Hyperspectral Super-Resolution Via Joint Regularization of Low-Rank Tensor Decomposition

Meng Cao 1, Wenxing Bao 1,2,*, and Kewen Qu 1,2

1 School of Computer Science and Engineering, North Minzu University, Yinchuan 750021, China; 20197237@stu.nun.edu.cn (M.C.); qukewen@nun.edu.cn (K.Q.)
2 The Key Laboratory of Images and Graphics Intelligent Processing of State Ethnic Affairs Commission: IGIPLab, North Minzu University, Yinchuan 750021, China
* Correspondence: baowenxing@nun.edu.cn

Abstract: The hyperspectral image super-resolution (HSI-SR) problem aims at reconstructing the high resolution spatial–spectral information of the scene by fusing low-resolution hyperspectral images (LR-HSI) and the corresponding high-resolution multispectral image (HR-MSI). In order to effectively preserve the spatial and spectral structure of hyperspectral images, a new joint regularized low-rank tensor decomposition method (JRLTD) is proposed for HSI-SR. This model alleviates the problem that the traditional HSI-SR method, based on tensor decomposition, fails to adequately take into account the manifold structure of high-dimensional HR-HSI and is sensitive to outliers and noise. The model first operates on the hyperspectral data using the classical Tucker decomposition to transform the hyperspectral data into the form of a three-mode dictionary multiplied by the core tensor, after which the graph regularization and unidirectional total variational (TV) regularization are introduced to constrain the three-mode dictionary. In addition, we impose the $l_1$-norm on core tensor to characterize the sparsity. While effectively preserving the spatial and spectral structures in the fused hyperspectral images, the presence of anomalous noise values in the images is reduced. In this paper, the hyperspectral image super-resolution problem is transformed into a joint regularization optimization problem based on tensor decomposition and solved by a hybrid framework between the alternating direction multiplier method (ADMM) and the proximal alternate optimization (PAO) algorithm. Experimental results conducted on two benchmark datasets and one real dataset show that JRLTD shows superior performance over state-of-the-art hyperspectral super-resolution algorithms.

Keywords: hyperspectral image super-resolution; fusion; tucker decomposition; joint regularization

1. Introduction

Hyperspectral images are obtained through hyperspectral sensors mounted on different platforms, which simultaneously image the target area in tens or even hundreds of consecutive and relatively narrow wavelength bands in multiple regions of the electromagnetic spectrum, such as the ultraviolet, visible, near-infrared and infrared, so it obtains rich spectral information along with surface image information. In other words, hyperspectral imagery combines image information and spectral information of the target area in one. The image information reflects the external characteristics such as size and shape of the sample, while the spectral information reflects the physical structure and chemical differences within the sample. In the field of hyperspectral image processing and applications, fusion [1] is an important element. Furthermore, the problem of hyperspectral image super-resolution (HSI-SR) is to fuse the hyperspectral image (LR-HSI) with rich spectral information and poor spatial resolution with a multispectral image (HR-MSI) with less spectral information but higher spatial resolution to obtain a high-resolution hyperspectral image (HR-HSI). It can usually be divided into two categories: hyper-sharpening and MSI-HSI fusion.
The earliest work on hyper-sharpening was an extension of pansharpening [2,3]. Pansharpening is a fusion method that takes a high-resolution panchromatic (HR-PAN) image and a corresponding low-resolution multispectral (LR-MSI) image to create a high-resolution multispectral image (HR-MSI). Meng et al. [4] first classified the existing pansharpening methods into component replacement (CS), multi-resolution analysis (MRA), and variational optimization (VO-based methods).

The steps of the CS [5] based methods are to first project the MSI bands into a new space based spectral transform, after which the components representing the spatial information are replaced with HR-PAN images, and finally the fused images are obtained by back-projection. Representative methods include principal component analysis (PCA) [6], Gram Schmidt (GS) [7], etc. The multi-resolution analysis (MRA) [8] method is a widely used method in pan-sharpening which is usually based on discrete wavelet transform (DWT) [9]. The basic idea is to perform DWT on MS and Pan images, then retain the approximate coefficients in MSI and replace the spatial detail coefficients with the approximate coefficients of PAN images to obtain the fused images. Representative algorithms are smoothing filter-based intensity modulation (SFIM) [10], generalized Laplace pyramid (GLP) [11], etc. VO-based [12] methods are an important class of pan-sharpening methods. Since the main fusion processes of regularization-based methods [13–17], Bayesian-based methods [18–20], model-based optimization (MBO) [21–23] methods and sparse reconstruction (SR) [24–26] based methods are all based on or transformed into an optimization of a variational model, they can be generalized to variational optimization (VO) based methods. In other words, the main process of such pan-sharpening methods is usually based on or transformed into an optimization of a variational model. A comprehensive review of VO methods based on the concept of super-resolution was first presented by Garzelli [27]. As the availability of HS imaging systems increased, pan-sharpening was extended to HSI-SR by fusing HSI with PANs, which is referred to as hyper-sharpening [28]. In addition, some hyper-sharpening methods have evolved from MSI-HSI fusion methods [13,14,29]. In this case, MSI consists of only a single band, so MSI can be simplified to PAN images [28], and a more detailed comparison of hyper-sharpening methods can be found in [28].

In recent years, several methods have been proposed to realize the hyper-sharpening process of hyperspectral data, such as: linear spectral unmixing (LSU)-based techniques [30,31], nonnegative matrix decomposition-based methods [29,32–37], tensor-based methods [38–41], and deep learning-based methods to improve the spatial resolution of hyperspectral data by using multispectral images. The LSU technique [30] is essentially a problem of decomposing remote sensing data into endmembers and their corresponding abundances. Song et al. [31] proposed a fast unmixing-based sharpening method, which uses unconstrained least squares algorithm to solve the endmember and abundance matrices. The innovation of the method is to apply the procedure to sub-images rather than to the whole data. Yokoya et al. [29] proposed a nonnegative matrix factorization (NMF)-based hyper-sharpening algorithm called coupled NMF (CNMF) by alternately unmixing low-resolution HS data and high-resolution MS data. In CNMF, the endmember matrix and the abundance matrix are estimated using the alternating spectral decomposition of NMF under the constraints of the observation model. However, the results of CNMF may not always be satisfactory; firstly, the solution of NMF is usually non-unique, and secondly, its solution process is very time-consuming because it needs to continuously alternate the application of NMF unmixing to low spatial resolution hyperspectral and high spatial resolution multispectral data, which yields a hyperspectral endmember and a high spatial resolution abundance matrix. Later, by combining these two matrices, fused data with high spatial and spectral resolution can be obtained. An HSI-SR method based on the sparse matrix decomposition technique was proposed in [33], which decomposes the HSI into a basis matrix and a sparse coefficient matrix. Then the HR-HSI was reconstructed using the spectral basis obtained from LR-HSI and the sparse coefficient matrix estimated by HR-MSI. Other NMF-based sharpening algorithms include spectral constraint NMF [34], sparse constraint NMF [35], joint-criterion NMF-based (JNMF) hyper-sharpening algorithm [36],
etc. Specifically, some of the NMF-based methods can also be applied to the fusion process of HS and PAN images, e.g., [34,35]. Furthermore, in order to obtain better fusion results, the work of [37] exploited both the sparsity and non-negativity constraints of HR-HSI and achieved good performance.

Although many methods based on matrix decomposition under different constraints have been proposed by researchers and yielded better performance, these methods based on matrix decomposition require the three-dimensional remote sensing data to be expanded into the form of a two-dimensional matrix, which makes it difficult for the algorithms to take full advantage of the spatial spectral correlation of HSI. HSI-SR method based on tensor decomposition has become a hot topic in MSI-HSI fusion research because of its excellent performance. The main idea of its fusion is to treat HR-HSI as a three-dimensional tensor and to redefine the HSI-SR problem as the estimation of the core tensor and dictionary in three modes. Dian et al. [38] first proposed a non-local sparse tensor factorization method for the HSI-SR problem (called NLSTF), which treats hyperspectral data as a tensor of three modes and combines the non-local similarity prior of hyperspectral images to nonlocally cluster MSI images, and although this method produced good results, LR-HSI was only used for learning the spectral dictionary and not for core tensor estimation. Li et al. [39] proposed the coupled sparse tensor factorization (CSTF) method, which directly decomposes the target HR-HSI using Tucker decomposition and then promotes the sparsity of the core tensor using the high spatial spectral correlation in the target HR-HSI. In order to effectively preserve the spatial spectral structure in LR-HSI and HR-MSI, Zhang et al. [40] proposed a new low-resolution HS (LRHS) and high-resolution MS (HRMS) image fusion method based on spatial–spectral-graph-regularized low-rank tensor decomposition (SS-GLRTD). This method redefines the fusion problem as a low-rank tensor decomposition model by considering LR-HSI as the sum of HR-HSI and sparse difference images. Then, the spatial spectral low-rank features of HR-HSI images were explored using the Tucker decomposition method. Finally, the HR-MSI and LR-HSI images were used to construct spatial and spectral graphs, and regularization constraints were applied to the low-rank tensor decomposition model. Xu et al. [41] proposed a new HSI-SR method based on a unidirectional total variational (TV) approach. The method has decomposed the target HR-HSI into a sparse core tensor multiplied by a three-mode dictionary matrix using Tucker decomposition, and then applied the $l_1$-norm to the core tensor to represent the sparsity of the target HR-HSI and the unidirectional TV three dictionaries to characterize the piecewise smoothness of the target HR-HSI. In addition, tensor ring-based super-resolution algorithms for hyperspectral images have recently attracted the attention of research scholars. He et al. [42,43] proposed a HSI-SR method based on a constrained tensor ring model, which decomposes the higher-order tensor into a series of three-dimensional tensors. Xu et al. [44] proposed a super-resolution fusion of LR-HSI and HR-MSI using a higher-order tensor ring method, which preserves the spectral information and core tensor in a tensor ring to reconstruct high-resolution hyperspectral images.

Deep learning has received increasing attention in the field of HSI-SR with its superior learning performance and high speed. However, deep learning-based methods usually require a large number of samples to train the neural network to obtain the parameters of the network.

The Tucker tensor decomposition is a valid multilinear representation for high-dimensional tensor data, but it fails to take the manifold structures of high-dimensional HR-HSI into account. Furthermore, the graph regularization can perfectly preserve local information of high-dimensional data and achieve good performances in many fusion tasks. Moreover, the existing tensor decomposition-based methods are sensitive to outliers and noise, there is still much room for improvement. We propose a new method based on joint regularization low-rank tensor decomposition (JRLTD) in this paper to solve the HSI-SR problem from the tensor perspective. The model operates on hyperspectral data using the classical Tucker decomposition and introduces graph regularization and the unidirectional total variation regularization (TV), which effectively preserves the spatial and spectral
structures in the fused hyperspectral images while reducing the presence of anomalous noise values in the images, thus solving the HSI-SR problem. The main contributions of the paper are summarized as follows.

(1) In the process of recovering high-resolution hyperspectral images (HR-HSI), joint regularization is considered to operate on the three-mode dictionary. The graph regularization can make full use of the manifold structure in LR-HSI and HR-MSI, while the unidirectional total variational regularization fully considers the segmental smoothness of the target image, and the combination of the two can effectively preserve the spatial structure information and the spectral structure information of HR-HSI.

(2) Based on the unidirectional total variational regularization, the $l_{2,1}$-norm is used. The $l_{2,1}$-norm is not only sparse for the sum of the absolute values of the matrix elements, but also requires row sparsity.

(3) During the experiments, not only the standard dataset of hyperspectral fusion is adopted, but also the dataset about the local Ningxia is used, which makes the algorithm more widely suitable and the performance more convincing.

