Article

Weighted Group Sparsity-Constrained Tensor Factorization for Hyperspectral Unmixing

Xinxi Feng¹*, Le Han² and Le Dong³

1 College of Artificial Intelligence, Yangzhou University, Fuzhou 350015, China
2 College of Computer Science and Technology, Zhejiang University, Hangzhou 310007, China; hanle@zju.edu.cn
3 School of Artificial Intelligence, Xidian University, Xi'an 710071, China; dongle@xidian.edu.cn

* Correspondence: xyda@yyu.edu.cn

Abstract: Recently, unmixing methods based on nonnegative tensor factorization have played an important role in the decomposition of hyperspectral mixed pixels. According to the spatial prior knowledge, there are many regularizations designed to improve the performance of unmixing algorithms, such as the total variation (TV) regularization. However, these methods mostly ignore the similar characteristics among different spectral bands. To solve this problem, this paper proposes a group sparse regularization that uses the weighted constraint of the $L_{2,1}$ norm, which can not only explore the similar characteristics of the hyperspectral image in the spectral dimension, but also keep the data smooth characteristics in the spatial dimension. In summary, a non-negative tensor factorization framework based on weighted group sparsity constraint is proposed for hyperspectral images. In addition, an effective alternating direction method of multipliers (ADMM) algorithm is used to solve the algorithm proposed in this paper. Compared with the existing popular methods, experiments conducted on three real datasets fully demonstrate the effectiveness and advancement of the proposed method.

Keywords: hyperspectral image; unmixing; nonnegative tensor factorization; total variation; group sparsity

1. Introduction

Hyperspectral remote sensing images have a wide range of applications in the field of earth observation because of the rich spectral information [1–3]. However, the increase in spectral resolution usually results in a decrease in spatial resolution due to the influence of spectral splitting [4]. As a result, hyperspectral images generally have low spatial resolution, which leads to the inclusion of multiple substances in the image pixels. Therefore, mixed pixels are widely present in the hyperspectral images (HSI). The purpose of hyperspectral unmixing is to extract all pure materials (called endmembers) in the HSI and the percentage of endmembers in each pixel (called abundances) [5].

To solve the above-mentioned unmixing problem, numerous models have been proposed, the most popular of which is the linear mixed model (LMM). Under the LMM, each of the mixed pixels is regarded as the linear expression form of the endmembers, and the expression coefficients of endmembers constitute the abundance matrix [6] of the image. Among them, nonnegative matrix factorization (NMF) has been favored by researchers for its clear physical meaning and simple optimization process. The purpose of NMF is to find a base matrix of the original data and the expression coefficients of endmembers constitute the abundance matrix [6] of the image. In unmixing task, the optimized base matrix is the endmember matrix while the expression coefficient matrix is the abundance matrix. However, the solution of the simple NMF model is a non-convex problem, which leads to unstable unmixing results [8]. Therefore, to improve the performance of unmixing, many improved NMF-based unmixing methods have been proposed, for instance, sparse-based and spatial-based.
According to the prior of the distribution of ground objects, the endmembers in the mixed pixels in HSI are much smaller than the endmember spectral library [9]. Therefore, the abundance matrix is reasonably sparse. As a result, the sparsity constraint of abundance matrix becomes one of the most popular constraints. The $L_0$ regularization is the most ideal sparse regularizer [10]. Unfortunately, the NMF-based unmixing algorithm with $L_0$ sparsity constraint is an NP-hard problem [11]. As an improvement, the $L_p$ norm is widely used in the unmixing frameworks based on NMF to improve the stability and effectiveness of the algorithm. Guo et al. [12] uses the $L_1$ regularization to relax the $L_0$ regularization while [13] uses the $L_2$ regularization to constrain the abundance matrix. $L_{1/2}$ has also been proven to make the abundance matrix get a good sparse result [14,15]. Some methods also use the $L_{2,1}$ regularization to simultaneously explore the sparse structure in the abundance map [16], which has been proven to have good performance in the unmixing algorithm.

Besides, the spatial distribution of ground objects has the characteristics of local similarity and global similarity [17,18]. That is, there is a certain spatial similarity structure in HSI. Therefore, spatial similarity constraints are also really popular technics to improve the effectiveness of NMF-based methods. Lu et al. [19] propose a manifold regularization, which learns the global similarity structure of data to constrain the abundance matrix and force the abundance vectors of similar pixels to be similar. Manifold constraints have also proved by [20,21] to be a really effective way to explore the spatial structures. In [22], it is considered that the abundance of a pixel is the consistence as the average value of its neighboring pixels to maintain the smooth characteristics of HSI. In addition to these, the total variation regularization [23] and its variation [24] are also applied to the NMF-based unmixing algorithms, which can effectively retain the smoothing relationship of the image in the abundance matrix. In [25], the clustering algorithm is used to explore the global similarity relationship of pixels in the HSI, and it is combined with the NMF framework to improve the performance of unmixing.

The NMF-based methods mentioned above have significantly improved the performance of unmixing. Unfortunately, this kind of algorithms has a great flaw. The NMF-based framework reshapes the original 3-D HSI into a 2-D matrix before unmixing. This method of image reshaping severely destroys the 3-D cube structure of HSI, and it is difficult to completely explore the missing spatial information even if spatial regularization constraints are subsequently used [26]. To slow down the loss of information and make better use of all the spatial and spectral information in HSI, many researchers have begun to use nonnegative tensor factorization (NTF) instead of NMF framework to carry out unmixing tasks. The NTF framework directly treats HSI as a 3-D tensor, and regards the image as a whole during the optimization process, which can completely transfer all the spatial relationships of the original image to the abundance tensor.

Zhang et al. [27] innovatively applies tensor factorization (TF) to unmixing for the first time, which obtains performance comparable to NMF framework. The canonical polyadic decomposition-based (CPD) method is used to extract endmember matrix in [28]. However, this kind of methods decomposes the image into a finite number of rank-1 tensors, the number of which is generally much larger than the actual number of endmembers. Another TF method, tucker factorization, is also applied to the unmixing task in [29]. Regrettably, tucker factorization often fails to correspond the unmixing result with the actual distribution of objects. To make the result of TF method has a clear physical meaning, Qian et al. [26] propose a matrix-vector-based NTF (MV-NTF) method, which decomposes the HSI into multiple component tensors, then each component tensor is expressed by the outer product of a matrix and a vector. This method greatly promotes the application of NTF framework in unmixing task. In addition, many researchers add some effective regularizations to improve the effectiveness of the MV-NTF method. For instance, a spatial low-rank tensor regularization is proposed by [30,31] to maintain the sparsity of the abundance tensor. Sun et al. [32] uses a local spatial similarity constraint to explore the local smooth features of the data. A regularization for exploring the low-rank spatial structure of data is also introduced in [33]. The purpose is to preserve the HSI properties completely in the abun-
dance tensor to improve the performance of unmixing. In addition, TV regularization has also been added to the NTF framework, such as [23,34].

Although the above methods have good performance in hyperspectral unmixing, the NTF-based models still have room for improvement in combining spatial and spectral information. According to [24], the difference image obtained after the image is differentiated along the horizontal and vertical directions is also sparse, which not only has zero elements, but also has zero connected regions in space. Therefore, a weighted group sparsity-constrained tensor factorization (WSCTF) method is proposed in this paper. Specifically, the HSI is differentiated along the horizontal and vertical directions at the same time to obtain two three-dimensional difference images firstly, which can explore the sparse structure of the HSI. Subsequently, the difference images are folded along the mode-1 and mode-2 to obtain two sparse matrices. Then, the $L_{2,1}$ norm is used to constrain the sparse structure of difference matrix, which can not only maintain the sparseness of the difference image, but also constrain the zero connected region. To enhance the sparsity, the proposed method WSCTF also provides a weight for the sparsity constraint in vertical and horizontal directions. In addition, the $L_{2,1}$ regularization is used to constrain the sparsity of the abundance tensor. In summary, the contributions of this paper can be concluded as:

- A sparse constrained regularization is proposed to explore the sparse structure of the HSI differential image in the horizontal and vertical directions. In addition, weight coefficients are used to enhance sparsity.
- The $L_{2,1}$ norm is utilized to constrain the sparse regularizer, and the $L_{2,1}$ norm is embedded in the NTF framework to maintain the sparsity of the abundance tensor.
- The proposed algorithm WSCTF uses the alternating direction method of multipliers to iteratively optimize. Experiments on three real data prove the effectiveness of our algorithm.

