

Research Article

Remote Sensing Image Reconstruction Based on Improved Adaptive Tensor Nuclear Norm Minimization

Jinliang Li,¹ Yang Yang ,² Shuai Sun,² Chao Lv ,³ Peng Liu ,² Xiaohai Qu,² and Sai Lei⁴

¹College of Materials Science and Engineering, Jilin University, Changchun 130012, China

²Engineering Training Center, Jilin University, Changchun 130012, China

³College of Electronic Science and Engineering, Jilin University, Changchun, Jilin 130012, China

⁴Key Laboratory of Micro/Nano and Ultra-Precision Manufacturing (Jilin Province), School of Mechatronic Engineering, Changchun University of Technology, Changchun 130012, China

Correspondence should be addressed to Yang Yang; yyjl@jlu.edu.cn and Chao Lv; chaolv@jlu.edu.cn

Received 21 May 2022; Revised 15 September 2022; Accepted 28 October 2022; Published 1 December 2022

Academic Editor: Ana C. Teodoro

Copyright © 2022 Jinliang Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Incomplete remote sensing image observation data require proper processing to restore the original data. Nevertheless, the lack of information, due to noise and dead pixels, along with the quality of the acquired images that may be severely degraded, limits the accuracy of later processing. Therefore, this paper proposes a remote sensing image restoration scheme that employs the low-rank structure of the Bayesian probability tensor latent space and a new tensor completion method, which has strong robustness for the model selection process of the following remote sensing image data analysis. Considering the data space, we incorporate an adaptive nuclear norm regularization into the processing of latent factors that performs singular value decomposition on a more accurate scale. The proposed model adopts two algorithms: Augmented Lagrange Multiplier (ALM) and Alternating Least Squares (ALS). Several synthetic and real-data-based experiments illustrate the inherent ability of our method in restoring original information and preventing overfitting problems, even when a vast number of items are missing. In addition, the combination of adaptive nuclear norm regularization and tensor decomposition is stable, settling a high sensitivity of rank selection and low computational efficiency in traditional tensor reconstruction methods.

1. Introduction

Remote sensing instruments have aroused widespread concern in the research field, as these are highly appealing for observing, learning, and understanding Earth. Nevertheless, severe weather conditions and sensor failures cause the remote sensing images to be heavily disturbed during the acquisition process, significantly reducing their visual quality and affecting the accuracy of the applications where these images are applied [1, 2]. Consequently, extracting raw data from remote sensing images containing dead pixels, clouds, and noise is a significant study subject [3–5]. However, to exploit efficiently raw remote sensing imagery, it is necessary to forecast and reconstruct missing information [6].

In recent years, advanced technologies such as wavelet compression and anisotropic diffusion have been applied in remote sensing image restoration, yet they neglect the relativity between spatial pixels resulting in poor post-processing results. Literature offers numerous relevant studies, such as proposing an effective nonlocal method [7] that analyzes the structural similarity through the gray correlation of coefficients and then setting the similarity weight function. Li et al. used a multitemporal dictionary learning algorithm to recover remote sensing images polluted by large thick clouds and the consequent cloud shadows and verified the effectiveness of the two methods [8]. Lorenzi et al. propose three alternatives for reconstructing missing regions in images acquired by very high spatial resolution (VHR) sensors based on the inpainting

method [9]. Zhuang and Bioucas–Dias develop a very sparse and compact Hyper-Spectral Imagery (HSI) representation using HSI's low-rank self-similar feature [3]. Xue et al. proposed a new HSI noise reduction method based on CP decomposition modeling with tensor nonlocal LR and automatic rank determination. The method considers the a priori knowledge of GCS and NSS in space and combines the rank estimation of CP decomposition. The results show that the method can effectively reduce the complex noise and maintain the fine weave structure of HSI [10]. Zeng et al. proposed an HSI recovery model based on local tensor kernel parametric minimization and global L1-2SSTV regularization for the HSI recovery problem. The results of extensive experimental simulations show that the model has a good quality evaluation index and visual effect of model recovery results [11]. Dian and Li proposed a subspace-based low-tensor multirank regularization method to improve the spatial resolution of HSI. Finally, the superiority of the method is demonstrated by experimental results on simulated and real datasets [12]. Yang et al. proposed a variational depth HSI-MSI fusion model, and then expanded the fusion model into a DL network, in which the degenerate model and the prior data can be represented by the DL network, and validated the effectiveness of the method from experiments on several HSI datasets [13]. The purpose of both methods is to eliminate one or two types of noise. However, the acquired remote sensing images are generally mixed with various noise types, i.e., edge destruction, dashed noise, impulse noise, and Gaussian noise, and therefore a solution restoring the image and enhance its quality regardless of the noise type is mandatory.

Low-order tensors are an effective means to describe the inherent geometric characteristics of remote sensing images. Some methods involve low-rank matrix modeling, e.g., Zhao and Yang exploit the low rank of high-order matrices and establish a low-rank matrix approximation based on small blocks [14]. Hao and Su suggest noise adjustment and propose an iterative low-rank matrix approximation denoising method, which uses the nonlocal self-similarity of the image. This method adopts a new patch-based model to restore the low-rank tensor applied to denoise multiframe images [15]. Image restoration is essentially the process of estimating the original image, and its purpose is to obtain the optimum estimate of the original image under specific objective standards. The prior knowledge of the image degradation process is related to the effect of image restoration. For reconstructing remote sensing images, low-rank tensor completion is one of the most effective methods to integrate prior knowledge [15, 16]. For instance, both the Tucker and CandeComp/Parafac (CP) models rely on prior knowledge of the model's original multidimensional data [17]. The CP decomposition can be solved by Alternating Least Squares (ALS), that is, updating a single variable by other fixed variables at each round of iteration. However, this model is nonconvex, subject to local minima, and requires good initialization to get better results [18]. For fixed n -rank, the uniqueness of Tucker decomposition cannot be guaranteed, and some constraints are generally added, such as orthogonality

constraints on the factor units obtained from the decomposition [19, 20]. Therefore, we need a new tensor completion method to enhance the robustness of the model selection process for subsequent remote sensing image data analysis.

Low-rank tensors utilize rank minimization to estimate low-rank approximations of incomplete data [21–23]. However, low-order tensors may cause overfitting problems due to neglecting critical information of the original data. The Bayesian probability inference is used to overcome this, as the Bayesian probability model requires fewer observation data to form inferences. It can overcome extremely scarce conditions and can be effectively combined with tensor decomposition [24, 25]. Consequently, the matrix/tensor decomposition probability model has received widespread attention in collaborative filtering and matrix/tensor completion. Hence, [26] proposes the probability matrix factorization combined with the Markov Chain Monte Carlo (MCMC) inference to perform complete Bayesian processing. Babacan [27] uses variational Bayesian inference to perform complete Bayesian processing on the tensor structure, while [28] further expands the nonparametric variables and robust variables. Chu and Ghahramani [29] represent the probability framework of tensor decomposition, with other variants involving the extension of the nonparametric Bayesian model [30] and the exponential family model [31].

This article introduces a complete Bayesian probability tensor decomposition model that is appropriate for the TR decomposition framework. Precisely, based on the predicted distribution, incomplete tensor, and the noise of missing items, a potential multilinear factor is calculated to implicitly and automatically determine the rank of the real tensor. To this end, a sparse-induced hierarchical prior based on multiple factor matrices is proposed, and the hyperparameters of each factor matrix are related to each potential dimension so that the elements in the factor matrix are limited and minimal. All model parameters are considered variables that can be assumed, including noise accuracy and corresponding priors. Aiming at the shortcomings of inherent high sensitivity and low efficiency of rank selection of traditional tensors, we design a combination of adaptive tensor nuclear normalization methods. The model comprises Bayesian probability tensor decomposition and an adaptive kernel model and affords good performance, robustness, and high computational efficiency. The Alternating Least Square (ALS) and Augmented Lagrange Multiplier (ALM) algorithms are adopted to optimize this model.

This paper primarily studies the problem of remote sensing image reconstruction under the influence of mixed noise. The main contributions of this work are summarized as follows:

- (i) A Bayesian probability tensor decomposition method is used to fully exploit the global space information heavily affected by complex noise. This strategy affords a low computational burden, and the tensor decomposition model is constructed using actual observation images.

