, T.-Z. Huang, X.-L. Zhao, T.-X. Jiang, M.K. Ng, Multi-dimensional visual data completion via low-rank tensor representation under coupled transform, in: IEEE Trans. Image Process, doi:10.1109/TIP.2021.3062995.
[6] W. He, N. Yokoya, L. Yuan, Q. Zhao, Remote sensing image reconstruction using tensor ring completion and total variation, IEEE Trans. Geosci. Remote Sens. 57 (11) (2019) 8998–9009.
[7] J.-H. Yang, X.-L. Zhao, T.-H. Ma, C. Yong, T.-Z. Huang, M. Ding, Remote sensing image destriping using unidirectional hybrid total variation and non-convex low-rank regularization, J. Comput. Appl. Math. 363 (2020) 124–144.
[8] Y.-B. Zheng, T.-Z. Huang, X.-L. Zhao, Y. Chen, W. He, Double-factor-regularized low-rank tensor factorization for mixed noise removal in hyperspectral image, IEEE Trans. Geosci. Remote Sens. 58 (12) (2020) 8450–8464.
[9] W. He, Q. Yao, C. Li, N. Yokoya, Q. Zhao, H. Zhang, L. Zhang, Non-local meets global: an integrated paradigm for hyperspectral image restoration, IEEE Trans. Pattern Anal. Mach. Intell. doi:10.1109/TPAMI.2020.3027563.
[10] X.-L. Zhao, W. Wang, T.-Y. Zeng, T.-Z. Huang, M.K. Ng, Total variation structured total least squares method for image restoration, SIAM J. Sci. Comput. 35 (6) (2013) 1304–1320.
[11] X.-L. Zhao, F. Wang, M.K. Ng, A new convex optimization model for multiplicative noise and blur removal, SIAM J. Imaging Sci. 7 (1) (2014) 456–475.
[12] Y.-Y. Liu, X.-L. Zhao, Y.-B. Zheng, T.-H. Ma, H. Zhang, Hyperspectral image restoration by tensor fibered rank constrained optimization and plug-and-play regularization, IEEE Trans. Geosci. Remote Sens. doi:10.1109/TGRS.2020.3045169.
[13] J. Yao, D. Meng, Q. Zhao, W. Cao, Z. Xu, Nonconvex-sparsity and nonlocal-smoothness-based blind hyperspectral unmixing, IEEE Trans. Image Process. 28 (6) (2019) 2991–3006.
[14] D. Manolakis, G. Shaw, Detection algorithms for hyperspectral imaging applications, IEEE Signal Process. Mag. 19 (1) (2002) 29–43.
[15] L. Zhuang, L. Gao, B. Zhang, X. Fu, Bioucas-Dias J.M., Hyperspectral image denoising and anomaly detection based on low-rank and sparse representations, IEEE Trans. Geosci. Remote Sens., doi:10.1109/TGRS.2020.3040221.
[16] J. Gao, L. Xu, A. Shi, F. Huang, A kernel-based block matrix decomposition approach for the classification of remotely sensed images, Appl. Math. Comput. 228 (2014) 531–545.
[17] J. Wright, A. Ganesh, S. Rao, Y. Peng, Y. Ma, Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization, in: in Proceedings of the NIPS, 2009, pp. 2080–2088.
[18] H. Zhang, W. He, L. Zhang, H. Shen, Q. Yuan, Hyperspectral image restoration using low-rank matrix recovery, IEEE Trans. Geosci. Remote Sens. 52 (8) (2014) 4729–4743.
[19] W. He, H. Zhang, L. Zhang, H. Shen, Total-variation-regularized low-rank matrix factorization for hyperspectral image restoration, IEEE Trans. Geosci. Remote Sens. 54 (1) (2016) 178–188.
[20] H. Liu, P. Sun, Q. Du, Z. Wu, Z. Wei, Hyperpsectral image restoration based on low-rank recovery with a local neighborhood weighted spectral-spatial total variation model, IEEE Trans. Geosci. Remote Sens. 57 (3) (2019) 1409–1422.
[21] S. Gu, Q. Xie, D. Meng, W. Zuo, X. Feng, L. Zhang, Weighted nuclear norm minimization and its applications to low level vision, Int. J. Comput. Vision 121 (2) (2017) 183–208.
[22] J. Xue, Y. Zhao, W. Liao, S.G. Kong, Joint spatial and spectral low-rank regularization for hyperspectral image denoising, IEEE Trans. Geosci. Remote Sens. 56 (4) (2018) 1940–1958.
[23] A. Buades, B. Coll, J.M. Morel, A non-local algorithm for image denoising, in: Proceedings of the CVPR, 2005, pp. 60–65.