The paper is arranged as follows. Section 2 introduces the concepts related to tensor and the unmixing framework based on NTF. The proposed method WSCTF and its optimization are described in detail in Section 3. The synthetic and real experimental results are analyzed in Section 4. Section 5 is the conclusion of this paper.

2. Related Work

2.1. Notation

According to experiences, a matrix is a two-dimensional array, and a tensor is an extension of the matrix. In this paper, an italic letter is used to represent a scalar, e.g., $y$ and $I$. A vector is expressed by a regular lowercase letter, e.g., $y \in \mathbb{R}^{I \times 1}$. A matrix is given using a bold captial letter, e.g., $M \in \mathbb{R}^{I_1 \times I_2}$, and a tensor is defined as a calligraphic letter, e.g., $Y \in \mathbb{R}^{I_1 \times I_2 \times \ldots \times I_N}$.

**Tensor Mode**: Each dimension of a tensor is called a mode. For instance, the tensor $Y \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ has 3 modes. **Mode-n Unfolding**: It can reshape a tensor to a matrix form. For example, given the tensor $Y \in \mathbb{R}^{I \times J \times K}$, the mode-n unfolding of $Y$ is represented as:

\[
\begin{align*}
(Y_{<1>})_{(j-1)K+k,j} &= y_{ijk}, \\
(Y_{<2>})_{(k-1)I+j,k} &= y_{ijk}, \\
(Y_{<3>})_{(i-1)J+j,k} &= y_{ijk},
\end{align*}
\]

in which $Y_{<1>}$ is the mode-1 unfolding of tensor $Y$. $Y_{<2>}$ and $Y_{<3>}$ are the mode-2 and mode-3 unfolding, respectively.

**Frobenius Norm**: The Frobenius Norm of tensor $Y \in \mathbb{R}^{I \times J \times K}$ is given by:

\[
||Y||_F = \left( \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} |y_{ijk}|^2 \right)^{1/2}.
\]
Mode-n Product: Given a tensor $\mathcal{Y}$, it can be expressed as $\mathcal{Y} = A \times_n M$, in which $A \times_n$ denotes the mode-n unfolding of tensor $A$.

2.2. NTF Unmixing Method

According to [35], the linear spectral mixing model can be represented as a linear combination by the endmember matrix and abundance maps. Specifically, given hyperspectral data $Y \in \mathbb{R}^{I \times J \times K}$, then it can be written as:

$$Y = A \times_3 M + N,$$

where $A \in \mathbb{R}^{I \times J \times P}$ is the abundance tensor, $A \times_3$ denotes the mode-3 unfolding of tensor $A$. And $M \in \mathbb{R}^{I \times J}$ and $N \in \mathbb{R}^{I \times J \times K}$ are the endmember matrix and noise tensor existing in HSI, respectively. Besides, $I$, $J$ and $K$ are the dimensions of hyperspectral data, and $P$ is the number of endmembers. As a result, the unmixing model based on tensor factorization can be given by

$$J(A, M) = \frac{1}{2} \|Y - A \times_3 M\|_F^2,$$

in which $\| \cdot \|_F$ denotes the Frobenius norm. Since hyperspectral unmixing is a practical image processing process, there must be non-negative constraint here. In addition, the sum of the abundance coefficients inside each pixel should be one. These two properties create two constraints in the function (4). Specifically, the objective function (4) can be updated as:

$$J(A, M) = \frac{1}{2} \|Y - A \times_3 M\|_F^2,$$

s. t. $A \geq 0$, $A \times_1 P = 1_{I \times J}$,

(5)

in which $A \times_1$ is mode-1 unfolding of the abundance tensor $A$.

Unfortunately, the above function (5) is a non-convex problem, which needs to be optimized for both the abundance $A$ and the endmember matrix $M$. To improve the performance of the algorithm, Ref. [36] adds the $L_{2,1}$ sparsity constraint to the NTF unmixing model. Therefore, it can be given as:

$$J(A, M) = \frac{1}{2} \|Y - A \times_3 M\|_F^2 + \lambda_S \|A\|_{2,1}$$

s. t. $A \geq 0$, $A \times_1 P = 1_{I \times J}$,

(6)

In addition, some spatial constraints are also added to the NTF unmixing model to explore the spatial information of HSI, especially TV regularization [34]. Specifically, the NTF unmixing model based on the TV regularization can be described as follows:

$$J(A, M) = \frac{1}{2} \|Y - A \times_3 M\|_F^2 + \lambda_S \|A\|_{2,1} + \lambda_{TV} \|A\|_{TV}$$

s. t. $A \geq 0$, $A \times_1 P = 1_{I \times J}$,

(7)

where $\|A\|_{TV}$ is the TV regularization, and $\lambda_{TV}$ denotes the weight parameter. Since the TV regularization has achieved excellent results in the spatial relationships of data, especially in maintaining the smoothness of the data, it has achieved comparable performance on the hyperspectral unmixing, which has been proved in [37,38]. However, the existing methods are still unable to fully explore some important characteristics of data. For instance, few work consider continuity in the direction of the spectral dimension. In addition, according to [39], the three-dimensional difference images obtained after also have group sparsity. Based on these ideas, this paper explores the three-dimensional spatial structure information of the image, including the spectral similarity of the difference images and the sparsity of the spatial group.
3. Proposed Method

In response to the questions mentioned above, a new regularization is designed to explore the spectral continuity and group sparsity of difference images of HSI. As a result, a weighted group sparsity constraint non-negative tensor factorization (WSCTF) method for hyperspectral unmixing is proposed. The proposed algorithm and its optimization method are described in detail in this section.

3.1. WSCTF Model

According to the existing experience, the TV regularization is widely used to explore the spatial smooth structure in HSI. Specifically, the TV regularizer can be expressed as follows:

$$\|A\|_{TV} = \|D_1 A\|_1 + \|D_2 A\|_1,$$

(8)

where $D$ denotes the difference operator. $D_1$ and $D_2$ are vertical and horizontal difference, respectively. $\| \cdot \|_1$ is the $l_1$ norm. In detail, given the spatial coordinate of an abundance tensor $(i, j, p)$, then the difference results in the vertical and horizontal directions can be described as:

$$D_1 A(i, j, p) = A(i, j+1, p) - A(i, j, p),$$

$$D_2 A(i, j, p) = A(i+1, j, p) - A(i, j, p).$$

(9)

Based on the above analysis, it can be concluded that the difference images in the horizontal and vertical directions are also three-dimensional tensors. According to the prior, the difference image has continuity in the spectral dimension. In addition, the difference image has extremely high sparsity. In the difference image, not only many elements are zeros, but also some connected regions with zeros elements. Therefore, a new weighted sparsity regularization is designed to explore these properties. Specifically, the regularization can be given as:

$$\mathcal{R}(A) = \|W_1 \odot (D_1 A_{\times 1})\|_{2,1} + \|W_2 \odot (D_2 A_{\times 2})\|_{2,1},$$