- (ii) The designed adaptive nuclear norm regularizes the tensor's decomposition rank, affording our algorithm to obtain a stable solution. The results manifest that the regularization enhances the remote sensing image information and better reserves the specific information that remote sensing images capture.
- (iii) The Augmented Lagrange Multiplier (ALM) and Alternating Least Square (ALS) algorithms are employed to solve the adaptive nuclear norm regularized Bayesian probability tensor decomposition model. A considerable amount of simulations and actual data experiments demonstrate that the algorithm is more effective than current commonly used methods, both in visual comparison and quantitative evaluation.

This essay is structured as follows: Section 2 introduces the Bayesian Gaussian TR decomposition and the remote sensing image reconstruction problem formulation, which will introduce our model. In Section 3, the adaptive nuclear norm regularized Bayesian probability tensor decomposition model is introduced, and the ALS and ALM algorithms are adopted to solve the proposed model. Extensive experimental simulations of both methods are implemented in Section 4; while finally, Section 5 concludes this work.

2. Bayesian Gaussian TR Decomposition and Problem Formulation

2.1. Notations. In this paper, the notations of Kolda and Bader [32] are utilized, i.e., tensors of order $N \geq 3$ are represented by $\chi \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$, matrices by $X \in \mathbb{R}^{I \times J}$, vectors by $x \in \mathbb{R}^I$, scalars $X \in \mathbb{R}$. $X(i_1, \dots, i_N)$ represents the element of tensor χ at position (i_1, \dots, i_N) . Furthermore, two kinds of tensor expansion processing types are adopted, where the first mode- n expansion [31] of tensor $\chi \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ is represented by $X_{(n)} \in \mathbb{R}^{I_n \times I_1 \dots I_{n-1} I_{n+1} \dots I_N}$, and the second one, which is usually adopted in Symmetric tensor decomposition processing, is represented by $X_{(n)} \in \mathbb{R}^{I_n \times I_{n+1} \dots I_N I_1 \dots I_{n-1}}$.

2.2. Bayesian Gaussian TR Decomposition. Tensor decomposition is widely used to reduce the storage cost and operational complexity of high-dimensional data and involves a linear scale transformation of the input dimension instead of an exponential scale transformation. $\chi \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$ refers to a higher-order tensor, where is the size of χ . The TR decomposition of χ can be considered as follows:

$$\begin{aligned} \chi(i_1, \dots, i_N) &= TR(\mathcal{T}_1(i_1) \cdots \mathcal{T}_N(i_N)) \\ &= TR\left\{\prod_{d=1}^N \mathcal{T}_d(i_d)\right\}, \end{aligned} \quad (1)$$

where $X(i_1, \dots, i_N)$ denotes the (i_1, \dots, i_N) -th element of the tensor and $\mathcal{T}_d(i_d)$ is the i_d -th lateral slice matrix of the latent tensor \mathcal{T}_d .

Regarding the Bayesian probability model, it is assumed that the noise contained in the observed entry obeys a Gaussian distribution:

$$\tilde{\chi}(i_1, i_2, \dots, i_N) \sim \mathfrak{M}(\chi(i_1, i_2, \dots, i_N), \varsigma^{-1}), \quad (2)$$

where $\mathfrak{M}(\cdot)$ represents the Gaussian distribution and ς is the accuracy.

To better fit the tensor data, a flexible prior part is placed on $\mathcal{T}_N(i_N)$ and ς , and it is assumed that this prior distribution applies to all factor matrices and is multivariate Gaussians:

$$\mathcal{T}(i_1, i_2, \dots, i_N) \sim \mathfrak{M}(\varphi_i, \nu_i^{-1}). \quad (3)$$

To boost the model's robustness and efficiency, the conjugate Gaussian-Wishart priors for hyperparameters φ_i and ν_i are placed [33].

$$(\varphi_i, \nu_i) \sim \text{Gaussian} - \text{Wishart}(\rho_0, \kappa_0, W_0, \eta_0)$$

$$p(\varphi_i, \nu_i | -) = \mathfrak{M}(\varphi_i | \rho_0, (\kappa_0 \nu_i)^{-1}) \times \text{Wishart}(\nu_i | W_0, \eta_0), \quad (4)$$

where η_0 is the degree of freedom of the Wishart distribution and W_0 is given by the following equation:

$$\text{Wishart}(\nu_i | W_0, \eta_0) = \frac{1}{C} |\nu_i|^{(\eta_0 - d - 1/2)} \exp\left\{-\frac{1}{2} \sum_{i=1}^d W_0^{-1} \nu_i\right\}. \quad (5)$$

The entire graphical structure characterizing the Bayesian probabilistic TR decomposition data generation process is illustrated in Figure 1. Given that the model has a ringlike graphical appearance, it is intuitively called tensor ring decomposition. In this graph model, χ is the observation data, $\mathcal{T}_d(i_d)$ and ς are the parameters, and φ_i and ν_i are the superparameters. For further information on setting these superparameters when initializing the sampling algorithm, the reader is referred to [23].

2.3. Problem Formulation. In many practical situations, the observed remote sensing image will be contaminated by various noise types [25, 34, 35]. Hence, a remote sensing image y can be represented by a linear regression model:

$$\mathcal{Y} = \chi + \varepsilon, \quad (6)$$

where \mathcal{Y} and ε have the same size representing the observed vertical image and the mixed noise items, respectively. The purpose of rereflecting the image is to predict the complete reflection of image X based on the observed incomplete data Y .

In this work, the Gaussian noise term \mathcal{E} and the sparse noise term P are divided from noise E . Considering that the sparse noise term includes stripes and impulsive noise:

$$\mathcal{Y} = \chi + \mathcal{E} + P. \quad (7)$$

Therefore, the Frobenius norm and l_1 norm can be used naturally to simulate these two noise terms \mathcal{E} and P , respectively.

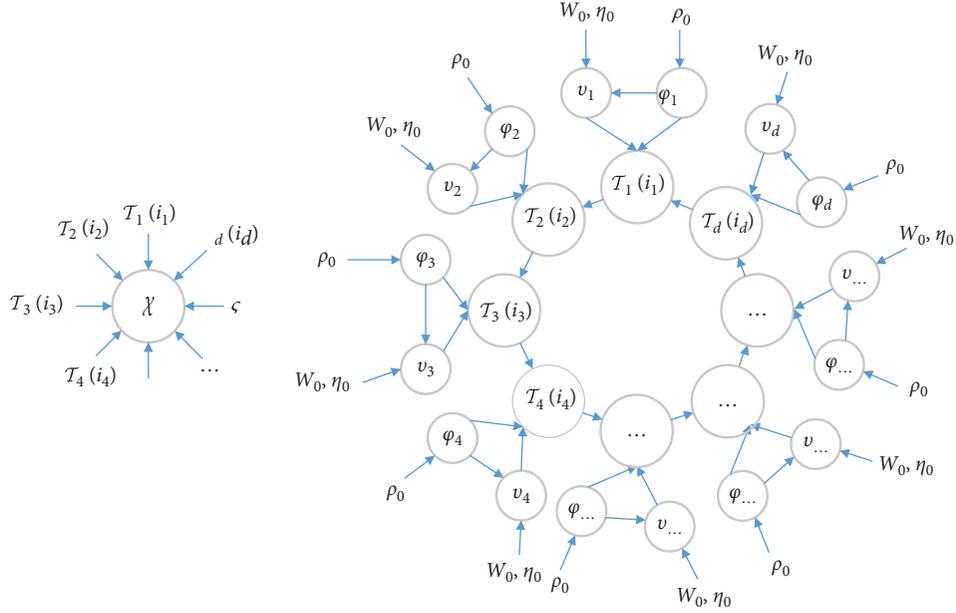


FIGURE 1: The probabilistic graphical models of TR decomposition for tensors.

3. Remote Sensing Image Restoration

3.1. Proposed Model. The purpose of tensor decomposition is to search for latent factors in the tensor data, which can be regarded as latent features of tensor data. The observation object is represented by a low-rank structure that is generated through tensor decomposition. According to the definition of tensor decomposition, an optimization model is drawn up:

$$\min_{\mathcal{T}} \left\| E_{\Omega}(\mathcal{T}) - E_{\Omega} \left(TR \left(\prod_{d=1}^N \mathcal{T}_d(i_d) \right) \right) \right\|_F^2, \quad (8)$$

where $E_{\Omega}(\mathcal{T})$ represents all observed entries, and a set of observation indices is Ω . $TR(\cdot)$ denotes the approximate tensor produced by the latent factor and $[\mathcal{T}_d(i_d)]$ a potential

factor generated by the Bayesian probability TR tensor decomposition.