The remainder of this paper is organized as follows. Section 2 presents theoretical model and related work. Section 3 describes the solution to the optimization model. Section 4 describes our experimental results and evaluates the algorithm. Conclusions and future research directions are presented in Section 5.

2. Related Works

We introduce the definition and representation of the tensor, discuss the basic problems of image fusion, and introduce the concept of joint regularization.

2.1. Tensor Description

In this paper, the capital flower font $\mathcal{T} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ denotes the $N$th order tensor, and each element in the tensor can be obtained by fixing the subscript: $T_{i_1,i_2,\cdots,i_N} \in \mathbb{R}$. In addition, to distinguish the tensor representation, this paper uses the capital letter to denote the matrix, e.g., $X \in \mathbb{R}^{I_1 \times I_2}$; the lower case letter denotes the vector, e.g., $x \in \mathbb{R}^I$. Tensor vectorization is the process of transforming a tensor into a vector. For example, a tensor $\mathcal{T} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ of order $N$ is tensorized to a vector $\mathcal{T} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, which can be expressed as $\tau = \text{vec}(\mathcal{T})$. The elemental correspondence between them is as follows:

$$T_{i_1,i_2,\cdots,i_N} = \tau_{1+i_1+(i_2-1)+\cdots+i_1+i_2+\cdots+i_{N-1}+(i_N-1)} \quad (1)$$

An $n$-mode expansion matrix is defined by arranging the $n$-mode fibers of a tensor as columns of a matrix, e.g., $T_{(n)} = \text{unfold}(\mathcal{T}) \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N \times N}$. Conversely, the inverse of the expansion can be defined as $\mathcal{T} = \text{fold}(T_{(n)})$. The $n$-mode product of a tensor $\mathcal{T} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ and a matrix $P \in \mathbb{R}^{I \times I_N}$, denoted $\mathcal{T} \times_n P$, is a tensor $\mathcal{A}$ of size $I_1 \times \cdots \times I_{n-1} \times J \times I_{n+1} \times \cdots \times I_N$. The $n$-mode product can also be expressed as each $n$-mode fiber multiplied by a matrix, denoted $A_{(n)} = PT_{(n)}$.

For tensor data, as the dimensionality and order increase, the number of parameters will exponentially skyrocket, which is called dimensional catastrophe or dimensional curse, and tensor decomposition can alleviate this problem well. Commonly used tensor decomposition methods include CP decomposition, Tucker decomposition, Tensor Train decomposition, and tensor Ring decomposition. In this paper, the Tucker decomposition method is mainly adopted to operate the tensor data. Tucker decomposition, also known as a form of higher-order principal component analysis, decomposes a tensor into a core tensor multiplied by a factor matrix along each modality, with the following equation:

$$\mathcal{T} = \mathcal{C} \times_1 P_1 \times_2 P_2 \times \cdots \times_N P_N \quad (2)$$
where \( P_i \in \mathbb{R}^{l_i \times r_i} \) denotes the factor matrix along the \( i \)th order modality. The core tensor describing the interaction of the different factor matrices can be denoted by \( C \in \mathbb{R}^{l_1 \times l_2 \times \cdots \times l_N} \). The matrixed form of the Tucker decomposition can be defined as:

\[
T_{(i)} = P_i C_{(i)} (P_N \otimes \cdots \otimes P_{i+1} \otimes P_{i-1} \otimes \cdots \otimes P_1)^T
\]

(3)

where \( \otimes \) is the Kronecker product. The \( l_1 \)-norm of the tensor is defined as \( \|T\|_1 = \sum_{i_1, \ldots, i_N |_{\tau_1, \ldots, \tau_N}} \|C\|_{F} \) and the \( F \)-norm is defined as \( \|T\|_F = \sqrt{\sum_{i_1, \ldots, i_N |_{\tau_1, \ldots, \tau_N}}\|C\|_{F}^2} \).

2.2. Observation Model

The desired HR-HSI can be defined as \( X \in \mathbb{R}^{N_W \times N_H \times N_S} \), the LR-HSI can be denoted as \( Y \in \mathbb{R}^{N_{W_0} \times N_H \times N_S} \) \((0 < N_{w_0} < N_W, 0 < N_h < N_H)\), the HR-MSI can be defined as \( Z \in \mathbb{R}^{N_{W_0} \times N_{H_0} \times N_S} \) \((0 < N_{w_0} < N_W)\). The dimensions of the spatial pattern are \( N_W \) and \( N_H \), and \( N_S \) denotes the dimension of the spectral mode. From the definition of tensor decomposition, we can derive the basic form of hyperspectral high resolution, i.e.,

\[
X = C \times_1 P_1 \times_2 P_2 \times_3 P_3
\]

(4)

The LR-HSI \( Y \) can be expressed as the spatial down-sampling form of the desired HR-HSI \( X \), i.e.,

\[
Y = C \times_1 \hat{P}_1 \times_2 \hat{P}_2 \times_3 P_3
\]

(5)

The HR-MSI \( Z \) can be expressed as the spectral down-sampling form of the desired HR-HSI \( X \), i.e.,

\[
Z = C \times_1 P_1 \times_2 \hat{P}_2 \times_3 \hat{P}_3
\]

(6)

where \( C \in \mathbb{R}^{N_{w_0} \times n_h \times n_s} \) is the core tensor, \( S_1 \in \mathbb{R}^{N_{w_0} \times N_W}, S_2 \in \mathbb{R}^{N_h \times N_H}, S_3 \in \mathbb{R}^{n_h \times N_S} \) are the down-sampling matrices, and \( P_1 \in \mathbb{R}^{N_{W_0} \times n_w}, P_2 \in \mathbb{R}^{N_H \times n_h}, P_3 \in \mathbb{R}^{n_h \times n_s} \) are the dictionaries in the three modes, then, \( \hat{P}_1, \hat{P}_2, \hat{P}_3 \) are the down-sampling dictionaries in the three modes, which can be derived from the following equation:

\[
\hat{P}_1 = S_1 P_1 \in \mathbb{R}^{N_{w_0} \times n_w}, \hat{P}_2 = S_2 P_2 \in \mathbb{R}^{N_h \times n_h}, \hat{P}_3 = S_3 P_3 \in \mathbb{R}^{n_h \times n_s}
\]

(7)

2.3. Joint Regularization

Based on the Tucker decomposition and the factor matrix processed along the tri-mode downsampling, the HSI-SR problem can be expressed by the following equation:

\[
\min_{P_1, \hat{P}_2, \hat{P}_3, C} \|Y - C \times_1 \hat{P}_1 \times_2 \hat{P}_2 \times_3 P_3\|_F^2 + \|Z - C \times_1 P_1 \times_2 \hat{P}_2 \times_3 \hat{P}_3\|_F^2
\]

\[\text{s.t.} \|C\|_0 \leq N\]

(8)

where \( \|\cdot\|_F \) denotes the Frobenius norm and \( N \) denotes the number of nonzero entries in matrix. Clearly, the optimization problem in (8) is non-convex. Aiming for a tractable and scalable approximation optimization, we impose the \( l_1 \)-norm on the core tensor instead of the \( l_0 \)-norm to formulate the unconstrained version and describe the sparsity in spatial and spectral dimensions.

\[
\min_{P_1, \hat{P}_2, \hat{P}_3, C} \|Y - C \times_1 \hat{P}_1 \times_2 \hat{P}_2 \times_3 P_3\|_F^2 + \|Z - C \times_1 P_1 \times_2 \hat{P}_2 \times_3 \hat{P}_3\|_F^2 + \lambda_1 \|C\|_1
\]

(9)

Regardless, problem (9) is still a non-convex problem of discomfort. Therefore, to solve problem (9), some prior information about the target HR-HSI is needed. In this paper, we consider the spectral correlation and spatial coherence of hyperspectral and multispectral images.

As we all know, HSI suffers from high correlation and redundancy in the spectral space and retains the fundamental information in the low-dimensional subspace. Because
of the lack of appropriate regularization item, the fusion model in (9) is sensitive to outliers and noise. Therefore, to accurately estimate the HSI, we used a joint regularization (graph regularization and unidirectional total variation regularization) in the form of a constraint on the HR-MSI and LR-HSI. To obtain accurate results for the target HR-HSI, we first assume that the spatial and spectral manifold information between HR-MSI and LR-HSI is similar to the embedded in the target HR-HSI, and describe the manifold information present in HR-MSI and LR-HSI in the form of two graphs: one based on the spatial dimension and the other on the spectral dimension. Thus, the spatial and spectral information from HR-MSI and LR-HSI can be transferred to HR-HSI by spatial and spectral graph regularization, which can preserve the intrinsic geometric structure information of HR-MSI and LR-HSI as much as possible. After that, we used a unidirectional total variation regularization model to manipulate the three-mode dictionary for the purpose of eliminating noise in the images.

2.3.1. Graph Regularization

We know that the pixels in HR-MSI do not exist independently and the correlation between neighboring pixels is very high. Scholars generally use a block strategy to define adjacent pixels, but this ignores the spatial structure and consistency of the image. As a hyper segmentation method, the hyper-pixel technique not only captures image redundant information, but also adaptively adjusts the shape and size of spatial regions. Considering the compatibility and computational complexity of superpixels, the entropy rate superpixel (ERS) segmentation method is employed in this paper to find spatial domains adaptively. The construction of the spatial graph consists of four steps: generating intensity images, superpixel segmentation, defining spatial neighborhoods, and generating spatial graphs. In contrast, for LR-HSI, its neighboring bands are usually contiguous, meaning that the neighboring bands have extremely strong correlation in the spectral domain. To further maintain the correlation and consistency in HR-HSI, we leverage the nearest neighbor strategy to establish the spectral graph.

2.3.2. Unidirectional Total Variation Regularization

Hyperspectral images are susceptible to noise, which seriously affects the image visual quality and reduces the accuracy and robustness of subsequent algorithms for image recognition, image classification and edge information extraction. Therefore, it is necessary to study effective noise removal algorithms. Common algorithms have the following three categories: the first type of methods is filtering method, including spatial domain filtering and transform domain filtering; the second type of methods is matching method, including moment matching method and histogram matching method; the third type of methods is variation method.

The best known of the variation methods is the total variation (TV) model, an algorithm that has proven to be one of the most effective image denoising techniques. The total variation model is an anisotropic model that relies on gradient descent for image smoothing, hoping to smooth the image as much as possible in the interior of the image (with small differences between adjacent pixels), while not smoothing as much as possible at the edges of the image. The most distinctive feature of this model is that it preserves the edge information of the image while removing the image noise. In general, scholars impose the $l_1$-norm on the total variation model to obtain better denoising effect by improving the total variation model or combining the total variation model with other algorithms. When $l_1$-norm is used in the model, it is insensitive to smaller outliers but sensitive to larger ones; when $l_2$-norm is used, it is insensitive to larger outliers and sensitive to smaller ones; and when $l_\sigma$-norm is used, it can be adjusted by tuning the parameters to be between $l_2$-norm and $l_1$-norm, so that the robustness of both $l_1$-norm and $l_2$-norm is utilized regardless of whether the outliers are large or small, but the burden of tuning parameters $\sigma$ is increased. In order to solve the above problem, the $l_{2,1}$-norm makes the total variation model better handle outliers and reduce the burden of tuning parameters, acting as a flexible embedding.
without the burden of tuning parameters of the \( l_\infty \)-norm. Therefore, in this paper, we impose the \( l_{2,1} \)-norm on the unidirectional total variation model to achieve the purpose of noise removal.