(10)

in which $W_1$ and $W_2$ are aiming at enhancing the sparsity of difference images. $A_{\times 1}$ and $A_{\times 2}$ are the mode-1 and mode-2 unfolding of tensor $A$. $\| \cdot \|_{2,1}$ denotes the $l_{2,1}$ norm, which can be used to explore the group sparsity of difference image. Therefore, the objective function (7) can be improved by:

$$J(A, M) = \frac{1}{2} \| Y - A_{\times 3} M \|_F^2 + \lambda_S \| A \|_{2,1} + \lambda \mathcal{R}(A)$$

s. t. $A \geq 0, \ A_{\times 1} 1_p = 1_{I \times J}$,

(11)

where $\lambda$ is regularization parameter. In addition, to be more prepared to explore the sparse structure of the abundance tensor $A$, a weight constraint is also added to the sparse regularization. Finally, the objective function of WSCTF proposed in this paper can be expressed as:

$$J(A, M) = \frac{1}{2} \| Y - A_{\times 3} M \|_F^2 + \lambda_S \| W_S \odot A \|_{2,1} + \lambda \mathcal{R}(A)$$

s. t. $A \geq 0, \ A_{\times 1} 1_p = 1_{I \times J}$,

(12)

in which $W_S$ is the weight to promote the sparsity of abundance tensor. In the proposed method, not only the spatial neighborhood smoothing information is considered, but also the sparsity characteristics of the difference image, especially the group sparsity, are described. The optimization method of objective function (12) is explained in detail in the next subsection.

3.2. Optimization

The Formula (12) is a non-convex problem, which is directly difficult to solve. According to the exist experience, the ADMM algorithm is used to update each variable, which can be solved by decomposing the original problem into multiple sub-problems.
(1) **Update M**: For the function (12), it first considers the update of the endmember matrix \( M \). At this time, the abundance tensor \( A \) is fixed, so the objective function can be expressed as:

\[
J(A, M) = \frac{1}{2} \| Y - A \times_3 M \|_F^2.
\]  

(13)

It can directly obtain the updated rule of endmember \( M \) as:

\[
M \leftarrow M \ast (Y \times_3 A \times_3^T) \ast (MA \times_3 A \times_3^T),
\]  

(14)

in which \( Y \times_3 \) and \( A \times_3 \) are the mode-3 unfolding of tensors. \( \ast \) and \( \ast / \) are multiplication and division operators.

In addition, by introducing six intermediate variables \( G_1, G_2, G_3, G_4, G_5 \) and \( G_6 \), the Formula (12) can be rewritten as:

\[
\begin{align*}
\min_{A, M} & \frac{1}{2} \| G_1 - Y \|_F^2 + \lambda_s \| W_S \odot G_2 \|_{2,1} + \lambda \| W_1 \odot G_3 \|_{2,1} + \lambda \| W_2 \odot G_4 \|_{2,1} + l_{R^+}(G_5) + l_{(1)}(G_6), \\
\text{s.t.} & \quad G_1 = A \times_3 M, \\
& \quad G_2 = A, \\
& \quad G_3 = D_1 A \times_1, \\
& \quad G_4 = D_2 A \times_2, \\
& \quad G_5 = A, \\
& \quad G_6 = A.
\end{align*}
\]  

(15)

In the above Formula (15), \( G_5 \) and \( G_6 \) are used to meet the abundance nonnegative constraint (ANC) and abundance sum-to-one constraint (ASC). They are defined as:

\[
\begin{align*}
l_{R^+}(X) = \left\{ \begin{array}{ll}
0, & \text{if } X \geq 0, \\
+\infty, & \text{otherwise}
\end{array} \right. \\
l_{(1)}(X) = \left\{ \begin{array}{ll}
0, & \text{if } \text{sum}(X) = 1, \\
+\infty, & \text{otherwise}.
\end{array} \right.
\end{align*}
\]  

(16)

in which \( 1 \) denotes a matrix with all ones. Hence, the Lagrangian function of objective function (15) is:

\[
\begin{align*}
\min_{A, M, G_1} & \frac{1}{2} \| G_1 - Y \|_F^2 + \lambda_s \| W_S \odot G_2 \|_{2,1} + \lambda \| W_1 \odot G_3 \|_{2,1} + \lambda \| W_2 \odot G_4 \|_{2,1} + l_{R^+}(G_5) + l_{(1)}(G_6) \\
& + \frac{\mu}{2} \| G_1 - A \times_3 M + H_1 \|_F^2 + \frac{\mu}{2} \| G_2 - A + H_2 \|_F^2 + \frac{\mu}{2} \| G_3 - D_1 A \times_1 + H_3 \|_F^2 \\
& + \frac{\mu}{2} \| G_4 - D_2 A \times_2 + H_4 \|_F^2 + \frac{\mu}{2} \| G_5 - A + H_5 \|_F^2 + \frac{\mu}{2} \| G_6 - A + H_6 \|_F^2.
\end{align*}
\]  

(17)

(2) **Update A**: In problem (17), the subproblem of \( A \) can become:

\[
\begin{align*}
\min_{A} & \frac{\mu}{2} \| G_1 - A \times_3 M + H_1 \|_F^2 + \frac{\mu}{2} \| G_2 - A + H_2 \|_F^2 + \frac{\mu}{2} \| G_3 - D_1 A \times_1 + H_3 \|_F^2 \\
& + \frac{\mu}{2} \| G_4 - D_2 A \times_2 + H_4 \|_F^2 + \frac{\mu}{2} \| G_5 - A + H_5 \|_F^2 + \frac{\mu}{2} \| G_6 - A + H_6 \|_F^2.
\end{align*}
\]  

(18)

In problem (18), it can be directly solved by:

\[
A \times_3 M^T M + \sum_{k=1}^2 A \times k (D_k^T D_k) + 3A = \mathcal{K}
\]  

(19)

where

\[
\mathcal{K} = (G_1 + H_1) \times_3 M^T + (G_3 + H_3) \times_1 D_1^T + (G_4 + H_4) \times_2 D_2^T + (G_2 + H_2 + G_5 + H_5 + G_6 + H_6)
\]  

(20)
(19) is equal as
\[ A_{(3)}^T M^T M + CA_{(3)}^T + 3A_{(3)}^T = K_{(3)}^T \]  
(21)

where \( C = [I \otimes D_1^T D_1] + [D_2^T D_2 \otimes I] \). \( C \) is a matrix with a locking loop structure. It can be easily concluded that \( M^T M \) is a symmetric matrix. According to [39], 2-D fast Fourier transformation and SVD is employed to solve \( C \) and \( M^T M \), respectively. Then it can be obtained that \( C = F_2^T \Psi F_2 \) and \( M^T M = U_2^T \Sigma U_2^T \), in which \( F_2 \) is a 2-D discrete Fourier transformation matrix. Then the update rule of \( A \) is
\[ A = \text{Fold}_3((F_2^T ((1 \otimes T_2) \odot (F_2 K_{(3)}^T U_2^T)))U_2^T)^T \]  
(22)

in which \( T_2 = (\text{diag}(\Psi_2), \text{diag}(\Psi_2), \ldots, \text{diag}(\Psi_2)) + (\text{diag}(\Sigma), \text{diag}(\Sigma), \ldots, \text{diag}(\Sigma))^T \).