In particular, the Bayesian probability tensor factorization model adopts a hierarchical probability model to represent the TR factorization. Specifically, several potential factors are introduced into the space induced prior, the complete Bayesian method is adopted, and the appropriate prior is introduced to determine the rank. A remote sensing image can be combined with several noise types, like fringe, Gaussian noise, dashed line, and pulse noise, to simulate real-world scenarios. The Bayesian probability tensor decomposition model obtains the suboptimal rank, and consequently, the rank minimization tensor complement model is introduced:

$$\min_{\chi, \mathcal{T}, \mathcal{E}, \mathcal{P}} \text{Rank}(\chi) + \frac{\alpha}{2} \left\| E_{\Omega}(\mathcal{T}) - E_{\Omega} \left(TR \prod_{d=1}^N \mathcal{T}_d(i_d) \right) \right\|_F^2 + \frac{\beta}{2} \|\mathcal{E}\|_1 + \frac{\gamma}{2} \|\mathcal{P}\|_F^2, \quad (9)$$

where χ is the low-rank tensor after recovery. Therefore, the model can discover the low-rank structure of the data and approximate the recovered tensor. Additionally, an adaptive tensor nuclear norm is designed as follows:

$$\text{Rank}(\chi) = \sum_{i=1}^N \ell_i \|\chi_i\|_{\omega, *}, \quad (10)$$

where $\text{Rank}(\chi)$ is the weighted nuclear norm of χ and ℓ_i is a nonnegative weight related to the Gaussian noise variance in each set of observations. Depending on the importance of the image center point or field, different positions are given different weights. Considering the weight of the edge position of the observation data as one, while the internal position increases proportionally, the center position has the most prominent weight to ensure that each weight is an

integer and reduces the amount of calculation. Embedding this in the model of rank minimization tensor completion,

$$\min_{\mathcal{X}, \mathcal{T}, \mathcal{G}, \mathbf{P}} \sum_{i=1}^N \ell_i \|\chi\|_{\omega, *}, + \frac{\alpha}{2} \left\| E_{\Omega}(\mathcal{T}) - E_{\Omega} \left(TR \prod_{d=1}^N \mathcal{T}_d(i_d) \right) \right\|_F^2 + \frac{\beta}{2} \|\mathcal{G}\|_1 + \frac{\gamma}{2} \|\mathbf{P}\|_F^2. \quad (11)$$

3.2. Optimization. To optimize the proposed solution, we employ the ALM and ALS algorithms shortly introduced as follows.

3.2.1. ALM. Equation (11) is rewritten as an equivalent minimization problem by introducing some auxiliary variables:

$$\min_{\psi, \mathcal{T}, \mathcal{G}, \mathbf{P}} \sum_{n=1}^N \ell_i \|\psi\|_{\omega, *}, + \frac{\alpha}{2} \left\| E_{\Omega}(\mathcal{T}) - E_{\Omega}(\Phi[\mathcal{T}_d(i_d)]) \right\|_F^2 + \frac{\beta}{2} \|\mathcal{G}\|_1 + \frac{\gamma}{2} \|\mathbf{P}\|_F^2 \quad (12)$$

s.t. $\mathcal{Y} = \mathcal{X} + \mathcal{G} + \mathbf{P}$, $\mathcal{X} = \mathcal{L}$, $D_{\omega}(\mathcal{L}) = \psi$,

where $D_{\omega}(\mathbb{R}) = TR\{\prod_{d=1}^N \omega_d \mathcal{T}_d(i_d)\}$ is the weighted N-dimensional difference operator. The problem of equation (12) is transformed into the minimized augmented Lagrangian function of equation (13), under the constraints $\mathcal{X} = C \cdot TR\{\prod_{d=1}^N \mathcal{T}_d(i_d)\}$, where μ is the penalty parameter, and $\Gamma_i(1, 2, 3)$ are the Lagrange multipliers.

Consequently, the augmented Lagrangian function (equation (13)) can perform alternative optimizations to one variable while modifying other variables. Variables involved in the model of equation (8) can be updated in the k -th iteration as follows:

$$\begin{aligned} \mathcal{L}(\mathcal{X}, \mathcal{G}, \mathbf{P}, \mathcal{L}, \psi, \Gamma_1, \Gamma_2, \Gamma_3) = & \sum_{n=1}^N \ell_i \|\psi\|_{\omega, *}, + \langle \Gamma_3, D_{\omega}(\mathcal{L}) - \psi \rangle \\ & + \frac{\alpha}{2} \left\| E_{\Omega}(\mathcal{T}) - E_{\Omega} \left(TR \left(\prod_{d=1}^N \mathcal{T}_d(i_d) \right) \right) \right\|_F^2 \\ & + \langle \Gamma_1, \mathcal{Y} - \mathcal{X} - \frac{1}{2} \mathcal{G} - \frac{1}{2} \mathbf{P} - \frac{1}{2} \left(E_{\Omega}(\mathcal{T}) - E_{\Omega} \left(TR \left(\prod_{d=1}^N \mathcal{T}_d(i_d) \right) \right) \right) \rangle \\ & + \frac{\mu}{2} \left\| \mathcal{Y} - \mathcal{X} - \frac{1}{2} \mathcal{G} - \frac{1}{2} \mathbf{P} - \frac{1}{2} \left(E_{\Omega}(\mathcal{T}) - E_{\Omega} \left(TR \left(\prod_{d=1}^N \mathcal{T}_d(i_d) \right) \right) \right) \right\|_F^2 \\ & + \frac{\mu}{2} \|\mathcal{X} - \mathcal{L}\|_F^2 + \frac{\mu}{2} \|D_{\omega}(\mathcal{L}) - \psi\|_F^2 + \frac{\gamma}{2} \|\mathbf{P}\|_F^2 + \frac{\beta}{2} \|\mathcal{G}\|_1 + \langle \Gamma_2, \mathcal{X} - \mathcal{L} \rangle. \end{aligned} \quad (13)$$

From the augmented Lagrangian function (13), all terms containing \mathcal{X} are extracted as follows:

$$\begin{aligned} \min_{U_i^T U_i=I, \mathcal{X}=C \cdot TR\{\prod_{d=1}^N \mathcal{F}_d(i_d)\}} & \langle \Gamma_1^{(k)}, \mathcal{Y} - \mathcal{X} - \frac{1}{2}\mathcal{G}^{(k)} - \frac{1}{2}\mathbf{P}^{(k)} - \frac{1}{2}\left(E_\Omega(\mathcal{T}) - E_\Omega\left(TR\left(\prod_{d=1}^N \mathcal{F}_d(i_d)\right)\right)\right)^{(k)} \rangle \\ & + \langle \Gamma_2^{(k)}, \mathcal{X} - \mathcal{Z}^{(k)} \rangle + \frac{\mu}{2} \|\mathcal{X} - \mathcal{Z}^{(k)}\|_F^2 \\ & + \frac{\mu}{2} \left\| \mathcal{Y} - \mathcal{X} - \frac{1}{2}\mathcal{G}^{(k)} - \frac{1}{2}\mathbf{P}^{(k)} - \frac{1}{2}\left(E_\Omega(\mathcal{T}) - E_\Omega\left(TR\left(\prod_{d=1}^N \mathcal{F}_d(i_d)\right)\right)\right)^{(k)} \right\|_F^2. \end{aligned} \quad (14)$$

Equation (14) can be transformed into the following equal problem:

$$\min_{U_i^T U_i=I} \mu \left\| C \prod_{i=1}^N U_i - \frac{1}{2} \left(\mathcal{Y} - \frac{1}{2}\mathcal{G}^{(k)} - \frac{1}{2}\mathbf{P}^{(k)} + \mathcal{Z}^{(k)} - \frac{1}{2}\left(E_\Omega(\mathcal{T}) - E_\Omega\left(TR\left(\prod_{d=1}^N \mathcal{F}_d(i_d)\right)\right)\right)^{(k)} + \frac{(\Gamma_1^{(k)} - \Gamma_2^{(k)})}{\mu} \right) \right\|_F^2. \quad (15)$$

Parameters $C^{(k+1)}$ and $U_i^{(k+1)}$ ($i = 1, 2, \dots, N$) can be calculated using the classical Bayesian tensor algorithm. Then,

$$\mathcal{X}^{(k+1)} = C^{(k+1)} \left(TR \left(\prod_{d=1}^N \mathcal{F}_d(i_d) \right) \right)^{(k+1)}. \quad (16)$$

