

Article

Non-Negative Matrix Factorization Based on Smoothing and Sparse Constraints for Hyperspectral Unmixing

Xiangxiang Jia and Baofeng Guo *

School of Automation, Hangzhou Dianzi University, Hangzhou 310018, China; jxxhdu@163.com

* Correspondence: gbf@hdu.edu.cn

Abstract: Hyperspectral unmixing (HU) is a technique for estimating a set of pure source signals (end members) and their proportions (abundances) from each pixel of the hyperspectral image. Non-negative matrix factorization (NMF) can decompose the observation matrix into the product of two non-negative matrices simultaneously and can be used in HU. Unfortunately, a limitation of many traditional NMF-based methods, i.e., the non-convexity of the objective function, may lead to a sub-optimal solution. Thus, we put forward a new unmixing method based on NMF under smoothing and sparse constraints to obtain a better solution. First, considering the sparseness of the abundance matrix, a weight sparse regularization is introduced into the NMF model to ensure the sparseness of the abundance matrix. Second, according to the similarity prior of the same feature in the adjacent pixels, a Total Variation regularization is further added to the NMF model to improve the smoothness of the abundance map. Finally, the signatures of each end member are modified smoothly in spectral space. Moreover, it is noticed that discontinuities may emerge due to the removal of noisy bands. Therefore, the spectral data are piecewise smooth in spectral space. Then, in this paper, a piecewise smoothness constraint is further applied to each column of the end-member matrix. Experiments are conducted to evaluate the effectiveness of the proposed method based on two different datasets, including a synthetic dataset and the real-life Cuprite dataset, respectively. Experimental results show that the proposed method outperforms several state-of-the-art HU methods. In the Cuprite hyperspectral dataset, the proposed method's Spectral Angle Distance is 0.1694, compared to the TV-RSNMF method's 0.1703, $L_{1/2}$ NMF method's 0.1925, and VCA-FCLS method's 0.1872.

Keywords: hyperspectral unmixing; non-negative matrix factorization; reweighted sparseness; total variation; piecewise smoothness constraint



Citation: Jia, X.; Guo, B. Non-Negative Matrix Factorization Based on Smoothing and Sparse Constraints for Hyperspectral Unmixing. *Sensors* **2022**, *22*, 5417. <https://doi.org/10.3390/s22145417>

Academic Editors: Nikolaos Doulamis and Marc Brecht

Received: 22 April 2022

Accepted: 18 July 2022

Published: 20 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The hyperspectral sensor simultaneously acquires hundreds of continuous spectral bands in the wavelength, ranging from 0.4 to 2.5 μm . Therefore, the hyperspectral images (HSI) have a high spectral resolution, which promotes the development of quantitative remote sensing applications [1]. However, due to the widespread presence of mixed pixels resulting from the low spatial resolution of the sensor, i.e., a pixel is mixed with several end-member signatures, the accuracy of identification of the images can be affected [2]. Thus, the mixed pixels should be decomposed into a series of end members and their proportional fractions to improve the performance of HSI analysis [3]. HU is a technique to find out the pure spectra (i.e., end members) and their specific percentage (i.e., abundance) for every pixel, which has been extensively studied in recent years to solve the mixed pixel problem.

According to the previous studies, HU algorithms can be divided into two categories. The first group of methods consists of several independent steps, namely end-member estimation and abundance estimation. The end-member estimation algorithms mainly include PPI [4], N-FINDR [5], and VCA [6], which are usually based on the pure pixel assumption, i.e., it is assumed that there exists one pure pixel at least for each sort of

material in HSI. However, in some cases, the assumption is not reliable. To solve this problem, some improved methods have been proposed by related researchers. For example, Miao et al. [7] proposed to find the simplex that can contain all pixels with minimum volume, and the vertices corresponding to the simplex are the desired end members. After the extraction of the end member, it is followed by the abundance estimate. The fully constrained least squares (FCLS) method [8] is the most common method of estimating abundance. Satisfactory results can be obtained using this kind of end-member estimation combined with abundance estimation, but the performance of HU is heavily influenced by the accuracy of end-member estimation.