(3) **Update** the intermediate variables \( \mathcal{G}_i \):

\[
\begin{align*}
&\min_{\mathcal{G}_i} \frac{1}{2} \| \mathcal{G}_1 - Y \|_F^2 + \lambda_{12}(\mathcal{G}_3) + \lambda_{11}(\mathcal{G}_6) + \frac{\mu}{2} \| \mathcal{G}_1 - A \times_3 M + \mathcal{H}_1 \|_F^2 + \frac{\mu}{2} \| \mathcal{G}_2 - A + \mathcal{H}_2 \|_F^2 \\
&+ \frac{\mu}{2} \| \mathcal{G}_3 - D_1 A_{1 \times 1} + \mathcal{H}_3 \|_F^2 + \frac{\mu}{2} \| \mathcal{G}_4 - D_2 A_{2 \times 2} + \mathcal{H}_4 \|_F^2 + \frac{\mu}{2} \| \mathcal{G}_5 - A + \mathcal{H}_5 \|_F^2 + \frac{\mu}{2} \| \mathcal{G}_6 - A + \mathcal{H}_6 \|_F^2.
\end{align*}
\]  
(23)

It can be easily concluded that the update formula of \( \mathcal{G}_1 \)
\[ \mathcal{G}_1 \leftarrow \frac{1}{1 + \mu} [Y + \mu (A \times_3 M - \mathcal{H}_1)]. \]  
(24)

According to [40], the update rules related to \( \mathcal{G}_2 \) can be expressed as:
\[
\mathcal{G}_2 = \begin{cases} 
\frac{||H_2(i,j,:)||_2 - W_2(i,j)\lambda_2}{||H_2(i,j,:)||_2}, & \text{if } W_2(i,j)\lambda_2 < ||H_2(i,j,:)||_2, \\
0, & \text{otherwise}.
\end{cases}
\]  
(25)

Observing the problem (19), it can be seen the intermediate variables \( \mathcal{G}_2, \mathcal{G}_3, \) and \( \mathcal{G}_4 \) have similar forms. Hence, their solving rules are also similar, then the closed solutions of \( \mathcal{G}_3 \) and \( \mathcal{G}_4 \) can be directly obtained as:
\[
\begin{align*}
\mathcal{G}_3 &= \begin{cases} 
\frac{||C_1(i,j,:)||_2 - W_1(i,j)\lambda_2}{||C_1(i,j,:)||_2}, & \text{if } W_1(i,j)\lambda_2 < ||C_1(i,j,:)||_2, \\
0, & \text{otherwise}.
\end{cases}
\end{align*}
\]  
(26)

where \( C_1 = D_1 A_{1 \times 1} + \mathcal{H}_5. \)
\[
\begin{align*}
\mathcal{G}_4 &= \begin{cases} 
\frac{||C_2(i,j,:)||_2 - W_2(i,j)\lambda_2}{||C_2(i,j,:)||_2}, & \text{if } W_2(i,j)\lambda_2 < ||C_2(i,j,:)||_2, \\
0, & \text{otherwise}.
\end{cases}
\end{align*}
\]  
(27)

where \( C_2 = D_2 A_{2 \times 2} + \mathcal{H}_4. \)

It is worth noting that both \( W_3, W_1 \) and \( W_2 \) are weight matrixes, which are used to increase the degree of the sparsity. Therefore, their update rules can be described as:
\[
W_3(i,j,k) = \frac{1}{||\mathcal{G}_2(i,j,k)||_2 + \epsilon},
\]  
(28)

\[
W_1(i,j) = \frac{1}{||\mathcal{C}_1(i,j,:)||_2 + \epsilon},
\]  
(29)

\[
W_2(i,j) = \frac{1}{||\mathcal{C}_2(i,j,:)||_2 + \epsilon},
\]  
(30)
in which \( \epsilon \) is a small positive number.
Besides, \( G_5 \) and \( G_6 \) are aiming at ensuring ANC and ASC of abundance. Therefore, the update rule of them can be written as:

\[
G_5 \leftarrow \max(A - H_5, 0),
\]

\[
G_6 \leftarrow (A - H_6) + \text{repmat}(N, I(1), 1),
\]

\[
N = \frac{1}{I}[1 - \text{sum}(A - H_6)].
\]

(32)

(4) **Update** the Lagrange multipliers \( H_i \):

\[
H_1 = H_1 - A_{1\times2}M + G_1,
\]

\[
H_2 = H_2 - A + G_2,
\]

\[
H_3 = H_3 - D_1A_{1\times3} + G_3,
\]

\[
H_4 = H_4 - D_2A_{1\times2} + G_4,
\]

\[
H_5 = H_5 - A + G_5,
\]

\[
H_6 = H_6 - A + G_6.
\]

In general, the framework of the algorithm WSCTF proposed in this paper is summarized in Algorithm 1.

**Algorithm 1** The proposed method SCLT

**Input:**

- \( Y \in \mathbb{R}^{I\times J\times K} \)—the mixing hyperspectral data; 
- \( P \)—the number of endmembers; 
- Parameter \( \lambda_S \), and \( \lambda \).

1. Update the endmember matrix \( M \) by (14);
2. Update the abundance tensor \( A \) via (22);
3. Update the variables \( G_1 \) by (24), \( G_2 \) by (25), \( G_3 \) by (26), \( G_4 \) by (27), \( G_5 \) and \( G_6 \) via (31) and (32);
4. Update the Lagrangian multipliers \( H_1, H_2, H_3, H_4, H_5, H_6 \) by (33).

**Output:**

- \( A \)—the abundance tensor; 
- \( M \)—the endmember matrix.

**4. Experiments**

To verify the effectiveness of the WSCTF algorithm proposed in this paper, this section conducts related experiments on two synthetic data and real data sets of Cuprite, Jasper Ridge and Samon, and discusses the experimental results in detail. Since the proposed WSCTF is based on a non-negative tensor decomposition algorithm, the five comparison algorithms used here are: ULTRA-V [30], SULRSR-TV [24], MV-NTF [26], NMF-QMV [41], NL-TSUn [33]. It is worth noting that ULTRA-V, MV-NTF and NL-TSUn are all based on NTF, while SULRSR-TV and NMF-QMV are based on NMF. In addition, the SULRSR-TV algorithm also uses the TV regularization to constrain the local smooth structure of the abundance. The purpose of setting up these comparative experiments is to comprehensively verify the effectiveness of our algorithm.

Spectral angle distance (SAD) and root-mean-square error (RMSE) are the most widely used evaluation indicators in unmixing tasks. Among them, SAD is used to describe the spectral angular distance between the estimated endmember and the ground truth. Correspondingly, the distance of the estimated abundance and the ground truth is described by RMSE. They can be expressed in mathematical formulas as:

\[
\text{RMSE} = \sqrt{|A - \hat{A}|^2},
\]

\[
\text{SAD} = \cos^{-1}\left(\frac{\hat{M}^T M}{\|\hat{M}\|\|M\|}\right)
\]

(34)
in which $\hat{A}$ and $\hat{M}$ are the ground truth of abundances and endmembers, respectively. Besides, to ensure the fairness of comparison, all experiments use the same experimental platform and use the vertex component analysis (VCA) and fully constrained least squares (FCLS) algorithm to initialize the endmembers and abundances. In addition, it is worth noting that the average and variance of the results of all experiments have been presented after 20 runs, which can illustrate the stability of algorithms.

4.1. Synthetic Data

Firstly, experiments are conducted on two synthetic data. All pure endmember spectra in the data are randomly sampled from the USGS spectral library [42].

1. Dataset 1 (DS1): Set the number of endmembers $p = 5$, the size of the HSI is $75 \times 75 \times 224$, where 224 is the number of spectral band. The specific data synthesis method is similar to the paper [34]. It is worth noting that each pixel here should comply with the LMM and meet the abundance ANC and ASC at the same time. Figure 1a shows the false color image of DS1.

![Figure 1](image)

Figure 1. The false images of data. (a) DS1 (R = 25 band, G = 40 band, B = 175 band). (b) DS2 (R = 15 band, G = 20 band, B = 24 band). (c) Jasper Ridge (R = 20 band, G = 57 band, B = 100 band). (d) Samon (R = 65 band, G = 130 band, B = 125 band). (e) Cuprite (R = 55 band, G = 125 band, B = 120 band).

2. Dataset 2 (DS2): DS2 can form a control group with DS1. Here, the method described in [34] is used to synthesize DS2. Set the image size of DS2 to $90 \times 90 \times 224$, where 224 is the spectral band. Figure 1b shows the false color image of DS2. According to the above, two clean synthetic data can be obtained. Subsequently, 10 dB, 20 dB and 30 dB Gaussian noise are added to DS1 and DS2 respectively to form six data interfered by different noises. All experiments are all performed on these six data, which can prove the effectiveness of algorithms under different image sizes, different noise conditions, and different numbers of endmembers.