Considering also the augmented Lagrangian function (equation (13)), \mathcal{Z} can be updated to

$$\mathcal{Z}^{(k+1)} = \arg \min_z \langle \Gamma_2^{(k)}, \mathcal{X}^{(k+1)} - \mathcal{Z} \rangle + \langle \Gamma_3^{(k)}, D_\omega(\mathcal{Z}) - \psi^{(k+1)} \rangle + \frac{\mu}{2} \left(\|\mathcal{X}^{(k+1)} - \mathcal{Z}\|_F^2 + \|D_\omega(\mathcal{Z}) - \psi^{(k)}\|_F^2 \right), \quad (17)$$

where D_ω^* represents the adjoint operator of D_ω . The block-circulant structure of the matrix corresponding to the operator $D_\omega^* D_\omega$ can be diagonalized by the FFT matrix, and thus

$$\begin{cases} \mathcal{H}_{\mathcal{Z}} = \mu \mathcal{X}^{(k+1)} + \mu D_\omega^*(\psi^{(k)}) + \Gamma_2^{(k)} - D_\omega^*(\Gamma_3^{(k)}), \\ \mathcal{F}_{\mathcal{Z}} = \sum_{i=1}^N \omega_i^2 |fftfn(D_i)|^2, \\ \mathcal{Z}^{(k+1)} = ittfn\left(\frac{fftfn(H_{\mathcal{Z}})}{\mu + \mu \mathcal{F}_{\mathcal{Z}}}\right), \end{cases} \quad (18)$$

where $fftfn$ and $ittfn$ represent the fast Fourier transform and the inverse fast Fourier transform and $|\cdot|^2$ is the elementwise square. The division is also implemented elementwise.

By extracting from equation (13) all terms containing ψ ,

$$\begin{aligned} \psi^{(k+1)} &= \arg \min_{\psi} \sum_{n=1}^N \ell_n \|\psi\|_{w,*} + \langle \Gamma_3^{(k)}, D_\omega(\mathcal{Z}^{(k+1)}) - \psi \rangle \\ &+ \frac{\mu}{2} \|D_\omega(\mathcal{Z}^{(k+1)}) - \psi\|_F^2 \\ &= \arg \min_{\psi} \tau \|\psi\|_1 + \frac{\mu}{2} \left\| \psi - \left(D_\omega(\mathcal{Z}^{(k+1)}) + \frac{\Gamma_3^{(k)}}{\mu} \right) \right\|_F^2. \end{aligned} \quad (19)$$

Moreover, by introducing a threshold contraction process,

$$\psi^{k+1} = M_{(\tau/\mu)} \left(D_\omega(\mathcal{Z}^{k+1}) + \frac{\Gamma_3^k}{\mu} \right), \quad (20)$$

where $M_{(\tau/\mu)}(\cdot)$ represents the threshold shrinkage processing.

Similarly, by updating \mathcal{Z} ,

$$\begin{aligned}
\mathcal{Z}^{(k+1)} &= \arg \min_{\mathcal{Z}} \|\mathcal{Z}\|_1 + \langle \Gamma_1^{(k)}, \mathcal{Y} - \mathcal{X}^{(k+1)} - \frac{1}{2}\mathcal{Z} - \frac{1}{2}\mathbf{P}^{(k)} - \frac{1}{2}\left(E_{\Omega}(\mathcal{T}) - E_{\Omega}\left(\text{TR}\left(\prod_{d=1}^N \mathcal{T}_d(i_d)\right)\right)\right)^{(k)} \rangle \\
&\quad + \frac{\mu}{2} \left\| \mathcal{Y} - \mathcal{X}^{(k+1)} - \frac{1}{2}\mathcal{Z} - \frac{1}{2}\mathbf{P}^{(k)} \right\|_F^2 + \frac{\mu}{2} \left\| \mathcal{Y} - \frac{1}{2}\left(E_{\Omega}(\mathcal{T}) - E_{\Omega}\left(\text{TR}\left(\prod_{d=1}^N \mathcal{T}_d(i_d)\right)\right)\right)^{(k)} \right\|_F^2 \\
&= \arg \min_{\mathcal{Z}} \|\mathcal{Z}\|_1 + \frac{\mu}{2} \left\| \mathcal{Z} - \left(\mathcal{Y} - \mathcal{X}^{(k+1)} - \frac{1}{2}\mathbf{P}^{(k)} + \frac{\Gamma_1^{(k)}}{\mu} \right) \right\|_F^2 \\
&\quad + \frac{\mu}{2} \left\| \mathcal{Z} - \frac{1}{2}\left(E_{\Omega}(\mathcal{T}) - E_{\Omega}\left(\text{TR}\left(\prod_{d=1}^N \mathcal{T}_d(i_d)\right)\right)\right)^{(k)} \right\|_F^2.
\end{aligned} \tag{21}$$

Using the threshold shrinkage processing introduced above, the following formula is obtained:

$$\mathcal{Z}^{k+1} = M_{\tau} \left(\mathcal{Y} - \mathcal{X}^{k+1} - \mathbf{P}^k + \frac{\Gamma_1^k}{\mu} \right). \tag{22}$$

Then, by updating P through extracting from equation (8) all items related to P,

$$\begin{aligned}
\mathbf{P}^{(k+1)} &= \arg \min_{\mathbf{P}} \|\mathbf{P}\|_F^2 + \frac{\mu}{2} \left\| \mathcal{Y} - \mathcal{X}^{(k+1)} - \frac{1}{2}\mathcal{Z}^{(k+1)} - \frac{1}{2}\mathbf{P} - \frac{1}{2}\left(E_{\Omega}(\mathcal{T}) - E_{\Omega}\left(\text{TR}\left(\prod_{d=1}^N \mathcal{T}_d(i_d)\right)\right)\right)^{(k+1)} \right\|_F^2 \\
&\quad \cdot \langle \Gamma_1^{(k)}, \mathcal{Y} - \mathcal{X}^{(k+1)} - \frac{1}{2}\mathcal{Z}^{(k+1)} - \frac{1}{2}\mathbf{P} - \frac{1}{2}\left(E_{\Omega}(\mathcal{T}) - E_{\Omega}\left(\text{TR}\left(\prod_{d=1}^N \mathcal{T}_d(i_d)\right)\right)\right)^{(k)} \rangle \\
&= \arg \min_{\mathbf{P}} \frac{\mu}{2} \left\| \mathbf{P} - \left(\mathcal{Y} - \frac{1}{2}\left(E_{\Omega}(\mathcal{T}) - E_{\Omega}\left(\text{TR}\left(\prod_{d=1}^N \mathcal{T}_d(i_d)\right)\right)\right)^{(k+1)} \right) \right\|_F^2 + \frac{\mu}{2} \left\| \mathbf{P} - \frac{\Gamma_1^{(k)}}{\mu} \right\|_F^2 \\
&\quad + \frac{\mu}{2} \left\| \mathbf{P} - \left(\mathcal{Y} - \mathcal{X}^{(k+1)} - \frac{1}{2}\mathcal{Z}^{(k+1)} \right) \right\|_F^2.
\end{aligned} \tag{23}$$

Finally, the following formulas are used to update the multipliers based on the ALM algorithm:

$$\begin{cases}
\Gamma_1^{(k+1)} = \Gamma_1^{(k)} + \mu \left(\mathcal{Y} - \mathcal{X}^{(k+1)} - \frac{1}{2}\mathcal{Z}^{(k+1)} - \frac{1}{2}\mathbf{P}^{(k+1)} - \frac{1}{2}\left(E_{\Omega}(\mathcal{T}) - E_{\Omega}\left(\text{TR}\left(\prod_{d=1}^N \mathcal{T}_d(i_d)\right)\right)\right)^{(k+1)} \right), \\
\Gamma_2^{(k+1)} = \Gamma_2^{(k)} + \mu \left(\mathcal{X}^{(k+1)} - \mathcal{Z}^{(k+1)} \right), \\
\Gamma_3^{(k+1)} = \Gamma_3^{(k)} + \mu \left(D_{\omega}(\mathcal{Z}^{(k+1)}) - \psi^{(k+1)} \right).
\end{cases} \tag{24}$$

3.2.2. *ALS*. ALS is adopted to settle the model represented by equation (13) and includes two steps: kernel tensor update and adaptive nuclear norm regularization reconstruction.