2.4. Proposed Algorithm

Combining the observation model proposed in Section 2.2 with the joint regularization constraint proposed in Section 2.3, the following fusion model is obtained to solve the HSI-SR problem, i.e.,

\[
\min_{P_1, P_2, P_3, C} \|Y - C \times 1 \tilde{P}_1 \times 2 \tilde{P}_2 \times 3 \tilde{P}_3\|_F^2 + \|Z - C \times 1 P_1 \times 2 P_2 \times 3 P_3\|_F^2 + \lambda_1 \|C\|_1 + \beta \text{tr}(P_3^T D_2 P_3) + \gamma \text{tr}(P_2^T (P_2 \otimes P_1)^T D_2 (P_2 \otimes P_1)) + \lambda_2 P_{y1} \|_{2,1} + \lambda_3 \|D_y P_2\|_{2,1} + \lambda_4 \|D_y P_3\|_{2,1}
\]

\[
s.t. X = C \times 1 P_1 \times 2 P_2 \times 3 P_3
\]

where \( X \) denotes the desired HR-HSI, \( Y \) denotes the acquired LR-HSI, \( Z \) denotes the HR-MSI of the same scene, \( P_1, P_2, P_3 \) are the down-sampling dictionaries in the three modes, \( C \) is the core tensor, \( \tilde{P}_1, \tilde{P}_2, \tilde{P}_3 \) are the dictionaries in the three modes, \( D_2, D_3 \) are the graph Laplacian matrices, \( \beta, \gamma \) are the equilibrium parameters of the graph regularization, \( \lambda_i (i = 1, 2, 3, 4) \) are the positive regularization parameters, \( D_y \) is a finite difference operator along the vertical direction, given by the following equation:

\[
D_y = \begin{bmatrix}
1 & -1 & 0 & 0 & \cdots & 0 \\
0 & 1 & -1 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 1 & -1
\end{bmatrix}
\]

Next, we will give an effective algorithm for solving Model (10).

3. Optimization

The proposed model (10) is a non-convex problem by solving \( P_1, P_2, P_3 \) and \( C \) jointly, and we can barely obtain the closed-form solutions for \( P_1, P_2, P_3 \) and \( C \). We know that non-convex optimization problems are considered to be very difficult to solve because the set of feasible domains may have an infinite number of local optima; that is to say, the solution of the problem is not unique. However, with respect to each block of variables, the model proposed in (10) is convex while keeping the other variables fixed. In this context, we utilize the proximal alternating optimization (PAO) scheme [45,46] to solve it, which is ensured to converge to a stationary point under certain conditions. Concretely, the iterative update of model (10) is as follows:

\[
\begin{align*}
&\begin{aligned}
P_1 &= \arg \min_{P_1} f(P_1, P_2, P_3, C) + \rho \|P_1 - P_1^{\text{pre}}\|_F^2 \\
P_2 &= \arg \min_{P_2} f(P_1, P_2, P_3, C) + \rho \|P_2 - P_2^{\text{pre}}\|_F^2 \\
P_3 &= \arg \min_{P_3} f(P_1, P_2, P_3, C) + \rho \|P_3 - P_3^{\text{pre}}\|_F^2 \\
C &= \arg \min_{C} f(P_1, P_2, P_3, C) + \rho \|C - C^{\text{pre}}\|_F^2
\end{aligned}
\end{align*}
\]

where the objective function \( f(P_1, P_2, P_3, C) \) is the implicit definition of (10), and \((\cdot)^{\text{pre}}\) and \( \rho \) represent the estimated blocks of variables in the previous iteration and a positive number, respectively. Next, we present the solution of the four optimization problems in (12) in detail.
3.1. Optimization of $P_1$

With fixing $P_2$, $P_3$ and $C$, the optimization problem of $P_1$ in (12) is given by

$$
\arg\min_{P_1} \left\| Y - C \times_1 \hat{P}_1 \times_2 \hat{P}_2 \times_3 \hat{P}_3 \right\|_F + \left\| Z - C \times_1 P_1 \times_2 P_2 \times_3 \hat{P}_3 \right\|_F + \rho \left\| P_1 - P_1^{pre} \right\|_F^2 \\
\quad + \gamma tr\left( (P_2 \otimes P_1)^T P_D (P_2 \otimes P_1) \right) + \lambda_2 \left\| D_y P_1 \right\|_{2,1}
$$

(13)

where $P_1^{pre}$ denotes the estimated dictionary of width mode in the previous iteration and $D_y \in \mathbb{R}^{(N_y - 1) \times N_y}$ denotes the difference matrix along the vertical direction of $P_1$. Using the properties of n-mode matrix unfolding, problem (13) can be formulated as

$$
\arg\min_{P_1} \left\| Y_{(1)} - S_1 P_1 A_1 \right\|_F + \left\| Z_{(1)} - P_1 B_1 \right\|_F + \rho \left\| P_1 - P_1^{pre} \right\|_F^2 \\
\quad + \gamma tr\left( (P_2 \otimes P_1)^T P_D (P_2 \otimes P_1) \right) + \lambda_2 \left\| D_y P_1 \right\|_{2,1}
$$

(14)

where $Y_{(1)}$ and $Z_{(1)}$ are the width-mode (1-mode) unfolding matrix of tensors $Y$ and $Z$, respectively, $A_1 = (C \times_2 \hat{P}_2 \times_3 \hat{P}_3)_{(1)}$, and $B_1 = (C \times_2 \hat{P}_2 \times_3 \hat{P}_3)_{(1)}$.

3.2. Optimization of $P_2$

With fixing $P_1$, $P_3$ and $C$, the optimization problem of $P_2$ in (12) is given by

$$
\arg\min_{P_2} \left\| Y - C \times_1 \hat{P}_1 \times_2 P_2 \times_3 \hat{P}_3 \right\|_F + \left\| Z - C \times_1 P_1 \times_2 P_2 \times_3 \hat{P}_3 \right\|_F + \rho \left\| P_2 - P_2^{pre} \right\|_F^2 \\
\quad + \gamma tr\left( (P_2 \otimes P_1)^T P_D (P_2 \otimes P_1) \right) + \lambda_3 \left\| D_y P_2 \right\|_{2,1}
$$

(15)

where $P_2^{pre}$ denotes the estimated dictionary of height mode in the previous iteration and $D_y \in \mathbb{R}^{(N_y - 1) \times N_y}$ denotes the difference matrix along the vertical direction of $P_2$. Using the properties of n-mode matrix unfolding, problem (15) can be formulated as

$$
\arg\min_{P_2} \left\| Y_{(2)} - S_2 P_2 A_2 \right\|_F + \left\| Z_{(2)} - P_2 B_2 \right\|_F + \rho \left\| P_2 - P_2^{pre} \right\|_F^2 \\
\quad + \gamma tr\left( (P_2 \otimes P_1)^T P_D (P_2 \otimes P_1) \right) + \lambda_3 \left\| D_y P_2 \right\|_{2,1}
$$

(16)

where $Y_{(2)}$ and $Z_{(2)}$ are the height-mode (2-mode) unfolding matrix of tensors $Y$ and $Z$, respectively, $A_2 = (C \times_1 \hat{P}_1 \times_3 \hat{P}_3)_{(2)}$, and $B_2 = (C \times_1 \hat{P}_1 \times_3 \hat{P}_3)_{(2)}$.

3.3. Optimization of $P_3$

With fixing $P_1$, $P_2$ and $C$, the optimization problem of $P_3$ in (12) is given by

$$
\arg\min_{P_3} \left\| Y - C \times_1 \hat{P}_1 \times_2 \hat{P}_2 \times_3 P_3 \right\|_F + \left\| Z - C \times_1 P_1 \times_2 P_2 \times_3 \hat{P}_3 \right\|_F + \rho \left\| P_3 - P_3^{pre} \right\|_F^2 \\
\quad + \beta tr\left( P_3^T P_3 \right) + \lambda_4 \left\| D_y P_3 \right\|_{2,1}
$$

(17)

where $P_3^{pre}$ denotes the estimated spectral dictionary in the previous iteration and $D_y \in \mathbb{R}^{(N_y - 1) \times N_y}$ denotes the difference matrix along the vertical direction of $P_3$. Using the properties of n-mode matrix unfolding, problem (17) can be formulated as

$$
\arg\min_{P_3} \left\| Y_{(3)} - P_3 A_3 \right\|_F + \left\| Z_{(3)} - S_3 P_3 B_3 \right\|_F + \rho \left\| P_3 - P_3^{pre} \right\|_F^2 \\
\quad + \beta tr\left( P_3^T P_3 \right) + \lambda_4 \left\| D_y P_3 \right\|_{2,1}
$$

(18)
where \(Y^{(3)}\) and \(Z^{(3)}\) are the spectral-mode (3-mode) unfolding matrix of tensors \(Y\) and \(Z\), respectively, \(A_3 = (C \times 1 \hat{P}_1 \times 2 \hat{P}_2)^{(3)}\), and \(B_3 = (C \times 1 \hat{P}_1 \times 2 \hat{P}_2)^{(3)}\).

3.4. Optimization of \(C\)

With fixing \(P_1, P_2\) and \(P_3\), the optimization problem of \(C\) in (12) is given by

\[
\arg \min_C \|Y - C \times 1 \hat{P}_1 \times 2 \hat{P}_2 \times 3 \hat{P}_3\|^2_F + \|Z - C \times 1 \hat{P}_1 \times 2 \hat{P}_2 \times 3 \hat{P}_3\|^2_F + \lambda_1 \|C\|_1 + \rho \|C - C^{pre}\|^2_F
\]

(19)

where \(C^{pre}\) represents the estimated core tensor in the previous iteration.

It should be noted that problems (14), (16), (18) and (19) are convex problems. Therefore, all these four subproblems can be effectively solved using fast and accurate ADMM technique. Due to the similarity of the solution process of problems (14), (16), and (18), we include the solution details of the four subproblems and the optimization updates of each variable as appendices for more conciseness. In Appendix A, Algorithms A1–A4 draw a summary of the solution process of the four subproblems in (12).

Algorithm 1 specifies the steps of the JRLTD-based hyperspectral image super-resolution proposed in this section.

**Algorithm 1 JRLTD-Based Hyperspectral Image Super-Resolution.**

1: Initialize \(P_1, P_2\) through the DUC-KSVD algorithm [47];
2: Initialize \(P_3\) through the SISAL algorithm [48];
3: Initialize \(C\) through the Algorithm A4;
4: while not converged do
5: \hspace{1em} **Step 1** Update the width mode dictionary matrix \(P_1\) via Algorithm A1;
6: \hspace{1em} \(\hat{P}_1 = S_1 P_1, P_1^{pre} = P_1;\)
7: \hspace{1em} **Step 2** Update the height mode dictionary matrix \(P_2\) via Algorithm A2;
8: \hspace{1em} \(\hat{P}_2 = S_2 P_2, P_2^{pre} = P_2;\)
9: \hspace{1em} **Step 3** Update the spectral dictionary matrix \(P_3\) via Algorithm A3;
10: \hspace{1em} \(\hat{P}_3 = S_3 P_3, P_3^{pre} = P_3;\)
11: \hspace{1em} **Step 4** Update the core tensor \(C\) via Algorithm A4;
12: \hspace{1em} \(C^{pre} = C;\)
13: end while
14: Estimating target HR-HSI \(X\) via formula (4)

4. Experiments

4.1. Datasets

In this section, three datasets are used to test the performance of the proposed method. The first dataset is the Pavia University dataset, which was acquired by the Italian Reflection Optical System Imaging Spectrometer (ROSIS) optical sensor in the downtown area of the University of Pavia. The image size is 610 \(\times\) 340 \(\times\) 115, with a spatial resolution of 1.3 m. We reduced the number of spectral bands to 93 after removing the water vapor absorption band. For reasons related to the down-sampling process, only the 256 \(\times\) 256 \(\times\) 93 image in the upper left corner was used as a reference image in the experiment.

The second dataset is the Washington DC dataset, which is obtained from the Washington shopping mall acquired by the HYDICE sensor, intercepting images of size 1280 \(\times\) 307 for annotation. The spatial resolution is 2.5 m and contains a total of 210 bands. We intercept a part of the image with the size of 256 \(\times\) 256 \(\times\) 191 for the experiment and use it as a reference image.