The SAD and RMSE results of the proposed WSCTF and five comparison algorithms on these six data are shown in Tables 1 and 2. According to the results of Tables 1 and 2, we can analyze the results: (1) On the DS1 and DS2 data, as the noise continues to increase, the performance of all algorithms, including the proposed WSCTF, is declining. However, the performance of WSCTF method changes less under the influence of different noises and the performance is more stable than other methods, since it fully considers the spatial and spectral similarity. (2) MV-NTF performs poorly on both data, because MV-NTF is based on the NTF basic algorithm, and other algorithms are mostly improved methods on it. (3) Since DS2 is more complicated than DS1 and the number of endmembers is more, all algorithms perform slightly worse on DS2. The ULTRA-V algorithm is mainly proposed for the variability with endmembers, but the absence of variable endmembers in the settings of DS1 and DS2 leads to poor unmixing performance. (4) The lack of pure pixels in the DS2 is the main reason for the low performance of the SULRSR-TV method. SULRSR-TV assumed that there must be pure pixels in the image when unmixing. In summary, compared with other comparison algorithms, the proposed WSCTF has certain advantages under the conditions of different noises, different image sizes and different numbers of endmembers. The above analysis can also be verified in the visualization results of abundance. As shown...
in Figure 2, the abundance and groundtruth obtained by all algorithms are displayed on DS1 with 20 dB. The abundance map estimated by the WSCTF method is obviously better than other comparison algorithms. Due to the large number of endmembers in DS2 with 20 dB, only five abundances are shown in Figure 3 here.

Table 1. RMSE results on the synthetic data.

| Data | SNR  | SULRSR-TV | NMF-QMV | MV-NTF | NL-TSUn | ULTRA-V | WSCTF    |
|------|------|-----------|---------|--------|---------|---------|----------|
|      | 10 dB| 0.0995 ± 0.0049 | 0.0960  | 0.1906 ± 0.0041 | 0.1447 ± 0.0621 | 0.2339 ± 0.0031 | 0.0827 ± 0.0019 |
| DS1  | 20 dB| 0.0805 ± 0.0077 | 0.0723  | 0.1500 ± 0.0420 | 0.1024 ± 0.0352 | 0.1701 ± 0.0214 | 0.0682 ± 0.0241 |
|      | 30 dB| 0.0797 ± 0.0051 | 0.0624  | 0.1456 ± 0.0164 | 0.0828 ± 0.0614 | 0.2291 ± 0.0510 | 0.0621 ± 0.0146 |
|      | 10 dB| 0.1433 ± 0.0143 | 0.1200  | 0.1256 ± 0.0091 | 0.1326 ± 0.0312 | 0.1980 ± 0.0312 | 0.1126 ± 0.0321 |
| DS2  | 20 dB| 0.1366 ± 0.0106 | 0.1193  | 0.1580 ± 0.0217 | 0.1273 ± 0.0318 | 0.1876 ± 0.0021 | 0.1004 ± 0.0021 |
|      | 30 dB| 0.1344 ± 0.0054 | 0.0942  | 0.1005 ± 0.0032 | 0.0987 ± 0.0021 | 0.1690 ± 0.0410 | 0.0911 ± 0.0039 |

Table 2. SAD results on the synthetic data.

| Data | SNR  | SULRSR-TV | NMF-QMV | MV-NTF | NL-TSUn | ULTRA-V | WSCTF    |
|------|------|-----------|---------|--------|---------|---------|----------|
|      | 10 dB| 0.1115 ± 0.0139 | 0.3624  | 0.5127 ± 0.0164 | 0.1326 ± 0.0502 | 0.1427 ± 0.0216 | 0.0920 ± 0.0031 |
| DS1  | 20 dB| 0.0453 ± 0.0014 | 0.0867  | 0.4023 ± 0.0121 | 0.0952 ± 0.0143 | 0.1214 ± 0.0126 | 0.0434 ± 0.0108 |
|      | 30 dB| 0.0134 ± 0.0004 | 0.0920  | 0.3210 ± 0.0114 | 0.0623 ± 0.2432 | 0.0737 ± 0.0601 | 0.0126 ± 0.0245 |
|      | 10 dB| 0.1977 ± 0.0928 | 0.2325  | 0.5232 ± 0.0088 | 0.4329 ± 0.0621 | 0.3172 ± 0.0241 | 0.1920 ± 0.0314 |
| DS2  | 20 dB| 0.1688 ± 0.0794 | 0.0861  | 0.5161 ± 0.0128 | 0.4061 ± 0.0051 | 0.3202 ± 0.0152 | 0.0827 ± 0.0021 |
|      | 30 dB| 0.1423 ± 0.0366 | 0.0184  | 0.3237 ± 0.0672 | 0.2846 ± 0.0601 | 0.3110 ± 0.0031 | 0.0094 ± 0.0017 |

Figure 2. The abundance maps estimated by the five comparison algorithms and the proposed WSCTF method in DS1 with 20 dB.
4.2. Real Datasets

Secondly, experiments are conducted on three real datasets.

(1) Jasper Ridge Data: Jasper Ridge data is a data set widely used to test the performance of unmixing algorithms, covering the Jasper Ridge Natural Reserve in California, USA. The data contains 224 spectral bands from 300 nm to 2500 nm, and its spectral resolution is 9.46 nm. Figure 1c is a pseudo-color image of Jasper Ridge data. Here we use an image with a size of 100 $\times$ 100 pixels to verify the performance of all algorithms. Because some bands are severely interfered by moisture and noise during the acquisition process, 188 bands are finally used for testing after removing these bands. According to a priori, there are four end members in this area, including soil, water, road and trees.

Table 3 shows the SAD results of the endmembers extracted by the proposed WSCTF compared with ground truth. The bold font indicates the best result. It can be concluded that the WSCTF method proposed in this paper has great advantages in the extraction of two endmembers, and the final SAD result is also optimal. In addition, the endmember spectra extracted by all algorithms are shown in Figure 4. Compared with the reference spectrum, it can be seen that the proposed algorithm has considerable advantages in endmember extraction compared with other methods. Figure 5 is the comparison result of all abundances, which can be concluded that the WSCTF algorithm is the closest to the reference. In addition, compared with the distribution of ground objects in the original image in Figure 1c, the accuracy of the proposed algorithm can also be inferred.
Figure 4. The endmembers of Jasper Ridge Data extracted by the comparison methods and the proposed WSCTF. (a) Tree. (b) Soil. (c) Water. (d) Road.

Figure 5. The abundances of six algorithms and the Ground Truth on Jasper Ridge Data.