Initially, the core tensors are updated sequentially. For $n = 1, \dots, N$, the optimization is

$$(Q)_{(2)}^{(n)} = \arg \min_{\mathcal{Q}_{(2)}^{(n)}} \left\| \Omega_{\langle n \rangle} \circ \left(\chi_{\langle n \rangle} - \mathcal{Q}_{(2)}^{(n)} (\mathcal{Q}_{(2)}^{\#n})^T \right) \right\|_F^2, \quad (25)$$

where \circ is the elementwise multiplication. We adopt $\bar{\mathcal{Q}}$, $\bar{\Omega}$, $\bar{\mathcal{X}}$, and B to denote the matrices $\mathcal{Q}_{(2)}^{(n)}$, $\Omega_{\langle n \rangle}$, $\mathcal{X}_{\langle n \rangle}$, and $(\mathcal{Q}_{(2)}^{\#n})^T$, respectively, where the core $\bar{\mathcal{Q}}$ can be updated on a row-by-row basis. Let $i = 1, \dots, I_n$ be the row index of $\bar{\mathcal{Q}}$, then the update of $\bar{\mathcal{Q}}(i, :)$ is

$$\bar{\mathcal{Q}}(i, :) = \arg \min \left\| \bar{\chi}(i, \bar{\Omega}(i, :)) - \bar{\mathcal{Q}}(i, :) B(:, \bar{\Omega}(i, :)) \right\|_F^2, \quad (26)$$

where $\bar{\mathcal{X}}(i, \bar{\Omega}(i, :))$ is a vector extracted from the i -th row and the observed column provided by $\bar{\Omega}(i, :)$. To resolve equation (26), the least square method is adopted, and the optimal solution is given by the following equation:

$$\bar{\mathcal{Q}}(i, :) = \bar{\chi}(i, \bar{\Omega}(i, :)) (B(:, \bar{\Omega}(i, :)))^T. \quad (27)$$

Subsequently, once all core tensors are updated, the adaptive nuclear norm regularization is adopted to reconstruct tensor χ .

4. Experiments and Discussion

The validity of the proposed method is demonstrated through real-data and simulated experiments.

4.1. *Simulated Experiments*. Remote sensing datasets have already been used in the current literature. Synthetic data were first considered in [36] and involved geographic images acquired by remote sensing instruments, where the synthesized remote sensing image has a size of $145 \times 145 \times 724$ pixels. The next dataset employed is the Hydice urban dataset [37], with a raw image of $307 \times 307 \times 210$ pixels. In this dataset, the image presenting the entire city is affected by water absorption, the atmosphere, deadlines, and stripes.

To comprehensively evaluate the proposed algorithm, it is challenged against four different image reconstruction methods, i.e., using TV and low-rank matrix factorization tensor completion (LRTV) [35], adaptive weighted tensor completion (AWTC) [16], decomposable nonlocal tensor dictionary learning (TDL) [15] and stochastic gradient descent TT completion (TTSGD) [38, 39]. In the subsequent experiments, the parameters of all competitor techniques are manually adjusted based on the default strategy. In addition, the spectral bands of all remote sensing image sets are normalized to $[0, 1]$ and stretched to the raw level after restoration to facilitate visualization and numerical calculation.

4.1.1. *Experiments on the Synthetic Datasets*. To simulate noisy remote sensing images and quantitatively highlight the image recovery capabilities of each method, four noise cases

are artificially created in which the raw synthetic remote sensing data are polluted with several noise types. In any of these cases, the noise intensity of disparate frequency bands is the same. The added noise on each frequency band has the same distributed zero-mean Gaussian noise, while variance and data loss rate vary. Specifically, the values of variance and data loss rate are, for Case 1, 0.06 and 0.2, for Case 2, 0.03 and 0.2, for Case 3, 0.06 and 0.4, and for Case 4, 0.03 and 0.4, respectively.

The average structural similarity index (MSSIM) and the average peak signal-to-noise ratio (MPSNR) are employed to objectively evaluate the repair quality for understanding and analyzing the overall property of the disparate methods:

$$MPSNR = \frac{1}{N} \sum_{i=1}^N psnr_i MPS, \quad (28)$$

$$MSSIM = \frac{1}{N} \sum_{i=1}^N ssim_i,$$

where $psnr_i$ and $ssim_i$ represent the PSNR and SSIM values of the i -th band, respectively.

Figure 2 highlights that all methods have eliminated noise to a certain extent. The proposed method performs best, as it retains the original image's basic structure while effectively removing noise nuisances. Specifically, ALS obtains the best visual effect, while LRTV, AWTC, TTSGD, and TDL successfully reconstruct the missing stripes. However, there are still noticeable streaks in the results, as some details are smoothed out from the restored image utilizing TTSGD and TDL. Although LRTV produces some excessively smooth details, the methods based on regularization perform best in preserving details and reconstructing fringes. Combining the adaptive nuclear norm regularization algorithm and the Bayesian probability tensor modeling technology affords a decisive function in remote sensing image reconstruction.

Table 1 records three indicators of the restoration results of all competitor methods. The two methods proposed have achieved excellent results in all cases, with ALM being better than ALS, mainly because the adaptive nuclear norm contributes more to the ALM architecture. It is also evident that the improvement afforded by the nuclear norm regularization in the ALS method is not adequate. Additionally, when the missed detection rate is reduced, the AWTC method is more effective than TDL and TTSGD. The results of ALS and ALM are better than LRTL, mainly because the performance of the adaptive nuclear norm normalization is better than the traditional TV regularization. Considering the quantitative evaluation utilizing different indicators on the simulated restoration results, AWTC has the best MPSNR and SLM evaluation, but ALM manages the best MSSIM value.

4.1.2. *Experiments on Hydice Urban Datasets*. In these experiments, we select the tensor of the decomposed by Bayesian probability RANK [25, 10, 25], [25, 10, 15], [15, 10, 25], and [15, 10, 15], and the adaptive nuclear norm

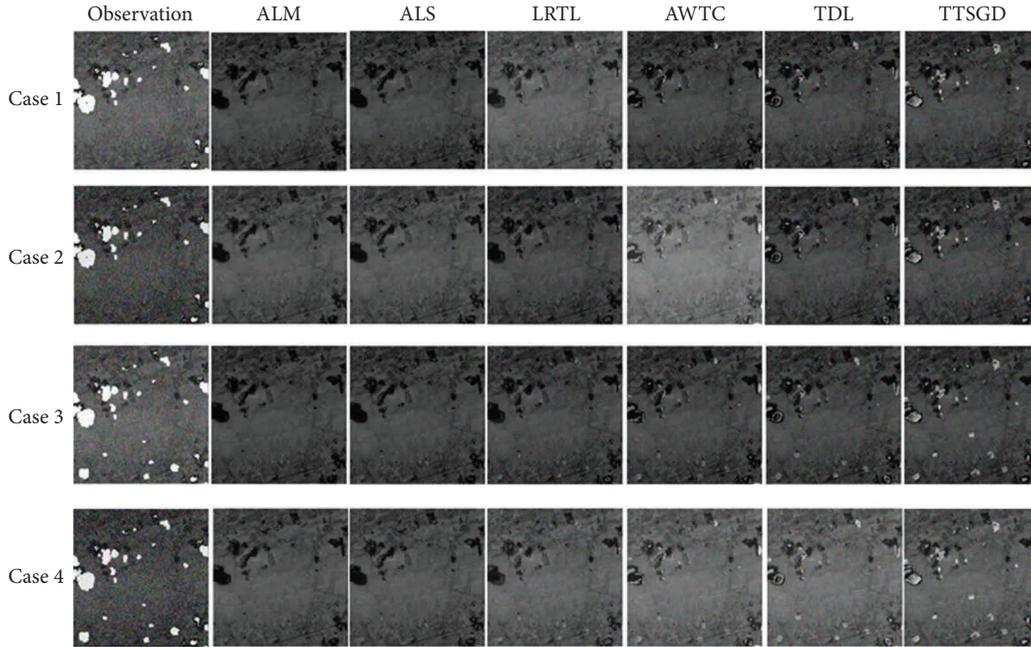


FIGURE 2: Denoised results by all the compared methods.

TABLE 1: Quantitative evaluation of disparate methods on the emulated data in disparate conditions.