The third dataset is the Sand Lake in Ningxia of China, which is a scene acquired from the GF-5 AHSI sensor during the flight activity in Ningxia. The original image size is 2774 \(\times\) 2554 \(\times\) 330, its spatial resolution is 30 m, and the image has 330 bands, and the
experiments reduce the spectral bands to 93 to obtain the reference image size of Sand Lake as $256 \times 256 \times 93$.

4.2. Compared Algorithms

We selected classical and currently popular fusion methods for comparison, including CNMF [29], HySure [18], NLSTF [38], CSTF [39], and UTV-HSISR [41]. The experiment was run on a PC equipped with an Intel Core i5-9300HF CPU, 16 GB RAM and NVIDIA GTX 1660Ti GPU. The Windows 10 x64 operating system was used and the programming application was Matlab R2016a.

4.3. Quantitative Metrics

For the evaluation of image fusion, it is more important to obtain more convincing values from objective metrics in addition to observing the results from subjective assumptions. To evaluate the fusion output in the numerical results, we use the following eight metrics, namely the peak signal-to-noise ratio (PSNR), which is an objective measure of image distortion or noise level; the error relative global dimensionless synthesis (ERGAS) to measure the comprehensive quality of the fused results; the spectral angle mapping (SAM) represents the absolute value of the spectral angle between two images; the root mean square error (RMSE) is used to measure the deviation between the predicted value and true value; the correlation coefficient (CC), which indicates the ability of the fused image to retain spectral information; the degree of distortion (DD), which is used to indicate the degree of distortion between the fused image and the ground truth image; the structural similarity (SSIM) and the universal image quality index (UIQI), which measures the degree of structural similarity between the two images.

The concept of mean squared deviation is first defined in the paper:

$$MSE = \frac{1}{N_W N_H} \sum_{i=0}^{N_W-1} \sum_{j=0}^{N_H-1} |I(i,j) - J(i,j)|^2$$  \hspace{1cm} (20)$$

where $N_W$ and $N_H$ denote the size of the image, $I$ denotes a noise-free image, and $J$ denotes a noisy image. Then PSNR is defined as:

$$PSNR = 10 \cdot \log_{10} \left( \frac{MAX^2}{MSE} \right)$$  \hspace{1cm} (21)$$

where $MAX$ denotes the maximum number of pixels of the image. After that, the metrics we use to evaluate the fused image can be expressed by the following equation:

$$PSNR(X, \hat{X}) = \frac{1}{N_S} PSNR(X_i, \hat{X}_i)$$  \hspace{1cm} (22)$$

$$ERGAS(X, \hat{X}) = \frac{100}{S} \sqrt{\frac{1}{N_S} \sum_{i=1}^{N_S} \frac{MSE(X, \hat{X})}{\text{MEAN}(X)}}$$  \hspace{1cm} (23)$$

$$SAM(X, \hat{X}) = \frac{1}{N_W N_H} \sum_{i=1}^{N_W N_H} \arccos \left( \frac{X(i,j) \cdot \hat{X}(i,j)}{\|X(i,j)\|_2 \cdot \|\hat{X}(i,j)\|_2} \right)$$  \hspace{1cm} (24)$$

$$RMSE(X, \hat{X}) = \sqrt{\frac{\|X, \hat{X}\|_F}{N_W N_H N_S}}$$  \hspace{1cm} (25)$$

$$CC(X, \hat{X}) = \frac{\sum_{i=1}^{N_W} \sum_{j=1}^{N_H} [X(i,j) - V_X] \cdot [\hat{X}(i,j) - V_{\hat{X}}]}{\sqrt{\sum_{i=1}^{N_W} \sum_{j=1}^{N_H} [X(i,j) - V_X]^2} \cdot \sqrt{\sum_{i=1}^{N_W} \sum_{j=1}^{N_H} [\hat{X}(i,j) - V_{\hat{X}}]^2}}$$  \hspace{1cm} (26)$$
SSIM \((X, \tilde{X}) = \frac{1}{M} \sum_{i=1}^{M} \frac{\left(2\bar{X}_i, \bar{\tilde{X}}_i + c_1\right)\left(2\sigma_{X_i}\sigma_{\tilde{X}_i} + c_2\right)}{\left(X_i\right)^2 + \left(\bar{X}_i\right)^2 + c_1\left(\sigma_{X_i}^2 + \sigma_{\tilde{X}_i}^2 + c_2\right)} \tag{27} \)

\[ DD(X, \tilde{X}) = \frac{1}{N_S} \left\| \text{vec}(X) - \text{vec}(\tilde{X}) \right\|_1 \tag{28} \]

UIQI \((X, \tilde{X}) = \frac{1}{M} \sum_{i=1}^{M} \frac{4\sigma_{X_i}\sigma_{\tilde{X}_i}}{\sigma_{X_i}^2 + \sigma_{\tilde{X}_i}^2 + \left(\bar{X}_i\right)^2 + \left(\bar{\tilde{X}}_i\right)^2} \tag{29} \]

where \(N_S\) denotes the number of bands; \(S\) denotes the spatial downsampling factor; \(X_i, \tilde{X}_i\) denote the value of the \(i\)th band of the ground truth image and the fused image, respectively; \(\text{MEAN}(Z)\) denotes the mean value of each band image; \(\bar{V}_X\) denotes the mean pixel value of the original image; \(\bar{V}_{\tilde{X}}\) is the mean pixel value of the fused image; \(M\) denotes the sliding window; \(\bar{X}_i, \bar{\tilde{X}}_i\) denotes the mean value of \(X_i, \tilde{X}_i\), respectively; \(\sigma_{X_i}, \sigma_{\tilde{X}_i}\) denotes the standard deviation of \(X_i, \tilde{X}_i\); \(c_1, c_2\) are constants; \(\sigma_{X_i, \tilde{X}_i}^2\) denotes the covariance of \(X_i, \tilde{X}_i\). Furthermore, \(\sigma_{X_i}^2, \sigma_{\tilde{X}_i}^2\) denotes the variance of \(X_i, \tilde{X}_i\), respectively. It should be noted that the best value of ERGAS, SAM, RMSE and DD is 0, the best value of CC, SSIM and UIQI is 1, and the best value of PSNR is \(\infty\).

4.4. Parameters Discussion

JRLTD is mainly related to the following parameters, i.e., the number of PAO iterations \(K\), the weights of the proximal terms \(\rho\), the sparse regularization parameters \(\lambda_1\), the smooth regularization parameters \(\lambda_2, \lambda_3\) and \(\lambda_4\), the graph regularization parameters \(\beta\) and \(\gamma\), and the number of three-mode dictionaries \(N_w, N_h\) and \(N_s\).

According to the description of Algorithm 1, we use the PAO scheme to solve the problem (10). The change of PSNR caused by the change in the number of PAO iterations \(K\) is shown in Figure 1. In Figure 1, all three datasets show a fast increasing trend of PSNR as \(K\) goes from 1 to 10. For the PAVIA dataset, there is a slight fluctuation in PSNR when \(K\) varies from 10 to 50, and the maximum number of iterations of PAVIA is set to 20 in the experiment. The Washington dataset reached the maximum PSNR when \(K = 25\), so we set the maximum number of iterations of the algorithm in Washington to 25. Similarly, we set the maximum number of iterations for Sand Lake as 20.

![Figure 1. PSNR values for different K.](image)

The parameter \(\rho\) is the weight of the proximal term in (12). For the evaluation of the influence of \(\rho\), we perform the method for different \(\rho\). Figure 2 presents the change of PSNR values of the fused HSIs of the three datasets with different log \(\rho\) values (the base of log is 10). In the experiments of this paper, we take the range of log \(\rho\) to be set to \([-3, 0]\). As is
displayed in Figure 2, there is a rise trend of PSNR for all three datasets as \( \log \rho \) varies from \(-3\) to \(-1\), reaches a maximum when \( \log \rho \) equals \(-1\), and decreases sharply when \( \log \rho \) is greater than \(-1\). Therefore, we set \( \log \rho \) to \(-1\), i.e., we take \( \rho = 0.1 \) for all three datasets.

Figure 2. PSNR values for different \( \log \rho \).

The regularization parameter \( \lambda_1 \) in (10) controls the sparsity of the core tensor, therefore, \( \lambda_1 \) affects the estimation of the HR-HSI. Higher values of \( \lambda_1 \) yield sparser core tensor. Figure 3 shows the PSNR values of the reconstructed HSI for the Pavia University dataset under different \( \log \lambda_1 \). In this work, we set the range \( \log \lambda_1 \) of to \([-9, -2]\). As shown in Figure 3, when \( \log \lambda_1 \) belongs to \([-9, -5]\), the PSNR stays relatively stable; when \( \log \lambda_1 \) belongs to \([-5, -4]\), the PSNR decreases slowly; and when \( \log \lambda_1 > -4 \), the PSNR decreases sharply. Therefore, we set \( \log \lambda_1 \) as \(-6\), that is, \( \lambda_1 = 10^{-6} \) for the Pavia University dataset. By the same token, the values for the Washington dataset and the Sand Lake dataset can be decided in the same way.

Figure 3. PSNR values for different \( \log \lambda_1 \).

The unidirectional total variation regularization parameters \( \lambda_2 \), \( \lambda_3 \) and \( \lambda_4 \) control the segmental smoothness of the width-mode, height-mode and spectral-mode dictionaries, respectively. Figure 4 shows the reconstructed PSNR values of HSI for the Pavia University dataset with different \( \log \lambda_2 \), \( \log \lambda_3 \) and \( \log \lambda_4 \). In the experiments of this paper, we set the range of values of \( \log \lambda_2 \) and \( \log \lambda_3 \) both to \([-9, -2]\) and the range of values of \( \log \lambda_4 \) to \([-4, 4]\). As shown in Figures 4 and 5, the PSNR reaches its peak value when \( \log \lambda_2 = -8 \), \( \log \lambda_3 = -7 \), and \( \log \lambda_4 = 2 \). Therefore, for Pavia University dataset, we set \( \log \lambda_2 \) as \(-8\), \( \log \lambda_3 \) as \(-7\), and \( \log \lambda_4 \) as 2. It is worth noting that the optimal value of \( \lambda_4 \) is relatively large compared of \( \lambda_2 \) and \( \lambda_3 \), due to the fact that HSI is continuous in the spectral dimension,
which leads to a potentially smaller full variation regularization of the dictionary along the spectral direction. Therefore, the optimal value of its regularization parameter should be relatively large. Similarly, the values of $\lambda_2$, $\lambda_3$ and $\lambda_4$ for the Washington and Sand Lake datasets can be determined in the same way.

The graph regularization parameters $\beta$ and $\gamma$ control the spectral structure of the spectral graph and the spatial correlation of the spatial graph, respectively. Figure 6 shows the reconstructed PSNR values of HSI for the Pavia University dataset under different $\beta$ and $\gamma$. In the experiments of this paper, we take the value range of both $\log \beta$ and $\log \gamma$ to $[-7, -1]$. As shown in Figure 6, the PSNR reaches its peak value when $\log \beta = -1$ and $\log \gamma = -1$. Therefore, for the Pavia University dataset, we set $\log \beta$ as $-1$ and $\log \gamma$ as $-1$. Similarly, the $\beta$ and $\gamma$ values for the Washington dataset and the Sand Lake dataset can be determined in the same way.