To avoid the above problems, related researchers have proposed a secondary category of the HU method from a perspective of statistical analysis, i.e., taking HU as a blind source separation problem. Such methods mainly include independent component analysis (ICA) [9] and non-negative matrix factorization (NMF) [10], where the ICA method assumes that the end-member matrix and the abundance matrix are independent of each other. However, the sum of the fractions proportional to each material is 1, so the independence assumption of ICA cannot be guaranteed in HSI. It can only be used as an approximate solution for HU in some cases. Hyperspectral data is decomposed into the end-member matrix and abundance matrix by HU, and the observation matrix is decomposed into a product of two non-negative matrices by the NMF-based method, which better fits the needs of HU and has no need to satisfy the assumption of pure pixels. Therefore, this method is widely used for HU.

Unfortunately, the objective function of the classic NMF method is non-convex, which means the result may fall into local minima. In order to solve this problem, different constraints based on the characteristics of the HSI were introduced. For example, the distribution of end member in HSI is clustered so that most pixels are mixed by some particular end members instead of all of them. Thus, the columns of the abundance matrix are sparse. Accordingly, researchers proposed to use L_1 norm [11] to promote the sparsity of the results. However, L_1 norm cannot enforce further sparsity when the full additivity constraint of the material abundances is used. Then, Qian et al. proposed an HU method based on the $L_{1/2}$ norm [12], which can obtain more accurate results than that of L_1 . However, this article only considers the sparse feature of HSI. For hyperspectral data, due to its low spatial resolution, the abundances of each end member vary smoothly in spatial space. However, discontinuities may occur due to abrupt changes in abundance. Therefore, the spectral data are piecewise smooth in the spatial domain. Hua et al. proposed an adaptive abundance smoothing (AAS) autoencoder network to promote abundance smoothing [13]. Specifically, a multilayer encoder is used to obtain the abundance, and then the input layer is constructed using a single-layer decoder. Finally, $L_{1/2}$ is applied to promote sparsity. In addition, He et al. proposed the use of Total Variation (TV) regularization to increase the piecewise smoothing of abundance [14], where TV regularization can be treated as a denoising process of the abundance map, which can improve the robustness to noise. In addition, the spectrum is smoothly varied in the wavelength space by the high spectral resolution of HSI [15]. However, discontinuities may occur owing to the removal of noisy bands in wavelength space. Therefore, the spectra data are actually piecewise smooth in the spectral domain. Thus, Sen et al. proposed a piecewise smoothness constraint for HU [16]. It is also noticed that a new variant of NMF-based algorithms was proposed in [17], which showed great potential for many image processing applications due to its sparseness constraint and natural incorporation of local information. Apart from the above studies, some new NMF-related hyperspectral unmixing methods have been put forward in recent years, such as the novel low-rank factorization-based methods [18–20], and the VCA-FCLS that is used as an initialization step in [18]. Future work will be carried out to explore the relationship between the novel low-rank tensor and the non-negative matrix factorization and compare their performances. Although the above methods have achieved better results, there is still extra room for improvement.

To further improve the accuracy of HU, piecewise smoothing constraints and sparsity constraints are integrated into the NMF objective function, and an improved HU model is put forward. Specifically, in this paper, the reweighted sparsity and TV norm constraint [14] are proposed to promote the sparsity and smoothness of the abundance map, followed by the end-member smoothing constraint. First, an unconstrained NMF decomposition is performed, and the spectral difference of the adjacent bands is calculated as a measure of smoothness according to the obtained result. In addition, different smoothing weights are assigned according to different smoothness to further reduce the solution space of the NMF-based method, and the optimal solution can be found more accurately. In summary, our main contribution is to extend the reweighted sparsity and the piecewise smoothing constraints in the abundance map, which were introduced in [14], to a more specific scenario, where smoothing on the end-element spectra is also required.

The rest of this paper is organized as follows: In Section 2, we first introduce the linear spectral mixing model, followed by a brief description of the NMF. The sparsity and piecewise smoothing of hyperspectral abundance maps and the piecewise smoothing of end-member spectral are discussed in Section 3. The model proposed in this paper is described at the end of this section. The numerical results for a synthetic dataset and a real dataset are reported in Section 4. Section 5 is the conclusion with suggestions for future work.

2. Linear Spectral Mixture Model

In the linear mixing model, the spectral signal of a pixel is mixed by a set of end-member spectral signatures linearly according to the abundance fractions. It is defined as

$$\mathbf{y} = \mathbf{A}\mathbf{s} + \mathbf{e} \quad (1)$$

where $\mathbf{y} \in \mathbf{R}^{B \times 1}$ denotes a B -dimensional spectral vector. B is the number of spectral bands. $\mathbf{A} \in \mathbf{R}^{B \times K}$ is an end-member matrix, and each column corresponds to an end-member signature. K is the number of the end member, $\mathbf{s} \in \mathbf{R}^{K \times 1}$ is the abundance vector of a pixel, and $\mathbf{e} \in \mathbf{R}^{B \times 1}$ represents the Gaussian noise.