Noise case	Evaluation index	Noise	ALS	ALM	LRTL	AWTC	TDL	TTSGD
Case 1	MPSNR (dB)	28.35	46.32	39.48	34.59	36.15	35.22	34.68
	MSSIM	0.3485	0.9856	0.9712	0.931	0.9564	0.9416	0.8934
	ERGAS	245.21	22.10	23.26	26.25	30.25	31.26	40.29
Case 2	MPSNR (dB)	22.65	44.26	42.15	39.25	34.21	36.21	32.15
	MSSIM	0.3324	0.9901	0.9851	0.9726	0.8901	0.9015	0.8541
	ERGAS	268.25	25.21	26.29	50.26	101.27	104.58	52.16
Case 3	MPSNR (dB)	15.16	43.25	43.26	37.25	29.54	25.62	30.21
	MSSIM	0.1815	0.9908	0.9874	0.9526	0.9154	0.8564	0.8705
	ERGAS	492.56	25.26	26.25	37.85	108.36	91.25	58.17
Case 4	MPSNR (dB)	17.25	42.65	41.10	39.32	27.24	28.11	33.58
	MSSIM	0.1814	0.9915	0.9899	0.9846	0.8436	0.8245	0.8801
	ERGAS	531.21	23.25	23.98	56.44	123.25	129.00	62.15

regularize parameter is defined as $\lambda = 10^{-6}$ and $\lambda = 10^{-5}$ for data missing percentages of 25% and 50%. Also, several methods are employed, such as LRTL, AWTC, TDL, and TTSGD.

Table 2 lists the quantitative evaluation results of a random loss of remote sensing images under different completion methods. The table highlights that ALM and ALS have considerable advantages over other comparison methods. The experimental results reveal that under the condition of missing information, the adaptive nuclear norm regularization can effectively enhance the property of low-rank tensor models. The AWTC method achieves better results with a lower miss rate, while the TDL method operates oppositely.

Figure 3 illustrates the composite figure after applying remote sensing image reconstruction when 50% of the data is missing. The figure shows that TDL and TTSGD do not recover some of the missing data, and LRTL and AWTC produce ambiguous results. However, ALM and ALS

produce better results. In summary, ALM and ALS perform best for remote sensing reconstruction, demonstrating the advantages of the Bayesian probability tensor decomposition combined with an adaptive nuclear norm adjustment.

4.2. Real-Data Experiments. This section involves experiments utilizing remote sensing camera data from a lake area under actual conditions, aiming to verify the reconstruction capability of the proposed method in real-world imagery. The simulated remote sensing image acquisition device of Figure 4 is used to obtain the corresponding image data, which are further processed to simulate the remote sensing image obtained by the remote sensing device in reality [40]. The gray value of each image band is normalized to $[0, 1]$ before restoration. The tested image is of size of $300 \times 300 \times 6 \times 6$, that is, the space size is 300×300 , the number of spectra is 6, and the time node is 6. We perform cloud pollution on the experimental images to simulate the

TABLE 2: Quantitative evaluation of disparate methods on the emulated data in disparate conditions.

Noise case	Evaluation index	Noise	ALS	ALM	LRTL	AWTC	TDL	TTSGD
Case 1	MPSNR (dB)	25.67	50.29	48.26	40.26	29.26	44.26	34.39
	MSSIM	0.2935	0.9526	0.9478	0.8963	0.9435	0.9528	0.8816
	ERGAS	195.25	19.26	20.35	24.59	27.36	28.26	44.28
Case 2	MPSNR (dB)	17.56	39.96	39.56	32.56	34.75	33.24	30.19
	MSSIM	0.2486	0.9825	0.9545	0.9878	0.8548	0.9245	0.8274
	ERGAS	27.22	19.48	19.99	48.25	110.24	108.55	79.54
Case 3	MPSNR (dB)	12.01	48.65	37.55	34.26	29.08	26.55	27.25
	MSSIM	0.1938	0.9810	0.9744	0.9556	0.9121	0.8377	0.8648
	ERGAS	535.53	18.11	18.22	31.24	100.14	98.24	40.22
Case 4	MPSNR (dB)	11.99	39.22	38.65	37.14	28.26	29.33	31.02
	MSSIM	0.1639	0.9715	0.9156	0.9815	0.8021	0.7748	0.8125
	ERGAS	519.26	20.55	22.10	57.15	118.26	131.05	55.28

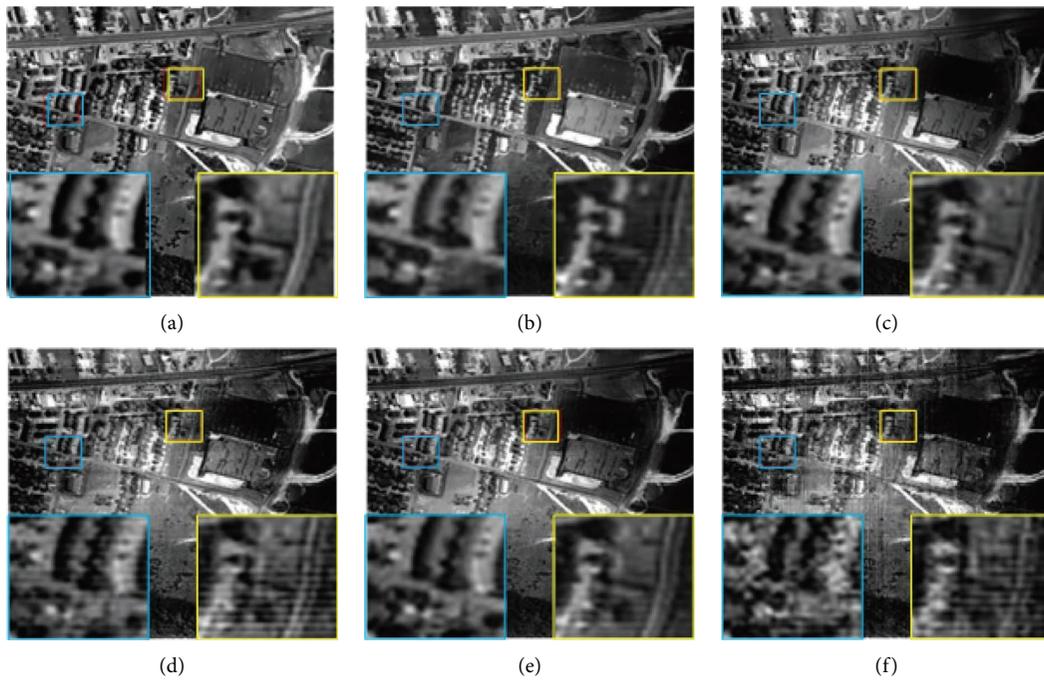


FIGURE 3: Denoised consequences by all the compared methods: (a) ALS, (b) ALM, (c) LRTL, (d) AWTC, (e) TDL, and (f) TTSGD.

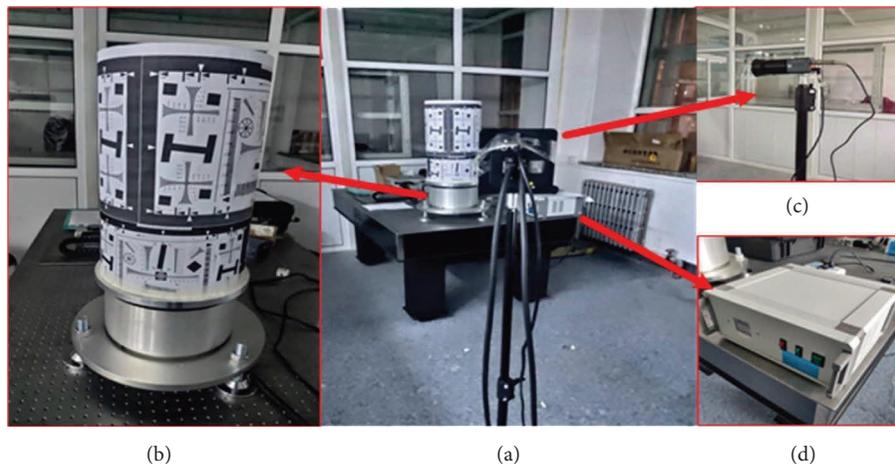


FIGURE 4: Remote sensing image shooting test bench: (a) Ground picture acquisition device, (b) ground simulator, (c) camera lens, and (d) speed control device.

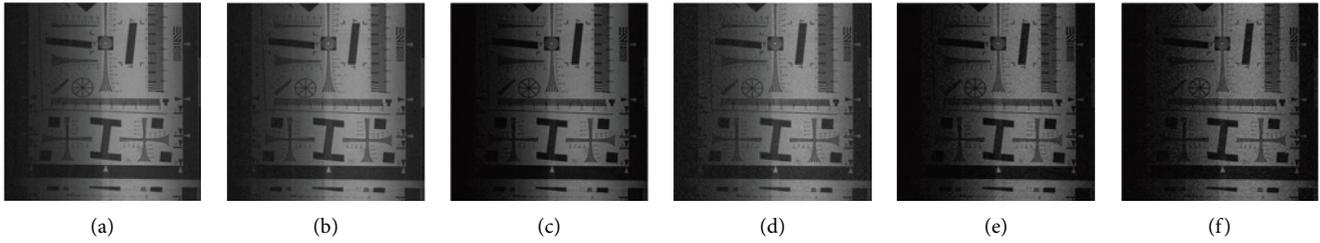


FIGURE 5: Denoised consequences by all the compared methods: (a) ALS, (b) ALM, (c) LRTL, (d) AWTC, (e) TDL, and (f) TTSGD.