![Figure 4. PSNR values for different log $\lambda_2$ and log $\lambda_3$.](image)

![Figure 5. PSNR values for different $\lambda_4$.](image)

The number of atoms in the three-model dictionaries are $n_w$, $n_h$ and $n_s$. Figure 7 shows the PSNR values of the fused HSI of the Pavia University dataset for different $n_w$ and $n_h$, and Figure 8 shows the PSNR values of the fused HSI of the Pavia University dataset for different $n_s$. In this paper, we set the range of values for both $n_w$ and $n_h$ to [260, 400], and set $n_s$ as [3, 21]. This is because the spectral features of HSI exist on the low-dimensional subspace. As shown in Figure 7, the PSNR increases sharply when $n_w$ is varied in the range [260, 360] and reaches a maximum at $n_w = 360$, while it tends to decrease when $n_w$ is varied in the range [360, 400]. Therefore, we set $n_w$ as 360. It should be noted that the PSNR reaches its peak value when $n_h$ is 400, but what we have to consider is the overall performance of other evaluation indicators, so we set $n_h$ as 380 in the paper. It can be seen
from Figure 8 that the PSNR decreases with $n_s > 15$. Therefore, we set $n_w = 360$, $n_h = 400$, and $n_s = 15$ for the Pavia University dataset. Similarly, the values of $n_w$, $n_h$ and $n_s$ for the Washington dataset and the Sand Lake dataset can be determined in the same way.

In Table 1, we give the tuning ranges for the 11 main parameters, give the values of each parameter for the three HSI datasets mentioned in Section 4.1, and show the recommended ranges for each parameter to easily tune the parameters.
Table 1. Discussion of the main parameters.

| Parameters | Tuning Ranges | Pavia University Dataset | Washington DC Dataset | Sand Lake Dataset | Suggested Ranges |
|-----------|---------------|--------------------------|-----------------------|------------------|-----------------|
| K         | [1, 50]       | 20                       | 25                    | 20               | [20, 50]        |
| ρ         | $10^{-3}, 10^0$ | $10^{-1}$               | $10^{-1}$             | $10^{-1}$        | $10^{-1}, 10^0$ |
| $\lambda_1$ | $10^{-5}, 10^{-1}$ | $10^{-6}$               | $10^{-6}$             | $10^{-7}$        | $10^{-7}, 10^{-6}$ |
| $\lambda_2$ | $10^{-5}, 10^{-2}$ | $10^{-7}$               | $10^{-7}$             | $10^{-8}$        | $10^{-8}, 10^{-7}$ |
| $\lambda_3$ | $10^{-9}, 10^{-2}$ | $10^{-6}$               | $10^{-6}$             | $10^{-8}$        | $10^{-6}, 10^{-5}$ |
| $\lambda_4$ | $10^{-4}, 10^4$ | $10^1$                  | $10^2$                | $10^1$           | $10^1, 10^2$    |
| $\beta$   | $10^{-7}, 10^{-1}$ | $10^{-1}$               | $10^{-1}$             | $10^{-3}$        | $10^{-3}, 10^{-1}$ |
| $\gamma$  | $10^{-7}, 10^{-1}$ | $10^{-2}$               | $10^{-2}$             | $10^{-3}$        | $10^{-2}, 10^{-1}$ |
| $N_w$     | [260, 400]    | 360                     | 340                   | 360              | [340, 360]      |
| $N_h$     | [260, 400]    | 380                     | 380                   | 380              | [380, 400]      |
| $N_s$     | [3, 21]       | 15                      | 15                    | 18               | [15, 18]        |

4.5. Experimental Results

In this section, we show the fusion results of the five tested methods for the Pavia University, Washington DC, and Sand Lake datasets.

4.5.1. Experiment On Pavia University

In order to better display more spatial detail information and fusion results, we select three bands (R:61, G:25, B:13) to be synthesized as pseudo-color image for display, and then compared with other methods, the fusion results of Pavia University dataset are shown in the first row of Figure 9. In addition, to show the fusion performance more visually, we generate difference images to present the discrepancy between the reference image and the fused image. The second row in Figure 9 shows the difference image of the Pavia University dataset, which correspond to the fusion results in the first row.

Figure 9. Comparison of fusion results on the Pavia University dataset. (a) Reference Image; (b) LR-HSI; (c) CNMF; (d) HySure; (e) NLSTF; (f) CSTF; (g) UTV-HSISR; (h) JRLTD

From Figure 9, we can see that the spatial details in the fusion results of different methods are greatly enhanced. However, compared with the reference image, there are still some spectral differences and noise effects in the fused image. For example, in Figure 9c,d, the fusion results of CNMF [29] and Hysure [18] show spectral distortion. Compared with the fusion results in Figure 9e,f, the fused images in Figure 9g,h are able to provide better spectral information and preserve the spatial structure.

In addition, it can be seen from the difference images that the reconstruction errors is relatively large from the difference images of Figure 9c–e. Figure 9g,h are better and similar compared with Figure 9f. In other words, the UTV-HSISR algorithm [41] and the JRLTD algorithm proposed in the paper achieve better fusion results, that is, there is little noise.

The quality indicators of the comparison method are shown in Table 2, and the better results obtained in the experiment are highlighted in bold typeface. From the spectral features, the algorithm proposed in this paper has the smallest RMSE, the closest CC to 1, the smallest ERGAS, the smallest SAM, and the smallest DD, indicating that the algorithm proposed in this paper is closest to the reference image, has the smallest spectral distortion, and has the best spectral agreement with the reference image. From the results of signal-to-noise ratio, the algorithm in this paper has the highest PSNR, which indicates that the algorithm has the best effective suppression of noise. From the spatial characteristics, SSIM
is closest to 1, indicating that it is closest to the reference image in terms of brightness, contrast and structure; UIQI is closest to 1, indicating that the loss of relevant information reaches the minimum, the closer to the reference image.

Table 2. Quality evaluation for Pavia University dataset.

| Methods   | RMSE  | CC    | ERGAS | SAM   | DD    | PSNR  | SSIM  | UIQI  |
|-----------|-------|-------|-------|-------|-------|-------|-------|-------|
| BEST      | 0     | 1     | 0     | 0     | 0     | ∞     | 1     | 1     |
| CNMF      | 6.3889| 0.9702| 3.6300| 3.7427| 3.9586| 32.1227| 0.9366| 0.9492|
| HySure    | 4.0104| 0.9880| 2.2397| 3.3363| 2.5411| 36.4850| 0.9703| 0.9790|
| NLSTF     | 2.0265| 0.9966| 1.1602| 2.0873| 1.3064| 44.4323| 0.9706| 0.9928|
| CSTF      | 1.7673| 0.9974| 0.9886| 1.8391| 1.1610| 43.9473| 0.9881| 0.9942|
| UTV-HSISR | 1.6881| 0.9976| 0.9294| 1.7635| 1.0460| 44.6407| 0.9898| 0.9950|
| Proposed  | 1.6552| 0.9977| 0.9072| 1.7097| 1.0105| 44.8388| 0.9905| 0.9952|

4.5.2. Experiment on Washington DC

In order to better display more spatial detail information and fusion results, we select three bands (R:40, G:30, B:5) to be synthesized as pseudo-color image for display, and then compared with other methods, the fusion results of Washington DC dataset are shown in the first row of Figure 10. Besides, in order to show the fusion performance more visually, we generate difference images to present the discrepancy between the reference image and the fused image. The second row of Figure 10 shows the difference image of the Washington DC dataset.

![Figure 10](image_url)

Figure 10. Comparison of fusion results on the Washington DC dataset. (a) Reference Image; (b) LR-HSI; (c) CNMF; (d) HySure; (e) NLSTF; (f) CSTF; (g) UTV-HSISR; (h) JRLTD.

It can be seen that the spectral information is distorted in the results of CNMF [29] and HySure [18]. In addition, there are some blurring effects in the building regions in the results of NLSTF [38] when compared with Figure 10a. Compared with the fusion results of CSTF [39], the fused images of UTV-HSISR [41] and JRLTD are able to provide better spectral information and preserve the spatial structure. From the difference images, we can observe that the error of the UTV-HSISR algorithm [41] and the JRLTD algorithm proposed in the paper is smaller as a whole.

The quality evaluation results are shown in Table 3, and the better values obtained in the experiment are marked with bolded font. From Table 3, it can be seen that the algorithm proposed in this paper has the smallest RMSE, the closest CC to 1, the second minimum value of ERGAS, the smallest SAM, and the smallest DD in terms of spectral features. Collectively, the algorithm proposed in this paper is the closest to the reference image, with the smallest spectral distortion and the best spectral agreement with the reference image. From the results of signal-to-noise ratio, the algorithm in this paper has the highest PSNR, which indicates that the algorithm has the best effective suppression of noise. From the spatial characteristics, SSIM is closest to 1, which indicates that it is closest to the reference image in terms of brightness, contrast and structure; UIQI is closest to 1, which indicates that the loss of relevant information reaches the minimum, the closer to the reference image.
In summary, the JRLTD algorithm proposed in this paper outperforms other algorithms in most cases.

Table 3. Quality evaluation for Washington DC dataset.

| Methods         | RMSE | CC   | ERGAS | SAM  | DD   | PSNR  | SSIM  | UIQI  |
|-----------------|------|------|-------|------|------|-------|-------|-------|
| BEST            | 0    | 1    | 0     | 0    | ∞    | 37.5546 | 0.9585 | 0.9569|
| CNMF            | 4.1122 | 0.9745 | 3.4984 | 3.2825 | 2.9279 | 39.7109 | 0.9778 | 0.9749|
| HySure          | 3.0588 | 0.9837 | 3.7441 | 3.4822 | 1.9632 | 48.1596 | 0.9923 | 0.9919|
| NLSTF           | 1.2778 | 0.9947 | 2.2339 | 1.7381 | 0.7840 | 48.3925 | 0.9945 | 0.9926|
| CSTF            | 1.0618 | 0.9950 | 2.3983 | 1.5433 | 0.6865 | 48.3925 | 0.9945 | 0.9926|
| UTV-HSISR       | 0.9397 | 0.9962 | 2.0301 | 1.3421 | 0.5444 | 49.7023 | 0.9961 | 0.9945|
| Proposed        | 0.8847 | 0.9963 | 2.0478 | 1.2454 | 0.4871 | 50.2731 | 0.9966 | 0.9946|

4.5.3. Experiment on Sand Lake in Ningxia of China

In order to better display more spatial detail information and fusion results, we select three bands (R:41, G:25, B:3) to be synthesized as pseudo-color image for displaying, respectively, and then compared with other methods, the fusion results of Sand Lake dataset are shown in the first row of Figure 11. In addition, to show the fusion performance more visually, we generate difference images to present the discrepancy between the reference image and the fused image. The second row of Figure 11 shows the difference image of the Sand Lake dataset.

Figure 11. Comparison of fusion results on the Sand Lake dataset. (a) Reference Image; (b) LR-HSI; (c) CNMF; (d) HySure; (e) NLSTF; (f) CSTF; (g) UTV-HSISR; (h) JRLTD.

After corresponding the fusion results obtained in the first row of Figure 11 using different algorithms with the difference images in the second row, we can see that Figure 11c–e have spectral distortion compared to the reference image. In addition, we can observe that the Figure 11c–e are poorly reconstructed, so the difference images seems to have a lot of information. From the difference images, Figure 11g,h are better and similar compared to Figure 11f. In other words, the UTV-HSISR algorithm [41] and the JRLTD algorithm proposed in the paper achieve better fusion results, that is, there is little noise.

Furthermore, Table 4 displays the quantitative experimental evaluations with eight metrics. The better values obtained in the experiment are indicated in bold. As can be seen from Table 4, from the spectral features, the algorithm proposed in this paper has the smallest RMSE, the smallest ERGAS, the smallest SAM, the smallest DD, and CC values are the same as those obtained by the UTV-HSISR algorithm. Overall, it shows that the algorithm proposed in this paper is closest to the reference image, has the smallest spectral distortion, and has the best spectral agreement with the source image. From the results of the signal-to-noise ratio, the algorithm in this paper has the highest PSNR, which indicates that the algorithm has the best effective suppression of noise. From the spatial characteristics, SSIM is closest to 1, which indicates that it is closest to the reference image in terms of brightness, contrast and structure; UIQI is closest to 1, which indicates that the loss of relevant information reaches the minimum, the closer to the reference image. In
In general, the JRLTD algorithm proposed in this paper outperforms other algorithms in most cases.