Table 3. SAD results on the Jasper ridge data.

|          | SULRSR-TV | NMF-QMV | MV-NTF | NL-TSu | ULTRA-V | WSCTF |
|----------|-----------|---------|--------|--------|---------|-------|
| Tree     | 0.2012 ± 0.0187 | 0.2571  | 0.2416 ± 0.0136 | 0.2081 ± 0.0306 | 0.0502 ± 0.0003 | 0.0495 ± 0.0476 |
| Soil     | 0.2336 ± 0.0011 | 1.1890  | 0.2601 ± 0.0089 | 0.2334 ± 0.0001 | 0.1422 ± 0.0342 | 0.1071 ± 0.1660 |
| Water    | 0.6194 ± 0.1761 | 0.1386  | 0.1586 ± 0.0956 | 0.4929 ± 0.2715 | 0.5641 ± 0.0013 | 0.3977 ± 0.0072 |
| Road     | 0.2943 ± 0.0722 | 0.1840  | 0.4544 ± 0.0706 | 0.3665 ± 0.1179 | 0.0362 ± 0.0033 | 0.2461 ± 0.0303 |
| Mean     | 0.3863 ± 0.0357 | 0.4421  | 0.3024 ± 0.0396 | 0.3723 ± 0.0482 | 0.3002 ± 0.0010 | 0.2406 ± 0.0434 |

(2) Samon Data: According to [43], the samon data set is collected by the Florida Environment Research Institute through the Samon sensor. The data contains 156 bands and the spectral resolution is 3.13 nm. Here we are experimenting on an image with a size of 100 × 100 pixels. There are three types of ground features in this area, including water, soil and trees. Figure 1d is a pseudo-color image of Samon data. The SAD and its variance results extracted by the six methods on the Samon data are introduced in Table 4. Obviously, it can be concluded that all unmixing algorithms, including the WSCTF proposed in this paper, perform slightly worse on the endmember of water. Since the edge of water when it is mixed with the land is not easy to determine, the abundance map of the water body endmember is greatly disturbed. Similar conclusions can also be drawn from the visualization results in Figures 6 and 7. However, it is gratifying that the method proposed in this paper has a great advantage in the separation of the endmember of water compared to other algorithms, which proves the superiority and stability of WSCTF proposed in this paper.
Figure 6. The endmembers of Samon Data extracted by the five comparison methods and the proposed method. (a) Soil. (b) Tree. (c) Water.

Figure 7. The abundance maps of six methods and the Ground Truth on Samon Data.

Table 4. SAD results on the Samson data.

|          | SULRSR-TV | NMF-QMV | MV-NTF | NL-TSu | ULTRA-V | WSCTF |
|----------|-----------|---------|--------|--------|---------|-------|
| Tree     | 0.0247 ± 0.0121 | 0.0337 | 0.0433 ± 0.0129 | 0.0288 ± 0.0161 | 0.0340 ± 0.0272 | 0.0201 ± 0.0024 |
| Soil     | 0.0495 ± 0.0011 | 0.0732 | 0.0953 ± 0.0011 | 0.0496 ± 0.0001 | 0.0566 ± 0.0090 | 0.0685 ± 0.0102 |
| Water    | 0.1299 ± 0.0004 | 1.4831 | 0.2810 ± 0.0050 | 0.1289 ± 0.0003 | 0.2401 ± 0.0221 | 0.0675 ± 0.0021 |
| Mean     | 0.0818 ± 0.0019 | 0.8575 | 0.1733 ± 0.0036 | 0.0825 ± 0.0026 | 0.1438 ± 0.0154 | 0.0611 ± 0.0105 |

(3) Cuprite Data: This data set is also widely used in unmixing task. Its coverage area is a mining area, so there are many highly mixed pixels. Here, a subgraph with a size of 200 × 150 × 188 is intercepted to test the effectiveness of the algorithm, where 188 is the number of bands. Figure 1e is a pseudo-color image of Cuprite data. According to a prior, there are 12 pure mineral endmembers in this area. Unlike the Samon and Jasper Ridge data, the Cuprite data does not provide true values, so the ground truth of endmembers here are set using the method in [34].

Cuprite data mainly contains 12 minerals, namely Alunite, Kaolinite #1, Kaolinite #2, Sphene, Andradite, Nontronite, Muscovite, Pyrope, Chalcedony, Dumortierite, Montmorillonite, and Buddingtonite. Table 3 shows the SAD comparison results of the endmembers extracted by the six algorithms and the reference endmembers. As shown in Table 5, the algorithm proposed in this paper has obvious advantages in each endmember and average endmember results, and there are 7 optimal results. In addition, Figures 8 and 9 are the visualized results of endmembers and abundance obtained after unmixing, respectively. In Figure 8, compared with the other five methods, it is obvious that the endmember spectrum obtained by the WSCTF algorithm is closer to the reference spectrum. The abundance maps are shown in Figure 9. Compared with [34], it can be seen that the proposed WSCTF can also obtain more accurate abundance results.
Table 5. SAD results on the Aviris Cuprite data.

|                | SULRSR-TV | NMF-QMV | MV-NTF | NL-TSu | ULTRA-V | WSCTF |
|----------------|-----------|---------|--------|--------|---------|-------|
| Alunite        | 0.0966 ± 0.0739 | 0.0860 | 0.0748 ± 0.0875 | 0.0783 ± 0.0243 | 0.1897 ± 0.0722 | **0.0735 ± 0.0134** |
| Andradite      | 0.1353 ± 0.0157 | 0.1212 | 0.0661 ± 0.0207 | 0.0958 ± 0.0376 | 0.0971 ± 0.0323 | **0.0631 ± 0.0369** |
| Buddingtonite  | 0.0699 ± 0.0174 | 0.0797 | 0.0718 ± 0.0193 | 0.0982 ± 0.0183 | 0.1097 ± 0.0193 | **0.0669 ± 0.0217** |
| Chalcedony     | 0.1482 ± 0.0230 | 0.1428 | 0.1652 ± 0.0304 | 0.1253 ± 0.0444 | 0.1212 ± 0.0307 | **0.0829 ± 0.0317** |
| Dumortierite   | **0.0855 ± 0.0152** | 0.1108 | 0.0984 ± 0.0208 | 0.1709 ± 0.0240 | 0.0936 ± 0.0158 | 0.0944 ± 0.0258 |
| Kaolinite#1    | **0.0718 ± 0.0155** | 0.0795 | 0.0841 ± 0.0070 | 0.0738 ± 0.0095 | 0.1300 ± 0.0137 | 0.0859 ± 0.0104 |
| Kaolinite#2    | 0.0678 ± 0.0187 | 0.1014 | 0.0786 ± 0.0098 | **0.0625 ± 0.0126** | 0.0391 ± 0.0160 | 0.0638 ± 0.0154 |
| Montmorillonite| 0.0504 ± 0.0132 | 0.1114 | 0.0535 ± 0.0065 | 0.0720 ± 0.0124 | 0.0508 ± 0.0134 | **0.0501 ± 0.0143** |
| Muscovite      | 0.1066 ± 0.0450 | **0.0811** | 0.2518 ± 0.0498 | 0.0925 ± 0.0457 | 0.1257 ± 0.0389 | 0.1719 ± 0.0461 |
| Nontronite     | 0.0730 ± 0.0193 | 0.0715 | 0.1086 ± 0.0111 | 0.0813 ± 0.0156 | 0.0835 ± 0.0246 | **0.0702 ± 0.0271** |
| Pyrope         | 0.1876 ± 0.0521 | 0.1812 | 0.1431 ± 0.0598 | 0.1501 ± 0.0464 | 0.1737 ± 0.0535 | **0.1326 ± 0.0574** |
| Sphene         | **0.0663 ± 0.0564** | 0.1879 | 0.0700 ± 0.0673 | 0.1642 ± 0.0137 | 0.2306 ± 0.0506 | 0.0744 ± 0.0829 |
| Mean           | 0.1042 ± 0.0062 | 0.1190 | 0.1176 ± 0.0135 | 0.1054 ± 0.0066 | 0.1317 ± 0.0089 | **0.0925 ± 0.0134** |

Figure 8. The endmembers spectra of the Cuprite Date set for all tested method, and the black line is ground truth. (a) Montmorillonite. (b) Sphene. (c) Alunite. (d) Buddingtonite. (e) Dumortierite. (f) Andradite. (g) Muscovite. (h) Kaolinite #1. (i) Chalcedony. (j) Pyrope. (k) Nontronite. (l) Kaolinite #2.
Figure 9. Abundance maps of different endmembers using WSCTF on the AVIRIS Cuprite data. (a) Montmorillonite. (b) Sphene. (c) Alunite. (d) Buddingtonite. (e) Dumortierite. (f) Andradite. (g) Muscovite. (h) Kaolinite #1. (i) Chalcedony. (j) Pyrope. (k) Nontronite. (l) Kaolinite #2.