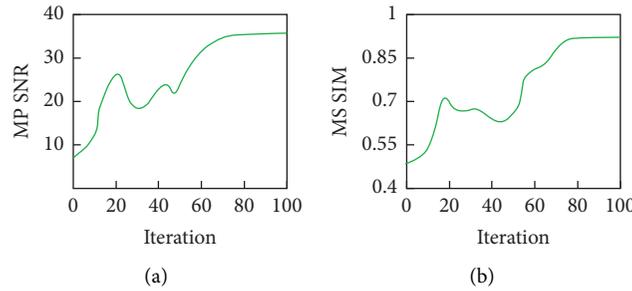


FIGURE 6: MPSNR and MSSIM value versus the iteration number: (a) change in the MPSNR value and (b) change in the MSSIM value.

effects of mixed noise during the remote sensing imagery acquisition.

This trial considers only the recovered images of the models for visual comparison as all competitor methods have similar performance in these actual data experiments. Figure 5 presents the remote sensing images of the time nodes after reconstruction utilizing various methods. The method proposed in this article manages better results compared to the competitor techniques, with ALS achieving the best visual effects. Despite LRTV, TTSGD, AWTC, and TDL can reconstruct the lacking stripes, some noticeable streaks remain. Considering TTSGD and TDL, some details are smoothed out from the enlarged image, and although LRTV produces some overly smooth details, regularization-based methods perform best in maintaining detail and reconstructing fringes.

4.3. Discussion. This subsection considers the computational burden, initialization, convergence, and parameter selection of ALM. Given that ALS and ALM have the same performance, analyzing the latter is neglected to optimize readability and maintain the paper at a reasonable length.

Convergence: Figure 6 presents the interplay between MPSNR as the iteration number in the disparate emulated experiments increases. As the number of iterations increases, the ALM converges for an increasing MPSNR, which at some point stabilizes. Hence, in all trials, 100 iterations are considered.

Sensitivity analysis of Parameter λ : Figure 7 highlights the relationship between MPSNR and parameter λ for real-data experiments. From the latter figure, when parameter λ starts to increase, the value of MPSNR increases and reaches

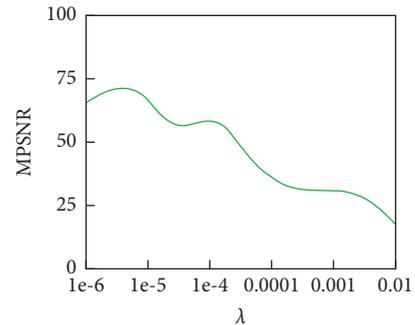


FIGURE 7: Changes in MPSNR values for ALM with an increase in the regularization parameter λ under disparate simulations.

a peak value at $5e-6$, indicating the efficiency of nuclear norm regularization. Nevertheless, the MPSNR value drops due to excessive regularization smoothness after the peak point.

The validity of the Rank Constraint: the Bayesian probability tensor factorization is employed to encode the low-rank prior during the ALM optimization process. Figure 8 shows the MSSIM and MPSNR values of the adaptive nuclear norm solver with disparate rank-constrained values. The results indicate that the values of MSSIM and MPSNR have a specific relationship with the estimated rank value, i.e., both values initially increase and afterward decrease because the estimated rank value increases.

Time Cost: the time cost of disparate methods corresponding to disparate simulation methods is reported in Table 3, demonstrating that the proposed algorithm (ALM) has a lower time complexity than competitor methods.

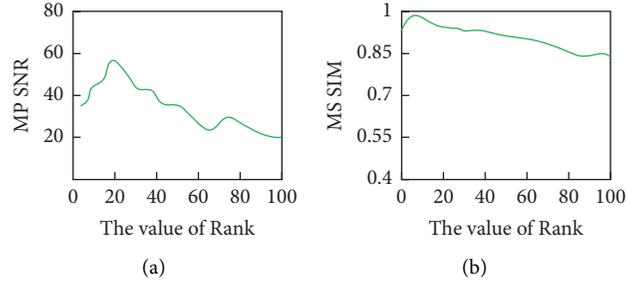


FIGURE 8: Sensitivity analysis of rank: (a) change in the MPSNR value and (b) change in the MSSIM value.

TABLE 3: Time cost (seconds) of disparate methods with disparate emulated data sets.

Data	ALM	LRTL	AWTC	TDL	TTSGD
The synthetic datasets	68.53	97.56	107.94	148.23	134.29
The Hydice urban datasets	64.59	93.68	101.29	159.58	134.59
Real-data	59.68	88.46	92.33	122.94	105.68

5. Conclusion

This article proposes a new tensor completion method, where the Bayesian probability tensor decomposition method is used for remote sensing image reconstruction to detect temporal, spectral, and spatial information. In our scheme, we adopt the adaptive kernel-mode regularization and Bayesian probability tensor decomposition for the spatial smoothness of remote sensing images. We demonstrate that the proposed method is more appealing than commonly used methods in visual comparison and quantitative evaluation through a series of simulations and real-data experiments. The conclusions derived from our research are as follows:

- (i) The suggested method is significantly more robust than the tensor completion method based on traditional tensor transformation. This is proven through our experiments utilizing real or synthetic data. The adaptive nuclear norm uses the gradient similarity between all remote sensing image bands to characterize the smooth structure further, thereby improving the performance and efficiency of the suggested method.
- (ii) The proposed combination of adaptive nuclear norm regularization and tensor decomposition manages appealing stability and can solve the inherent problems of traditional tensor reconstruction methods, such as high-rank selection sensitivity and low computational efficiency.

Future research shall enhance Bayesian probability tensor decomposition efficiency and extend our method to utilize the gradient similarity between all remote sensing image bands to characterize the smooth structure further.

Abbreviations

$\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}$:	Higher-order tensor
\mathcal{T}_d :	Latent tensor
$\mathfrak{N}(\cdot)$:	Gaussian distribution
ς :	Accuracy
φ_i, ν_i :	Hyperparameters
η_0 :	Degree of freedom of the Wishart distribution
\mathcal{Y} :	Vertical image
\mathcal{E} :	Mixed noise items
\mathcal{G} :	Gaussian noise term
\mathcal{P} :	Sparse noise term
$TR(\cdot)$:	Approximate tensor
$[\mathcal{T}_d(i_d)]$:	Potential factor
χ :	Low-rank tensor
$\text{Rank}(\chi)$:	Weighted nuclear norm
$D_\omega(\mathbb{R})$:	Weighted N-dimensional difference operator
μ :	Penalty parameter
$\Gamma_i(1, 2, 3)$:	Lagrange multipliers
$ \cdot ^2$:	Elementwise square
$M_{(\tau/\mu)}(\cdot)$:	Threshold shrinkage processing
\circ :	Elementwise multiplication.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Authors' Contributions

LJL investigated the study and reviewed and edited the article. YY and SS conceptualized and visualized the study, and reviewed and edited the manuscript. PL and XHQ visualized and investigated the study. CL supervised the study. SL proposed the methodology, was responsible for data curation, performed formal analysis, and wrote the original draft.

Acknowledgments

This research was financially supported by Changchun Science and Technology Bureau Key Research and Development Projects (Grant no. 21ZGG08).