Table 4. Quality evaluation for Sand Lake dataset.

| Methods     | Spectral Features | Signal-To-Noise Ratio | Spatial Features |
|-------------|-------------------|-----------------------|-----------------|
|             | RMSE              | CC                    | ERGAS           | SAM            | DD               | PSNR            | SSIM            | UIQI            |
| BEST        | 0                 | 1                     | 0               | 0              | ∞                | 37.6549         | 0.9681          | 0.9688          |
| CNMF        | 3.5512            | 0.9752                | 1.1293          | 1.1495         | 2.4822           | 39.6945         | 0.9732          | 0.9820          |
| HySure      | 2.9776            | 0.9935                | 1.8847          | 1.3881         | 1.9273           | 44.4597         | 0.9831          | 0.9828          |
| NLSTF       | 2.0026            | 0.9965                | 0.6263          | 1.1535         | 1.4592           | 44.7850         | 0.9860          | 0.9859          |
| CSTF        | 1.5303            | 0.9980                | 0.4853          | 0.9782         | 1.1343           | 50.5421         | 0.9956          | 0.9959          |
| UTV-HSISR   | 0.8926            | 0.9994                | 0.2932          | 0.5514         | 0.5054           | 51.0214         | 0.9962          | 0.9965          |
| Proposed    | 0.8452            | 0.9994                | 0.2786          | 0.5191         | 0.4606           | 51.0214         | 0.9962          | 0.9965          |

5. Conclusions

In this paper, a hyperspectral image super-resolution method using joint regularization as prior information is proposed. Considering the geometric structures of LR-HSI and HR-MSI, two graphs are constructed to capture the spatial correlation of HR-MSI and the spectral similarity of LR-HSI. Then, the presence of anomalous noise values in the images was reduced by smoothing the LR-HSI and HR-MSI using unidirectional total variational regularization. In addition, an optimization algorithm based on PAO and ADMM is utilized to efficiently solve the fusion model. Finally, experiments were conducted on two benchmark datasets and one real dataset. Compared with some fusion methods such as CNMF [29], HySure [18], NLSTF [38], CSTF [39], and UTV-HSISR [41], this fusion method produces better spatial details and better preservation of the spectral structure due to the superiority of joint regularization and tensor decomposition.

However, there are still some limitations, and there is room for improvement of the proposed JRLTD algorithm. For example, the proposed JRLTD algorithm has a high computational complexity, and this leads to a relatively long running time. In our future work, we aim to extend the method in two directions. On the one hand, since the model utilizes the ADMM algorithm, although it is possible to divide a large complex problem into multiple smaller problems that can be solved simultaneously in a distributed manner, leads to an increase in computational effort and a decrease in computational speed. Therefore, we will try to find a closed form solution for each sub-problem. Alternatively, it can be accelerated by using parallel computing techniques. On the other hand, there is non-local spatial similarity in HSI, that is, there are duplicate or similar structures in the image, and when processing blocks of images, we can use information from surrounding blocks of images that are similar to them. This prior information has been shown to be valid for image super-resolution problems. Therefore, we will investigate the incorporation of non-local spatial similarity into the JRLTD method.

Author Contributions: Funding acquisition, W.B.; Validation, K.Q.; Writing—original draft, M.C.; Writing—review & editing, W.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the Natural Science Foundation of Ningxia Province of China (Project No. 2020AAC02028), the Natural Science Foundation of Ningxia Province of China (Project No. 2021AAC03179) and the Innovation Projects for Graduate Students of North Minzu University (Project No.YCX21080).

Acknowledgments: The authors would like to thank the Key Laboratory of Images and Graphics Intelligent Processing of State Ethnic Affairs Commission: IGIPLab for their support.

Conflicts of Interest: The authors declare no conflict of interest.
Abbreviations

HSI-SR Hyperspectral image super-resolution
LR-HSI Low-resolution hyperspectral image
HR-MSI High-resolution multispectral image
HR-HSI High-resolution hyperspectral image
JRLTD Joint regularized low-rank Tensor decomposition method
TV Total variation
ADMM Alternating direction multiplier method
PAO Proximal alternate optimization
CNMF Coupled non-negative matrix factorization
HySure Hyperspectral image super-resolution via subspace-based regularization
NLSTF Non-local sparse tensor factorization
CSTF Coupled sparse Tensor factorization
PSNR Peak signal-to-noise ratio
ERGAS Error relative global dimensionless synthesis
SAM spectral angle mapping
RMSE Root mean square error
CC Correlation coefficient
DD Distortion degree
SSIM Structural similarity
UIQI Universal image quality index

Appendix A

Appendix A.1. Optimization of $P_1$

Problem (14) is convex and can be solved efficiently by ADMM. Thus, we introduce the variable $M = P_1$ and then the unconstrained optimization in (14) can be reexpressed as an equivalent constrained form, i.e.,

$$
\begin{align*}
\arg\min_{P_1} & \left\| Y(1) - S_1 P_1 A_1 \right\|_F^2 + \left\| Z(1) - P_1 B_1 \right\|_F^2 + \rho \left\| P_1 - P_{pre} \right\|_F^2 \\
+ \gamma & \text{tr} \left( (P_2 \otimes P_1)^T P_D (P_2 \otimes P_1) \right) + \lambda_2 \left\| D_2 P_1 \right\|_{2,1} \\
\text{s.t.} & \, P_1 = M
\end{align*}
$$

(A1)

It is easy to deduce that the augmented Lagrangian function for problem (A1) is

$$
\begin{align*}
L(P_1, M, V_1) &= \left\| Y(1) - S_1 P_1 A_1 \right\|_F^2 + \left\| Z(1) - P_1 B_1 \right\|_F^2 + \rho \left\| P_1 - P_{pre} \right\|_F^2 + \mu_1 \left\| P_1 - M - V_1 \right\|_F^2 \\
+ \gamma & \text{tr} \left( (P_2 \otimes M)^T P_D (P_2 \otimes M) \right) + \lambda_2 \left\| D_2 M \right\|_{2,1}
\end{align*}
$$

(A2)

where $V_1$ denotes the Lagrange multiplier and $\mu_1$ represents a positive penalty parameter. We solve (A2) using the ADMM algorithm:

$$
\begin{align*}
P_1^{(t+1)} &= \arg\min_{P_1} L(P_1, M^{(t)}, V_1^{(t)}) \\
M^{(t+1)} &= \arg\min_{M} L(P_1^{(t+1)}, M, V_1^{(t)}) \\
V_1^{(t+1)} &= \arg\min_{V_1} L(P_1^{(t+1)}, M^{(t+1)}, V_1)
\end{align*}
$$

(A3)

(1) $P_1$-Subproblem: From (A2), we have

$$
\begin{align*}
\arg\min_{P_1} & \left\| Y(1) - S_1 P_1 A_1 \right\|_F^2 + \left\| Z(1) - P_1 B_1 \right\|_F^2 + \rho \left\| P_1 - P_{pre} \right\|_F^2 + \mu_1 \left\| P_1 - M - V_1 \right\|_F^2
\end{align*}
$$

(A4)
The optimization problem in (A4) is quadratic, which has a unique solution, and it is equal to compute the following Sylvester matrix equation, i.e.,

\[ S^T_1 S_1 A_1^T + P_1 \left( B_1 B_1^T + \rho I \right) + \mu_1 I P_1 = S^T_1 Y_{(1)} A_1^T + Z_{(1)} B_1^T + \rho P^{pre}_1 + \mu_1 I (M + V_1) \]  

(A5)

We adopt the CG [46] to solve (A5) efficiently.

(2) M-Subproblem: From (A2), we have

\[ \arg \min_{M} \gamma \text{tr} \left( (P_2 \otimes M)^T P_D (P_2 \otimes M) \right) + \lambda_2 \| D_y M \|_{2,1} + \mu_1 \| P_1 - M - V_1 \|_F^2 \]  

(A6)

Note that it is complicated to solve due to the Kronecker product involved in the regularization of the spatial graph. Taking advantage of the symmetry and positive-semidefinite of Laplacian matrices, we simplify formula (A6) by implementing the Cholesky factorization [49] of \( P_D \). After that, we obtain a more contextually specific and brief function with respect to \( M \) as

\[ \arg \min_{M} \gamma \| M^T U_{(1)} \|_F^2 + \lambda_2 \| D_y M \|_{2,1} + \mu_1 \| P_1 - M - V_1 \|_F^2 \]  

(A7)

where \( U_{(1)} \) is the matrix obtained by performing the Cholesky decomposition of \( P_D \) and combining it with the Tucker2 decomposition model. The solution of the function (A7) can be obtained from the following equation:

\[ (\mu_1 I - \lambda_2 D_y \Sigma_D) M = \gamma U_{(1)} + \mu_1 I (P_1 - V_1) \]

\[ M = (\mu_1 I - \lambda_2 D_y \Sigma_D)^{-1} \left[ \gamma U_{(1)} + \mu_1 I (P_1 - V_1) \right] \]  

(A8)

where \( I \) denotes a unit matrix of appropriate size, \( \Sigma_D = \begin{bmatrix} 1 / \| M_1 \|_2 & 1 / \| M_2 \|_2 & \cdots & 1 / \| M_{nw} \|_2 \end{bmatrix} \).

(3) \( V_1 \)-Subproblem: From (A2), the Lagrangian multiplier \( V_1 \) can be updated by the following formula:

\[ V_1 = V_1 - (P_1 - M) \]  

(A9)

Specifically, the each step of solving \( P_1 \)-subproblem (13) by the ADMM is summarized in Algorithm A1.

**Algorithm A1** Solve \( P_1 \)-subproblem (13) with ADMM.

**Input:** \( Y, Z, P_2, \tilde{P}_2, P_3, \tilde{P}_3, C, D_y, P^{pre}_1, \rho > 0, \gamma > 0, \mu_1 > 0, \) and \( \lambda_2 > 0 \).