4.3. Parameter Analysis

The WSCTF algorithm proposed in this paper involves two sparse parameters $\lambda$ and $\lambda_s$. Here we take Jasper Ridge data as an example for parameter setting discussion. First, set $\lambda$ as 0.015, and $\lambda_s$ to $\{0, 1 \times 10^{-3}, 2 \times 10^{-3}, 3 \times 10^{-3}, 4 \times 10^{-3}, 5 \times 10^{-3}, 6 \times 10^{-3}, 7 \times 10^{-3}, 8 \times 10^{-3}, 9 \times 10^{-3}, 1 \times 10^{-2}\}$. The setting of 0 can verify the advantages of the regularization proposed by this algorithm. It can be seen from Figure 10a that when $\lambda_s=0.015$, WSCTF achieves the best results on Jasper Ridge data. Then set $\lambda_s$ as 0.015, $\lambda$ to $\{0, 1 \times 10^{-3}, 2 \times 10^{-3}, 3 \times 10^{-3}, 4 \times 10^{-3}, 5 \times 10^{-3}, 6 \times 10^{-3}, 7 \times 10^{-3}, 8 \times 10^{-3}, 9 \times 10^{-3}, 1 \times 10^{-2}\}$. The result is shown in Figure 10b, and when $\lambda = 0.02$, WSCTF achieves the best results on Jasper Ridge data. Similarly, apply the same method to other data sets to ensure the optimal parameters. Finally, the results of the parameters set in this paper are shown in Table 6. It is worth noting that the same parameters are set here for DS1 and DS2 under different noise conditions since the synthetic data is much simpler than the real data.

Figure 10. Parameter analysis on Jasper Ridge data. (a) Parameter $\lambda_s$; (b) Parameter $\lambda$. 
Table 6. Parameter settings.

| Data    | λ_s | λ  |
|---------|-----|----|
| DS1     | 0.015 | 0.010 |
| DS2     | 0.015 | 0.010 |
| Jasper  | 0.02 | 0.015 |
| Samon   | 0.015 | 0.010 |
| Cuprite | 0.010 | 0.010 |

4.4. Complexity Analysis

Here, three real datasets are taken as examples to analyze the time complexity of the proposed WSCTF and other comparison algorithms. All experiments in this paper are performed on Intel (R) Xeon (R) CPU E5-2697 v3 at 2.60 GHz and 64G RAM platform, which can guarantee the fairness of comparison. The running time of all algorithms is presented in Table 7. It can be seen from the table that the running time of MV-NTF will explode as the image size and the number of endmembers increase. The running time of the WSCTF algorithm proposed in this paper is very short, because it does not involve complex tensor operations. In summary, WSCTF also has certain advantages in terms of time complexity compared to other algorithms.

Table 7. Running Time (Seconds) of All Algorithms on Three Real Datasets.

| SULRST-TV | NMF-QMV | MV-NTF | NL-TSUn | ULTRA-V | WSCTF |
|-----------|---------|--------|---------|---------|-------|
| Jasper Ridge Data | 6 | 10 | 170 | 55 | 26 | 5 |
| Samon Data | 5 | 9 | 135 | 29 | 14 | 4 |
| Cuprite Data | 34 | 293 | 904 | 194 | 209 | 17 |

5. Conclusions

In this paper, a regularization for exploring the group sparse structure in difference images of HSI is proposed, which not only maintains the structural features of the raw data space in the abundance tensor, but also keeps the continuity in the spectral direction. First, the raw image can be constrained by the TV regularization to obtain the difference images, which are 3-D tensors. Then, the difference images with group sparse structures are separately reshaped in two spatial dimensions. In particular, the $L_{2,1}$ norm is used to constrain the sparsity, which can mine the global sparse connected structure. In addition, weights are also used to improve the sparsity. The final experiments prove the superiority of the proposed WSCTF compared with other state-of-the-art methods.

Author Contributions: Formal analysis, L.D.; Funding acquisition, X.F.; Methodology, L.H. and L.D.; Project administration, X.F.; Writing—original draft, X.F.; Writing—review & editing, L.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded in part by the Natural Science Basic Research Program of Shaanxi under Grant 2021JQ-193, and in part by the Project funded by China Postdoctoral Science Foundation under Grants 2021M692504 and 2021TQ0259 and in part by the Fundamental Research Funds of the Central Universities of China under Grant JB211907.

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

References

1. Lu, X.; Sun, H.; Zheng, X. A Feature Aggregation Convolutional Neural Network for Remote Sensing Scene Classification. *IEEE Trans. Geosci. Remote Sens.* 2019, 57, 7894–7906. [CrossRef]

2. Hong, D.; Gao, L.; Yao, J.; Zhang, B.; Plaza, A.; Chanussot, J. Graph Convolutional Networks for Hyperspectral Image Classification. *IEEE Trans. Geosci. Remote Sens.* 2020, 58, 8246–8257. [CrossRef]
3. Li, S.; Dian, R.; Fang, L.; Bioucas-Dias, J.M. Fusing hyperspectral and multispectral images via coupled sparse tensor factorization. *IEEE Trans. Image Process.* 2018, 27, 4118–4130. [CrossRef]

4. Lu, X.; Dong, L.; Yuan, Y. Subspace Clustering Constrained Sparse NMF for Hyperspectral Unmixing. *IEEE Trans. Geosci. Remote Sens.* 2020, 58, 3007–3019. [CrossRef]

5. Zhang, B.; Sun, X.; Gao, L.; Yang, L. Endmember Extraction of Hyperspectral Remote Sensing Images Based on the Ant Colony Optimization (ACO) Algorithm. *IEEE Trans. Geosci. Remote Sens.* 2011, 49, 2635–2646. [CrossRef]

6. Kizel, F.; Benediktsson, J.A. Spatially Enhanced Spectral Unmixing Through Data Fusion of Spectral and Visible Images from Different Sensors. *Remote Sens.* 2020, 12, 1255. [CrossRef]

7. Miao, L.; Qi, H. Endmember Extraction From Highly Mixed Data Using Minimum Volume Constrained Nonnegative Matrix Factorization. *IEEE Trans. Geosci. Remote Sens.* 2007, 45, 765–777. [CrossRef]

8. Zhang, Z.; Liao, S.; Zhang, H.; Wang, S.; Wang, Y. Bilateral Filter Regularized L2 Sparse Nonnegative Matrix Factorization for Hyperspectral Unmixing. *Remote Sens.* 2018, 10, 816. [CrossRef]

9. Dong, L.; Yuan, Y.; Lu, X. Spectral-Spatial Joint Sparse NMF for Hyperspectral Unmixing. *IEEE Trans. Geosci. Remote Sens.* 2020, 59, 2391–2402. [CrossRef]

10. Yuan, Y.; Feng, Y.; Lu, X. Projection-Based NMF for Hyperspectral Unmixing. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2015, 8, 2632–2643. [CrossRef]

11. Yuan, B. NMF hyperspectral unmixing algorithm combined with spatial and spectral correlation analysis. *J. Remote Sens.* 2018, 2, 7.

12. Guo, Z.; Wittman, T.; Osher, S. L1 unmixing and its application to hyperspectral image enhancement. In *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XV*; International Society for Optics and Photonics: Orlando, FL, USA, 2009; Volume 7334, p. 73341M.

13. Pauca, V.P.; Piper, J.; Plemmons, R.J. Nonnegative matrix factorization for spectral data analysis. *Linear Algebra Its Appl.* 2006, 416, 29–47. [CrossRef]

14. Xu, Z.; Zhang, H.; Wang, Y.; Chang, X.; Liang, Y. L 1/2 regularization. *Sci. China Inf. Sci.* 2010, 53, 1159–1169. [CrossRef]

15. Xu, Z.; Chang, X.; Xu, F.; Zhang, H. L 1/2 regularization: A thresholding representation theory and a fast solver. *IEEE Trans. Neural Netw. Learn. Syst.* 2012, 23, 1013. [PubMed]

16. Li, M.; Zhu, F.; Guo, A.J.X. A Robust Multilinear Mixing Model with L 2,1 norm for Unmixing Hyperspectral Images. In Proceedings of the 2020 IEEE International Conference on Visual Communications and Image Processing (VCIP), Macau, China, 1–4 December 2020.