References

- [1] F. Vuolo, W. T. Ng, and C. Atzberger, "Smoothing and gap-filling of high resolution multi-spectral time series: example of landsat data," *International Journal of Applied Earth Observation and Geoinformation*, vol. 57, pp. 202–213, 2017.
- [2] Q. Zhang, Q. Yuan, C. Zeng, X. Li, and Y. Wei, "Missing data reconstruction in remote sensing image with a unified spatial-temporal-spectral deep convolutional neural network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 8, pp. 4274–4288, 2018.
- [3] L. Zhuang and J. M. Bioucas-Dias, "Fast hyperspectral image denoising and inpainting based on low-rank and sparse representations," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 3, pp. 730–742, 2018.
- [4] Y. Wang, J. Peng, Q. Zhao, Y. Leung, X. L. Zhao, and D. Meng, "Hyperspectral image restoration via total variation regularized low-rank tensor decomposition," *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 4, pp. 1227–1243, 2018.
- [5] H. Lu, J. Wei, L. Wang et al., "Reference information based remote sensing image reconstruction with generalized non-convex low-rank approximation," *Remote Sensing*, vol. 8, no. 6, p. 499, 2016.
- [6] J. M. Bioucas-Dias and M. A. T. Figueiredo, "A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration," *IEEE Transactions on Image Processing*, vol. 16, no. 12, pp. 2992–3004, 2007.
- [7] H. Li and C. Y. Suen, "A Novel Non-local Means Image Denoising Method Based on Grey Theory," *Pattern Recognition*, vol. 49, pp. 237–248, 2016.
- [8] X. Li, H. Shen, L. Zhang, H. Zhang, Q. Yuan, and G. Yang, "Recovering quantitative remote sensing products contaminated by thick clouds and shadows using multitemporal dictionary learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 11, pp. 7086–7098, 2014.
- [9] L. Lorenzi, F. Melgani, and G. Mercier, "Multiresolution inpainting for reconstruction of missing data in VHR images," in *Proceedings of the 2011 IEEE International Geoscience and Remote Sensing Symposium*, vol. 8, no. 5, pp. 531–534, Vancouver, BC, Canada, July 2011.
- [10] J. Xue, Y. Zhao, W. Liao, and J. C. W. Chan, "Nonlocal low-rank regularized tensor decomposition for hyperspectral image denoising," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 7, pp. 5174–5189, 2019.
- [11] H. Zeng, X. Xie, H. Cui, H. Yin, and J. Ning, "Hyperspectral image restoration via global L_{1-2} spatial-spectral total variation regularized local low-rank tensor recovery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 4, pp. 3309–3325, 2021.
- [12] R. Dian and S. Li, "Hyperspectral image super-resolution via subspace-based low tensor multi-rank regularization," *IEEE Transactions on Image Processing*, vol. 28, no. 10, pp. 5135–5146, 2019.
- [13] J. Yang, L. Xiao, Y. Q. Zhao, and J. C. W. Chan, "Variational regularization network with attentive deep prior for hyperspectral-multispectral image fusion," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–17, 2022.
- [14] Y. Q. Zhao and J. Yang, "Hyperspectral image denoising via sparse representation and low-rank constraint," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 1, pp. 296–308, 2015.
- [15] R. Hao and Z. Su, "A patch-based low-rank tensor approximation model for multiframe image denoising," *Journal of Computational and Applied Mathematics*, vol. 329, pp. 125–133, 2018.
- [16] M. K. P. Ng, Q. Yuan, L. Yan, and J. Sun, "An adaptive weighted tensor completion method for the recovery of remote sensing images with missing data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 6, pp. 3367–3381, 2017.
- [17] Y. Chang, L. Yan, H. Fang, and C. Luo, "Anisotropic spectral-spatial total variation model for multispectral remote sensing image destriping," *IEEE Transactions on Image Processing*, vol. 24, no. 6, pp. 1852–1866, 2015.
- [18] J. Xue, Y. Zhao, Y. Bu, J. C. W. Chan, and S. G. Kong, "When laplacian scale mixture meets three-layer transform: a parametric tensor sparsity for tensor completion," *IEEE Transactions on Cybernetics*, pp. 1–15, 2022.
- [19] J. Xue, Y. Zhao, S. Huang, W. Liao, J. C. W. Chan, and S. G. Kong, "Multilayer sparsity-based tensor decomposition for low-rank tensor completion," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 11, pp. 6916–6930, 2022.
- [20] J. Xue, Y. Zhao, W. Liao, J. C. W. Chan, and S. G. Kong, "Enhanced sparsity prior model for low-rank tensor completion," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 11, pp. 4567–4581, 2020.
- [21] H. Shen, X. Li, L. Zhang, D. Tao, and C. Zeng, "Compressed sensing-based inpainting of aqua moderate resolution imaging spectroradiometer band 6 using adaptive spectrum-weighted sparse bayesian dictionary learning," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 2, pp. 894–906, 2014.
- [22] C. Zeng, H. Shen, and L. Zhang, "Recovering missing pixels for landsat ETM+ SLC-off imagery using multi-temporal regression analysis and a regularization method," *Remote Sensing of Environment*, vol. 131, pp. 182–194, 2013.
- [23] Z. Zhu, S. Wang, and C. E. Woodcock, "Improvement and expansion of the fmask algorithm: cloud, cloud shadow, and snow detection for landsats 4-7, 8, and sentinel 2 images," *Remote Sensing of Environment*, vol. 159, pp. 269–277, 2015.
- [24] P. Rai, Y. Wang, S. Guo, G. Chen, D. Dunson, and L. Carin, "Scalable bayesian low-rank decomposition of incomplete multiway tensors," in *Proceedings of the 31st International Conference on Machine Learning, ICML 2014*, vol. 5, pp. 3810–3820, Beijing, China, 2014.
- [25] Q. Zhao, L. Zhang, and A. Cichocki, "Bayesian CP factorization of incomplete tensors with automatic rank determination," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 9, pp. 1751–1763, 2015.
- [26] R. Salakhutdinov and A. Mnih, "Bayesian probabilistic matrix factorization using Markov chain Monte Carlo," *Proceedings of the 25th International Conference on Machine Learning, ACM, Helsinki, Finland*, 2008.
- [27] S. D. Babacan, M. Luessi, R. Molina, and A. K. Katsaggelos, "Sparse bayesian methods for low-rank matrix estimation," *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 3964–3977, 2012.

- [28] N. D. Lawrence and R. Urtasun, "Non-linear matrix factorization with Gaussian processes," *Proceedings of the 26th Annual International Conference on Machine Learning - ICML '09*, ACM, vol. 382, Montreal, QC, Canada, 2009.
- [29] W. Chu and Z. Ghahramani, "Probabilistic models for incomplete multi-dimensional arrays," *Journal of Machine Learning Research*, vol. 5, pp. 89–96, 2006.
- [30] Z. Xu, F. Yan, and Y. Qi, "Infinite tucker decomposition: nonparametric bayesian models for multiway data analysis," *Proceedings of the 29th International Conference on Machine Learning, ICML 2012*, Vol. 2, ACM, Edinburgh, Scotland, 2012.
- [31] K. Hayashi, T. Takenouchi, T. Shibata et al., "Exponential family tensor factorization for missing-values prediction and anomaly detection," in *Proceedings of the 2010 IEEE International Conference on Data Mining*, p. 216, Sydney, NSW, Australia, December 2010.
- [32] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009.
- [33] R. Salakhutdinov and A. Mnih, "Probabilistic matrix factorization," *Advances in Neural Information Processing Systems 20 - Proceedings of the 2007 Conference*, Curran Associates, Inc, Red Hook, NY, USA, 2009.
- [34] H. Zhang, W. He, L. Zhang, H. Shen, and Q. Yuan, "Hyperspectral image restoration using low-rank matrix recovery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 8, pp. 4729–4743, 2014.
- [35] W. He, H. Zhang, L. Zhang, and H. Shen, "Total-variation-regularized low-rank matrix factorization for hyperspectral image restoration," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 1, pp. 178–188, 2016.
- [36] D. Landgrebe and L. Biehl, *A Freeware Multispectral Image Data Analysis System*, Springer International Publishing, Manhattan, NY, USA, 2015.
- [37] S. Goslee, *Radiometric and Topographic Correction of Satellite Imagery*, 2015.
- [38] Y. Peng, D. Meng, Z. Xu, C. Gao, Y. Yang, and B. Zhang, "Decomposable Nonlocal Tensor Dictionary Learning for Multispectral Image Denoising," in *Proceedings of the IEEE Computer Society Conference On Computer Vision And Pattern Recognition*, pp. 2949–2956, Columbus, OH, USA, June 2014.
- [39] L. Yuan, Q. Zhao, L. Gui, and J. Cao, "High-order tensor completion via gradient-based optimization under tensor train format," *Signal Processing: Image Communication*, vol. 73, pp. 53–61, 2019.
- [40] S. Lei, M. Lu, J. Lin, X. Zhou, and X. Yang, "Remote sensing image denoising based on improved semi-soft threshold," *Signal, Image and Video Processing*, vol. 15, no. 1, pp. 73–81, 2021.