**Output:** Dictionary matrix \( P_1 \).

1: while not converged do
2:   **Step 1** Update the dictionary matrix \( P_1 \) via (A5);
3:   **Step 2** Update the variable \( M \) via (A8);
4:   **Step 3** Update the Lagrangian multiplier \( V_1 \) via (A9);
5: end while

**Appendix A.2. Optimization of \( P_2 \)**

Like problem (14), problem (16) can be solved efficiently with ADMM. Hence, we introduce the variable \( N = P_2 \) and then the unconstrained optimization in (16) can be rephrased into an equivalent constrained form, i.e.,
\[
\begin{align*}
\arg\min_{P_2} & \left\| Y_{(2)} - S_2 P_2 A_2 \right\|_F^2 + \left\| Z_{(2)} - P_2 B_2 \right\|_F^2 + \rho \left\| P_2 - P_2^{pre} \right\|_F^2 \\
& + \gamma \text{tr} \left( (P_2 \otimes P_1)^T P_D (P_2 \otimes P_1) \right) + \lambda_3 \left\| D_y P_2 \right\|_{2,1} \\
\text{s.t.} & P_2 = N
\end{align*}
\]

(A10)

It is easy to deduce that the augmented Lagrangian function for problem (A10) is

\[
L(P_2, N, V_2) = \left\| Y_{(2)} - S_2 P_2 A_2 \right\|_F^2 + \left\| Z_{(2)} - P_2 B_2 \right\|_F^2 + \rho \left\| P_2 - P_2^{pre} \right\|_F^2 + \mu_2 \left\| P_2 - N - V_2 \right\|_F^2 \\
+ \gamma \text{tr} \left( (N \otimes P_1)^T P_D (N \otimes P_1) \right) + \lambda_3 \left\| D_y N \right\|_{2,1}
\]

(A11)

where \( V_2 \) denotes the Lagrangian multiplier and \( \mu_2 \) represents a positive penalty parameter. We solve (A11) using the ADMM algorithm:

\[
\begin{cases}
P_2^{(t+1)} = \arg\min_{P_2} L \left( P_2, N^{(t)}, V_2^{(t)} \right) \\
N^{(t+1)} = \arg\min_{N} L \left( P_2^{(t+1)}, N, V_2^{(t)} \right) \\
V_2^{(t+1)} = \arg\min_{V_2} L \left( P_2^{(t+1)}, N^{(t+1)}, V_2 \right)
\end{cases}
\]

(A12)

(1) \( P_2 \)-Subproblem: From (A11), we have

\[
\begin{align*}
\arg\min_{P_2} & \left\| Y_{(2)} - S_2 P_2 A_2 \right\|_F^2 + \left\| Z_{(2)} - P_2 B_2 \right\|_F^2 + \rho \left\| P_2 - P_2^{pre} \right\|_F^2 + \mu_2 \left\| P_2 - N - V_2 \right\|_F^2 \\
& + \gamma \text{tr} \left( (N \otimes P_1)^T P_D (N \otimes P_1) \right) + \lambda_3 \left\| D_y N \right\|_{2,1} \quad \text{(A13)}
\end{align*}
\]

The optimization problem in (A13) is quadratic, which has a unique solution, and it is equal to compute the following Sylvester matrix equation, i.e.,

\[
S_2^T S_2 P_2 A_2 A_2^T + P_2 \left( B_2 B_2^T + \rho I \right) + \mu_2 I P_2 = S_2^T Y_{(2)} A_2^T + Z_{(2)} B_2^T + \rho P_2^{pre} + \mu_2 I (N + V_2) \quad \text{(A14)}
\]

We adopt the CG to solve (A14) efficiently.

(2) \( N \)-Subproblem: From (A11), we have

\[
\begin{align*}
\arg\min_{N} & \gamma \text{tr} \left( (N \otimes P_1)^T P_D (N \otimes P_1) \right) + \lambda_3 \left\| D_y N \right\|_{2,1} + \mu_2 \left\| P_2 - N - V_2 \right\|_F^2 \\
& \quad \text{subject to } N = N \quad \text{(A15)}
\end{align*}
\]

Note that the same calculation of the Kronecker product is needed here, and we can use the method of solving for \( M \) to solve for the solution with respect to \( N \).

(3) \( V_2 \)-Subproblem: From (A11), the Lagrangian multiplier \( V_2 \) can be updated by the following formula:

\[
V_2 = V_2 - (P_2 - N) \quad \text{(A16)}
\]

Specifically, the each step of solving \( P_2 \)-subproblem (15) by the ADMM is summarized in Algorithm A2.

**Algorithm A2** Solve \( P_2 \)-subproblem (15) with ADMM.

**Input:** \( Y, Z, P_1, \hat{P}_1, P_3, \hat{P}_3, C, D_y, P_2^{pre}, \rho > 0, \gamma > 0, \mu_2 > 0, \) and \( \lambda_3 > 0. \)

**Output:** Dictionary matrix \( P_2 \).

1. **while** not converged **do**
2. **Step 1** Update the dictionary matrix \( P_2 \) via (A14);
3. **Step 2** Update the variable \( N \) via (A15);
4. **Step 3** Update the Lagrangian multiplier \( V_2 \) via (A16);
5. **end while**
Appendix A.3. Optimization of $P_3$

Like problem (14), problem (18) can be solved efficiently with ADMM. Hence, we introduce the variable $O = P_3$ and then the unconstrained optimization in (18) can be rephrased into an equivalent constrained form, i.e.,

$$\arg \min_{P_3} \|Y(3) - S_3P_3A_3\|_F^2 + \|Z(3) - P_3B_3\|_F^2 + \rho \|P_3 - P_{3}^{\text{pre}}\|_F^2 + \beta \text{tr} \left(P_3^T P_3 P_3 \right) + \lambda_4 \|D_y P_3\|_{2,1}$$

s.t. $P_3 = O$  \hspace{1cm} (A17)

It is easy to deduce that the augmented Lagrangian function for problem (A17) is

$$L(P_3, O, V_3) = \|Y(3) - S_3P_3A_3\|_F^2 + \|Z(3) - P_3B_3\|_F^2 + \rho \|P_3 - P_{3}^{\text{pre}}\|_F^2 + \mu_3 \|P_3 - O - V_3\|_F^2 + \beta \text{tr} \left(O^T P_3 O \right) + \lambda_4 \|D_y O\|_{2,1}$$

where $V_3$ denotes the Lagrangian multiplier and $\mu_3$ represents a positive penalty parameter. We solve (A18) using the ADMM algorithm:

$$\begin{aligned}
    & P_3^{(t+1)} = \arg \min_{P_3} \left( \frac{\mu_3}{2} \|P_3 - O^{(t)}\|_F^2 + \beta \text{tr} \left(P_3^T P_3 P_3 \right) + \lambda_4 \|D_y P_3\|_{2,1} \right) \\
    & O^{(t+1)} = \arg \min_{O} \left( \frac{\mu_3}{2} \|P_3^{(t+1)} - O\|_F^2 + \lambda_4 \|D_y O\|_{2,1} \right) \\
    & V_3^{(t+1)} = \arg \min_{V_3} \left( \frac{\mu_3}{2} \|P_3^{(t+1)} - O^{(t+1)}\|_F^2 + \lambda_4 \|D_y P_3\|_{2,1} \right)
\end{aligned}$$

(1) $P_3$-Subproblem: From (A18), we have

$$\arg \min_{P_3} \|Y(3) - S_3P_3A_3\|_F^2 + \|Z(3) - P_3B_3\|_F^2 + \rho \|P_3 - P_{3}^{\text{pre}}\|_F^2 + \mu_3 \|P_3 - O - V_3\|_F^2$$  \hspace{1cm} (A20)

The optimization problem in (A20) is quadratic, which has a unique solution, and it is equal to compute the following Sylvester matrix equation, i.e.,

$$S_3^T S_3 P_3 A_3 A_3^T + P_3 \left( B_3 B_3^T + \rho I \right) + \mu_3 I P_3 = S_3^T Y(3) A_3 A_3^T + Z(3) B_3 B_3^T + \rho P_{3}^{\text{pre}} + \mu_3 I (O + V_3)$$  \hspace{1cm} (A21)

We adopt the CG to solve (A21) efficiently.

(2) $O$-Subproblem: From (A18), we have

$$\arg \min_{O} \beta \text{tr} \left(O^T P_3 O \right) + \lambda_4 \|D_y O\|_{2,1} + \mu_3 \|P_3 - O - V_3\|_F^2$$  \hspace{1cm} (A22)

After that, we obtain the closed solution of $O$

$$O = (2\beta P_3 + \mu_3 I + \lambda_4 D_y \Sigma_S)^{-1} \left[ \mu_3 I (P_3 + V_3) \right]$$  \hspace{1cm} (A23)

where $I$ denotes a unit matrix of appropriate size, $\Sigma_S = \begin{bmatrix} \frac{1}{\|\Omega_1\|_2} & 0 & \cdots \\ 0 & \frac{1}{\|\Omega_2\|_2} & \cdots \\ \vdots & \ddots & \ddots \\ 0 & \cdots & \frac{1}{\|\Omega_N\|_2} \end{bmatrix}$.

(3) $V_3$-Subproblem: From (A18), the Lagrangian multiplier $V_3$ can be updated by the following formula:

$$V_3 = V_3 - (P_3 - O)$$  \hspace{1cm} (A24)
Specifically, the each step of solving $P_3$-subproblem (17) by the ADMM is summarized in Algorithm A3.

**Algorithm A3** Solve $P_3$-Subproblem (17) with ADMM.

**Input:** $Y, Z, P_1, \hat{P}_1, P_2, \hat{P}_2, \mathcal{C}, D_y, P_3^{pre}, \rho > 0, \beta > 0, \mu_3 > 0$, and $\lambda_4 > 0$.

**Output:** Dictionary matrix $P_3$.

1: while not converged do
2:   **Step 1** Update the dictionary matrix $P_3$ via (A21);
3:   **Step 2** Update the variable $O$ via (A23);
4:   **Step 3** Update the Lagrangian multiplier $V_3$ via (A24);
5: end while

**Appendix A.4. Optimization of $\mathcal{C}$**

Problem (19) is convex and can be solved efficiently by ADMM algorithm by introducing two auxiliary variables $\mathcal{C}_1 = \mathcal{C}$ and $\mathcal{C}_2 = \mathcal{C}$ and then reformulate the problem (19) as follows:

$$\arg\min_{\mathcal{C}_1, \mathcal{C}_2} f(\mathcal{C}) + f(\mathcal{C}_1) + f(\mathcal{C}_2)$$

$$\text{s.t.} \mathcal{C}_1 = \mathcal{C}, \mathcal{C}_2 = \mathcal{C}$$

(A25)

where

$$f(\mathcal{C}) = \lambda_1 \|\mathcal{C}\|_1 + \rho \|\mathcal{C} - \mathcal{C}_{pre}\|_F^2$$

$$f_1(\mathcal{C}_1) = \|Y - \mathcal{C}_1 \times_1 \hat{P}_1 \times_2 \hat{P}_2 \times_3 \mathcal{P}_3\|_F^2$$

$$f_2(\mathcal{C}_2) = \|Z - \mathcal{C}_2 \times_1 \hat{P}_1 \times_2 \hat{P}_2 \times_3 \mathcal{P}_3\|_F^2$$

(A26)

It is easy to deduce that the augmented Lagrangian function for problem (A26) is

$$L(\mathcal{C}, \mathcal{C}_1, \mathcal{C}_2, V_4, V_5) = \lambda_1 \|\mathcal{C}\|_1 + \rho \|\mathcal{C} - \mathcal{C}_{pre}\|_F^2 + \|Y - \mathcal{C}_1 \times_1 \hat{P}_1 \times_2 \hat{P}_2 \times_3 \mathcal{P}_3\|_F^2$$

$$+ \mu_4 \|\mathcal{C} - \mathcal{C}_1 - V_4\|_F^2 + \|Z - \mathcal{C}_2 \times_1 \hat{P}_1 \times_2 \hat{P}_2 \times_3 \mathcal{P}_3\|_F^2 + \mu_4 \|\mathcal{C} - \mathcal{C}_2 - V_5\|_F^2$$

(A27)

where $V_4$, $V_5$ denotes the Lagrangian multiplier and $\mu_4$ represents a positive penalty parameter.