17. Salehani, Y.E.; Gazor, S. Smooth and Sparse Regularization for NMF Hyperspectral Unmixing. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2017, 10, 3677–3692. [CrossRef]

18. Yao, J.; Meng, D.; Zhao, Q.; Cao, W.; Xu, Z. Nonconvex-Sparsity and Nonlocal-Smoothness-Based Blind Hyperspectral Unmixing. *IEEE Trans. Image Process.* 2019, 28, 2991–3006. [CrossRef] [PubMed]

19. Lu, X.; Wu, H.; Yuan, Y.; Yan, P.; Li, X. Manifold Regularized Sparse NMF for Hyperspectral Unmixing. *IEEE Trans. Geosci. Remote Sens.* 2013, 51, 2815–2826. [CrossRef]

20. Yang, S.; Zhang, X.; Yao, Y.; Cheng, S.; Jiao, L. Geometric Nonnegative Matrix Factorization (GNMF) for Hyperspectral Unmixing. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2015, 8, 2696–2703. [CrossRef]

21. Guan, N.; Tao, D.; Luo, Z.; Yuan, B. Manifold Regularized Discriminative Nonnegative Matrix Factorization With Fast Gradient Descent. *IEEE Trans. Image Process. Publ. IEEE Signal Process. Soc.* 2011, 20, 2030–2048. [CrossRef]

22. Mei, S.; He, M.; Shen, Z.; Belkacem, B. Neighborhood preserving Nonnegative Matrix Factorization for spectral mixture analysis. In Proceedings of the 2013 IEEE International Geoscience and Remote Sensing Symposium—IGARSS, Melbourne, Australia, 21–26 July 2013; pp. 2573–2576.

23. Xiong, F.; Qian, Y.; Zhou, J.; Tang, Y.Y. Hyperspectral Unmixing via Total Variation Regularized Nonnegative Tensor Factorization. *IEEE Trans. Geosci. Remote Sens.* 2019, 57, 2341–2357. [CrossRef]

24. Li, H.; Feng, R.; Wang, L.; Zhong, Y.; Zhang, L. Superpixel-Based Reweighted Low-Rank and Total Variation Sparse Unmixing for Hyperspectral Remote Sensing Imagery. *IEEE Trans. Geosci. Remote Sens.* 2021, 59, 629–647. [CrossRef]

25. Lu, X.; Wu, H.; Yuan, Y. Double constrained NMF for hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* 2014, 52, 2746–2758. [CrossRef]

26. Qian, Y.; Xiong, F.; Zeng, S.; Zhou, J.; Tang, Y.Y. Matrix-Vector Nonnegative Tensor Factorization for Blind Unmixing of Hyperspectral Imagery. *IEEE Trans. Geosci. Remote Sens.* 2017, 55, 1776–1792. [CrossRef]

27. Zhang, Q.; Wang, H.; Plemmons, R.J.; Pauca, V.P. Spectral unmixing using nonnegative tensor factorization. In Proceedings of the Southeast Regional Conference, Winston-Salem, NC, USA, 23–24 March 2007.

28. Chatzichristos, C.; Kofidis, E.; Morante, M.; Theodoridis, S. Blind fMRI Source Unmixing via Higher-Order Tensor Decompositions. *J. Neurosci. Methods* 2018, 315, 17–47. [CrossRef]

29. Kiliç, L.B.; Pentiuc, S.G. Improving the Analysis of Hyperspectral Images Using Tensor Decomposition. In Proceedings of the 2020 International Conference on Development and Application Systems (DAS), Suceava, Romania, 21–23 May 2020.

30. Imbiriba, T.; Borsoi, R.A.; Bermudez, J.C.M. Low-Rank Tensor Modeling for Hyperspectral Unmixing Accounting for Spectral Variability. *IEEE Trans. Geosci. Remote Sens.* 2020, 58, 1833–1842. [CrossRef]
31. Imbiriba, T.; Borsoi, R.A.; Bermudez, J.C.M. A Low-rank Tensor Regularization Strategy for Hyperspectral Unmixing. In Proceedings of the 2018 IEEE Statistical Signal Processing Workshop (SSP), Freiburg im Breisgau, Germany, 10–13 June 2018; pp. 373–377.

32. Sun, L.; Wu, F.; Zhan, T.; Liu, W.; Wang, J.; Jeon, B. Weighted Nonlocal Low-Rank Tensor Decomposition Method for Sparse Unmixing of Hyperspectral Images. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2020, 13, 1174–1188. [CrossRef]

33. Huang, J.; Huang, T.Z.; Zhao, X.L.; Deng, L.J. Nonlocal Tensor-Based Sparse Hyperspectral Unmixing. IEEE Trans. Geosci. Remote Sens. 2020, 59, 6854–6868. [CrossRef]

34. Yuan, Y.; Dong, L.; Li, X. Hyperspectral Unmixing Using Nonlocal Similarity-Regularized Low-Rank Tensor Factorization. IEEE Trans. Geosci. Remote Sens. 2021, 1–14. [CrossRef]

35. Dobigeon, N.; Moussaoui, S.; Coulon, M.; Tourneret, J.; Hero, A.O. Joint Bayesian Endmember Extraction and Linear Unmixing for Hyperspectral Imagery. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2009, 57, 4355–4368. [CrossRef]

36. Dong, L.; Yuan, Y. Sparse Constrained Low Tensor Rank Representation Framework for Hyperspectral Unmixing. Remote Sens. 2021, 13, 1473. [CrossRef]

37. Wang, Y.; Peng, J.; Zhao, Q.; Leung, Y.; Zhao, X.L.; Meng, D. Hyperspectral Image Restoration Via Total Variation Regularized Low-Rank Tensor Decomposition. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2018, 11, 1227–1243. [CrossRef]

38. Sun, L.; Jeon, B.; Zheng, Y.; Chen, Y. Hyperspectral unmixing based on L1-L2 sparsity and total variation. In Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September 2016; pp. 4349–4353. [CrossRef]

39. Zheng, Y.B.; Huang, T.Z.; Zhao, X.L.; Chen, Y.; He, W. Double-Factor-Regularized Low-Rank Tensor Factorization for Mixed Noise Removal in Hyperspectral Image. IEEE Trans. Geosci. Remote Sens. 2020, 58, 8450–8464. [CrossRef]

40. Liu, G.; Lin, Z.; Yan, S.; Sun, J.; Yu, Y.; Ma, Y. Robust Recovery of Subspace Structures by Low-Rank Representation. IEEE Trans. Pattern Anal. Mach. Intell. 2013, 35, 171–184. [CrossRef] [PubMed]

41. Zhuang, L.; Lin, C.; Figueiredo, M.A.T.; Bioucas-Dias, J.M. Regularization Parameter Selection in Minimum Volume Hyperspectral Unmixing. IEEE Trans. Geosci. Remote Sens. 2019, 57, 9858–9877. [CrossRef]

42. Clark, R.N.; Swayze, G.A.; King, T.V.; Gallagher, A.J.; Calvin, W.M. The US Geological Survey, digital spectral reflectance library: Version 1: 0.2 to 3.0 microns. In Proceedings of the JPL, Summaries of the 4th Annual JPL Airborne Geoscience Workshop, Washington, DC, USA, 25–29 October 1993; Volume I, pp. 11–14.

43. Zheng, P.; Su, H.; Du, Q. Sparse and Low-Rank Constrained Tensor Factorization for Hyperspectral Image Unmixing. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2021, 14, 1754–1767. [CrossRef]