[24] J. Mairal, F. Bach, J. Ponce, G. Sapiro, A. Zisserman, Non-local sparse models for image restoration, in: Proceedings of the ICCV, 2009, pp. 2272–2279.
[25] Zhuang L., Fu X., Ng M.K., Bioucas-Dias J.M., Nonlocally centralized sparse representation for image restoration, IEEE Trans. Geosci. Remote Sens. doi:10.1109/TGRS.2020.3046038.
[26] Z. Kang, C. Peng, Q. Cheng, Robust PCA via nonconvex rank approximation, in: Proceedings of the ICDM, 2015, pp. 211–220.
[27] H. Ye, H. Li, B. Yang, F. Cao, Y. Tang, A novel rank approximation method for mixture noise removal of hyperspectral images, IEEE Trans. Geosci. Remote Sens. 57 (7) (2019) 4457–4469.
[28] Lin J., Huang T.-Z., Zhao X.-L., Jiang T.-X., Zhuang L., A tensor subspace representation-based method for hyperspectral image denoising, IEEE Trans. Geosci. Remote Sens., doi:10.1109/TGRS.2020.3032168.
[29] H. Fan, C. Li, Y. Guo, G. Kuang, J. Ma, Spatial-spectral total variation regularized low-rank tensor decomposition for hyperspectral image denoising, IEEE Trans. Geosci. Remote Sens. 56 (10) (2018) 6196–6213.
[30] T.G. Kolda, B.W. Bader, Tensor decompositions and applications, SIAM Rev. 51 (3) (2009) 455–500.
[31] L.R. Tucker, Some mathematical notes on three-mode factor analysis, Psychometrika. 31 (3) (1966) 279–311.
[32] M.E. Kilmer, C.D. Martin, Factorization strategies for third-order tensors, Linear Algebra Appl. 435 (3) (2011) 641–658.
[33] X. Liu, S. Bourennane, C. Fossati, Denoising of hyperspectral images using the PARAFAC model and statistical performance analysis, IEEE Trans. Geosci. Remote Sens. 50 (10) (2012) 3717–3724.
[34] X. Guo, X. Huang, L. Zhang, L. Zhang, Hyperspectral image noise reduction based on rank-1 tensor decomposition, ISPRS J. Photogramm. Remote Sens. 83 (2013) 50–63.
[35] N. Renard, S. Bourennane, J. Blanc-Talon, Denoising and dimensionality reduction using multilinear tools for hyperspectral images, IEEE Geosci. Remote Sens. Lett. 5 (2) (2008) 138–142.
[36] A. Karami, M. Yazdi, A.Z. Asli, Noise reduction of hyperspectral images using kernel non-negative tucker decomposition, IEEE J. Sel. Top. Signal Process. 5 (3) (2011) 487–493.
[37] D. Muti, S. Bourennane, Survey on tensor signal algebraic filtering, Signal Processing. 87 (2) (2007) 237–249.
[38] Y. Wang, J. Peng, Q. Zhao, Y. Leung, X.-L. Zhao, D. Meng, Hyperspectral image restoration via total variation regularized low-rank tensor decomposition, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens 11 (4) (2018) 1227–1243.
[39] J. Liu, P. Musialski, P. Wonka, J. Ye, Tensor completion for estimating missing values in visual data, IEEE Trans. Pattern Anal. Mach. Intell. 35 (1) (2012) 208–220.

[40] Z. Zhang, D. Liu, S. Aeron, A. Vetro, An online tensor robust PCA algorithm for sequential 2D data, in: Proceedings of the ICASSP, 2016, pp. 2434–2438.
[41] H. Fan, Y. Chen, Y. Guo, H. Zhang, G. Kuang, Hyperspectral image restoration using low-rank tensor recovery, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10 (10) (2017) 4589–4604.
[42] Y.-B. Zheng, T.-Z. Huang, X.-L. Zhao, T.-X. Jiang, T.-H. Ma, T.Y. Ji, Mixed noise removal in hyperspectral image via low-fibered-rank regularization, IEEE Trans. Geosci. Remote Sens. 58 (1) (2020) 734–749.
[43] Z. Lin, M. Chen, Y. Ma, The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices, in: arXiv. math.OC, 2010, pp. 116–127, doi:10.1016/j.jsb.2012.10.010.
[44] J.-F. Cai, E.J. Cands, Z. Shen, A singular value thresholding algorithm for matrix completion, SIAM J. Optim. 20 (4) (2010) 1956–1982.
[45] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Trans. Image Process. 13 (4) (2004) 600–612.
[46] L. Zhang, L. Zhang, X. Mou, D. Zhang, FSIM: A feature similarity index for image quality assessment, IEEE Trans. Image Process. 20 (8) (2011) 2378–2386.