ADMM iterations of (A27) are shown below:

$$\mathcal{C}^{(t+1)} = \arg\min_{\mathcal{C}} L(\mathcal{C}, \mathcal{C}_1^{(t)}, \mathcal{C}_2^{(t)}, V_4^{(t)}, V_5^{(t)})$$

$$\mathcal{C}_1^{(t+1)} = \arg\min_{\mathcal{C}_1} L(\mathcal{C}^{(t+1)}, \mathcal{C}_1^{(t)}, \mathcal{C}_2^{(t)}, V_4^{(t)}, V_5^{(t)})$$

$$\mathcal{C}_2^{(t+1)} = \arg\min_{\mathcal{C}_2} L(\mathcal{C}^{(t+1)}, \mathcal{C}_1^{(t+1)}, \mathcal{C}_2^{(t+1)}, V_4^{(t)}, V_5^{(t)})$$

$$V_4^{(t+1)} = \arg\min_{V_4} L(\mathcal{C}^{(t+1)}, \mathcal{C}_1^{(t+1)}, \mathcal{C}_2^{(t+1)}, \mathcal{C}_4^{(t+1)}, V_5^{(t)})$$

$$V_5^{(t+1)} = \arg\min_{V_5} L(\mathcal{C}^{(t+1)}, \mathcal{C}_1^{(t+1)}, \mathcal{C}_2^{(t+1)}, \mathcal{C}_4^{(t+1)}, V_4^{(t+1)}, V_5^{(t)})$$

(A28)

(1) $\mathcal{C}$-Subproblem: From (A27), we have

$$\arg\min_{\mathcal{C}} \lambda_1 \|\mathcal{C}\|_1 + \rho \|\mathcal{C} - \mathcal{C}_{pre}\|_F^2 + \mu_4 \|\mathcal{C} - \mathcal{C}_1 - V_4\|_F^2 + \mu_4 \|\mathcal{C} - \mathcal{C}_2 - V_5\|_F^2$$

(A29)

whose solution $\mathcal{C}$ can be easily derived by columnwise vector-soft threshold function as:

$$\mathcal{C} = \text{soft}_{\lambda_1} \left[ \frac{\mu_4 (\mathcal{C}_1 + V_4 + \mathcal{C}_2 + V_5) + \rho C_{pre}}{2\mu_4 + \rho} \right]$$

(A30)
where \( \text{soft}(x, y) = \text{sign}(x) \ast (|x| - y, 0) \).

(2) \( C_1 \)-Subproblem: From (A27), we have

\[
\arg\min_{C_1} \mu_4 \|C_1 - C + V_4\|_F^2 + \|Y - C_1 \times_1 \hat{P}_1 \times_2 \hat{P}_2 \times_3 \hat{P}_3\|_F^2
\]  

(A31)

Problem (A31) is equal to

\[
\arg\min_{C_1} \mu_4 \|c_1 - c + v_4\|_F^2 + \|y - Q_1 c_1\|_F^2
\]  

(A32)

where the vectors \( c_1 = \text{vec}(C_1) \), \( c = \text{vec}(C) \), \( v_4 = \text{vec}(V_4) \) and \( y = \text{vec}(Y) \) are generated by vectorizing the tensors \( C_1, C, V_4 \) and \( Y \), respectively, and \( Q_1 = \hat{P}_3 \otimes \hat{P}_2 \otimes \hat{P}_1 \).

Problem (A32) has the following closed-form solution, i.e.,

\[
c_1 = \left( Q_1^T Q_1 + \mu_4 I \right)^{-1} (Q_1^T y + \mu_4 c - \mu_4 v_4)
\]  

(A33)

Note that \( Q_1 \in \mathbb{R}^{N_a N_b N_c \times n_a n_b n_c} \) is extremely large, and formula in (A33) is complicated to solve. Fortunately, we find that

\[
\left( Q_1^T Q_1 + \mu_4 I \right)^{-1} = (D_3 \otimes D_2 \otimes D_1)(\Sigma_3 \otimes \Sigma_2 \otimes \Sigma_1 + \mu_4 I)^{-1} \times \left( D_3^T \otimes D_2^T \otimes D_1^T \right)
\]  

(A34)

where \( \Sigma_i \) and \( D_i(i = 1, 2, 3) \) are diagonal matrices and unitary matrices containing the eigenvalues and eigenvectors of \( P_1 \hat{P}_1 \), \( P_2 \hat{P}_2 \), and \( P_3 \hat{P}_3 \), respectively.

Therefore, \( (\Sigma_3 \otimes \Sigma_2 \otimes \Sigma_1 + \mu_4 I)^{-1} \) is a diagonal matrix and could be computed easily. Besides, the term \( Q_1^T y \) in (A33) can be computed by

\[
Q_1^T y = \text{vec}(Y \times_1 \hat{P}_1^T \times_2 \hat{P}_2^T \times_3 \hat{P}_3^T)
\]  

(A35)

(3) \( C_2 \)-Subproblem: From (A27), we have

\[
\arg\min_{C_2} \mu_4 \|C_2 - C + V_5\|_F^2 + \|Z - C_2 \times_1 P_1 \times_2 P_2 \times_3 \hat{P}_3\|_F^2
\]  

(A36)

Problem (A36) is equal to

\[
\arg\min_{C_2} \mu_4 \|c_2 - c + v_5\|_F^2 + \|z - Q_2 c_2\|_F^2
\]  

(A37)

where the vectors \( c_2 = \text{vec}(C_2) \), \( c = \text{vec}(C) \), \( v_5 = \text{vec}(V_5) \) and \( z = \text{vec}(Z) \) are generated by vectorizing the tensors \( C_2, C, V_5 \) and \( Z \), respectively, and \( Q_2 = \hat{P}_3 \otimes \hat{P}_2 \otimes \hat{P}_1 \).

Problem (A37) has the following closed-form solution, i.e.,

\[
c_2 = \left( Q_2^T Q_2 + \mu_4 I \right)^{-1} (Q_2^T z + \mu_4 c - \mu_4 v_5)
\]  

(A38)

Note that \( Q_2 \in \mathbb{R}^{N_a N_b N_c \times n_a n_b n_c} \) is extremely large, and formula in (A38) is complicated to solve. Fortunately, we find that

\[
\left( Q_2^T Q_2 + \mu_4 I \right)^{-1} = (D_3 \otimes D_2 \otimes D_1)(\Sigma_3 \otimes \Sigma_2 \otimes \Sigma_1 + \mu_4 I)^{-1} \times \left( D_3^T \otimes D_2^T \otimes D_1^T \right)
\]  

(A39)

where \( \Sigma_i \) and \( \bar{D}_i(i = 1, 2, 3) \) are diagonal matrices and unitary matrices containing the eigenvalues and eigenvectors of \( P_1^T \hat{P}_1 \), \( P_2^T \hat{P}_2 \), and \( P_3^T \hat{P}_3 \), respectively.
Therefore, \((\hat{\Sigma}_3 \otimes \hat{\Sigma}_2 \otimes \hat{\Sigma}_1 + \mu_4 I)^{-1}\) is a diagonal matrix and could be computed easily. (3) \(V_4\) and \(V_5\)-Subproblem: From (A27), the multipliers \(V_4\) and \(V_5\) can be updated by the following formulas:

\[
\begin{align*}
V_4 &= V_4 - (C - C_1) \\
V_5 &= V_5 - (C - C_2)
\end{align*}
\] (A40)

Specifically, the each step of solving \(C\)-subproblem (19) by the ADMM is summarized in Algorithm A4.

**Algorithm A4** Solve \(C\)-Subproblem (19) with ADMM.

**Input:** \(Y, Z, P_1, P_2, P_3, C_{\text{psr}}, \rho > 0, \mu_4 > 0, \) and \(\lambda_1 > 0.\)

**Output:** Core tensor \(C\).

1. while not converged do
2. \textbf{Step 1} Update \(C\) via (A30);
3. \textbf{Step 2} Update \(C_1\) via (A33);
4. \textbf{Step 3} Update \(C_2\) via (A38);
5. \textbf{Step 4} Update \(V_4\) and \(V_5\) via (A40);
6. end while

---

**References**

1. Bioucas-Dias, J.M.; Plaza, A.; Camps-Valls, G.; Scheunders, P.; Nasrabadib, N.; Chanussot, J. Hyperspectral remote sensing data analysis and future challenges. *IEEE Geosci. Remote. Sens. Mag.* 2013, 1, 6–36. [CrossRef]
2. Vivone, G.; Alparone, L.; Chanussot, J.; Dalla Mura, M.; Garzelli, A.; Licciardi, G.A.; Restaino, R.; Wald, L. A critical comparison among pansharpening algorithms. *IEEE Trans. Geosci. Remote Sens.* 2014, 53, 2565–2586. [CrossRef]
3. Aiazzi, B.; Alparone, L.; Baronti, S.; Garzelli, A.; Selva, M. Years of pansharpening: A critical review and new developments. In *Signal Image Processing for Remote Sensing, 2nd ed.*; Chen, C.H., Ed.; CRC Press: Boca Raton, FL, USA, 2012; Volume 25, pp. 533–548.
4. Meng, X.; Shen, H.; Li, H.; Zhang, L.; Fu, R. Review of the pansharpening methods for remote sensing images based on the idea of meta-analysis: Practical discussion and challenges. *Inf. Fusion* 2019, 46, 102–113. [CrossRef]
5. Vivone, G.; Restaino, R.; Licciardi, G.; Dalla Mura, M.; Chanussot, J. Multiresolution analysis and component substitution techniques for hyperspectral pansharpening. In *Proceedings of the 2014 IEEE Geoscience and Remote Sensing Symposium, Quebec City, QC, Canada, 13–18 July 2014*; pp. 2649–2652.
6. Kwarteng, P.; Chavez, A. Extracting spectral contrast in Landsat Thematic Mapper image data using selective principal component analysis. *Photogramm. Eng. Remote. Sens.* 1999, 55, 339–348.
7. Feng-Hua, H.; Lu-Ming, Y. Study on the hyperspectral image fusion based on the gram_schmidt improved algorithm. *Inf. Technol. J.* 2013, 12, 6694. [CrossRef]
8. Alparone, L.; Baronti, S.; Aiazzi, B.; Garzelli, A. Spatial methods for multispectral pansharpening: Multiresolution analysis demystified. *IEEE Trans. Geosci. Remote Sens.* 2016, 54, 2563–2576. [CrossRef]
9. Li, H.; Manjunath, B.; Mitra, S.K. Multisensor image fusion using the wavelet transform. *Graph. Model. Image Process.* 1995, 57, 235–245. [CrossRef]
10. Liu, J. Smoothing filter-based intensity modulation: A spectral preserve image fusion technique for improving spatial details. *Int. J. Remote Sens.* 2000, 21, 3461–3472. [CrossRef]
11. Aiazzi, B.; Alparone, L.; Baronti, S.; Garzelli, A.; Selva, M. MTF-tailored multiscale fusion of high-resolution MS and Pan imagery. *Photogramm. Eng. Remote Sens.* 2006, 72, 591–596. [CrossRef]
12. Yuan, Q.; Wei, Y.; Meng, X.; Shen, H.; Zhang, L. A multiscale and multidepth convolutional neural network for remote sensing imagery pan-sharpening. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2018, 11, 978–989. [CrossRef]
13. Bungert, L.; Coomes, D.A.; Ehrhardt, M.J.; Rasch, J.; Reisenhofer, R.; Schönlieb, C.B. Blind image fusion for hyperspectral imaging with the directional total variation. *Inverse Probl.* 2018, 34, 044003. [CrossRef]
14. Bajaj, C.; Wang, T. Blind Hyperspectral-Multispectral Image Fusion via Graph Laplacian Regularization. *arXiv* 2019, arXiv:1902.08224.
15. Ghaderpour, E. Multichannel antileakage least-squares spectral analysis for seismic data regularization beyond aliasing. *Acta Geophys.* 2019, 67, 1349–1363. [CrossRef]
16. Miao, J.; Cao, H.; Jin, X.B.; Ma, R.; Fei, X.; Niu, L. Joint sparse regularization for dictionary learning. *Cogn. Comput.* 2019, 11, 697–710. [CrossRef]
17. He, Z.; Wang, Y.; Hu, J. Joint sparse and low-rank multitask learning with laplacian-like regularization for hyperspectral classification. *Remote Sens.* 2018, 10, 322. [CrossRef]
48. Bioucas-Dias, J.M. A variable splitting augmented Lagrangian approach to linear spectral unmixing. In Proceedings of the 2009 First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, Grenoble, France, 26–28 August 2009; pp. 1–4.

49. Dereniowski, D.; Kubale, M. Cholesky factorization of matrices in parallel and ranking of graphs. In International Conference on Parallel Processing and Applied Mathematics; Springer: Berlin/Heidelberg, Germany, 2003; pp. 985–992.