The linear mixing model for mixed pixels in HSI is expressed in the following matrix form

$$\begin{aligned} \mathbf{Y} &= \mathbf{A}\mathbf{S} + \mathbf{E} \\ \text{s.t. } \mathbf{1}_K^T \mathbf{S} &= \mathbf{1}_N^T, \mathbf{A} \geq 0, \mathbf{S} \geq 0 \end{aligned} \quad (2)$$

where $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_N] \in \mathbf{R}^{B \times N}$ is the hyperspectral data with N pixels and B -bands, $\mathbf{S} = [\mathbf{s}_1, \dots, \mathbf{s}_N] \in \mathbf{R}^{K \times N}$ is the abundance matrix, $\mathbf{E} \in \mathbf{R}^{B \times N}$ represents the Gaussian noise, and $\mathbf{1}_K^T = [1, \dots, 1] \in \mathbf{R}^{1 \times K}$ is an all-one vector with size K .

The observation matrix is decomposed into the product of two non-negative matrices by the NMF-based method, and the classical NMF problem is expressed as

$$\mathbf{Y} \approx \mathbf{A}\mathbf{S} \quad (3)$$

The end-member spectral matrix $\mathbf{A} \in \mathbf{R}^{B \times K}$ and the abundance matrix $\mathbf{S} \in \mathbf{R}^{K \times N}$ can be solved by minimizing the difference between \mathbf{Y} and $\mathbf{A}\mathbf{S}$. The objective function of NMF based on the Euclidean distance is defined as follows:

$$\min_{\mathbf{A}, \mathbf{S}} \frac{1}{2} \|\mathbf{Y} - \mathbf{A}\mathbf{S}\|_F^2, \text{ s.t. } \mathbf{A} \geq 0, \mathbf{S} \geq 0 \quad (4)$$

where $\|\cdot\|_F$ is the Frobenius norm of the matrix.

Algorithms such as projective gradient and multiplicative iteration [21] are used to solve the NMF problems, these algorithms minimize the objective function starting from two non-negative matrices and iterate continuously, and the process decreases. Although the minimization problem of Equation (1) is separately convex in \mathbf{A} and \mathbf{S} , it is not simultaneously convex in both matrices. The widely used multiplicative algorithm presented

in [10] is simple to implement and performs well, and can be generated from the traditional gradient descent algorithm. However, Equation (4) has a non-unique solution due to its non-convex for A and S . For example, if the solutions to Equation (4) are A and S , then there exist invertible matrices B such that $A_1 = AB$ and $S_1 = B^{-1}S$ are also solutions of the equation. Therefore, the solutions are not unique, which is the biggest disadvantage of the NMF-based algorithm. The two most commonly used methods to solve this problem are to assign appropriate initial values and add constraints [21].

In this research, the traditional NMF method or the unconstrained NMF is applied, where no abundance and end-element constraints are imposed. The purpose of using the unconstrained NMF method is to consider a scenario when the smoothness of the end-element spectra is unknown. Therefore, an approximation of the smoothness for the end-element spectra can provide guidance or flag to drive the following piecewise smoothness on the end-element spectra. It is noticed that in the proposed method, the solution's uniqueness cannot be guaranteed. However, if the initial spectra are assumed appropriately, the following solution can be improved significantly and is acceptable for many applications.

3. Sparse and Smooth Constrained NMF Method

3.1. The Sparseness of the Abundance

In most cases, any end member does not contribute to all pixels in the scene. Thus, the abundance matrix is sparse. The L_0 norm is applied to the objective function to promote the sparsity of the abundance matrix, but it suffers from the NP-hard problem. L_1 or L_p ($0.5 \leq p < 1$) constraints are proposed by related researchers. Although L_1 regularization is widely used, the constraint of sum 1 is often not satisfied. In the proposed method, the L_1 norm is applied since no significant difference has been found between the choices of L_1 and $L_{1/2}$ norm in our simulations. It may indicate that the selection of a proper norm in spectral unmixing is a dataset-dependent problem.

Reweighted sparse regularization [14] is applied to the NMF model in this paper to promote the sparsity of the abundance matrix. Specifically, the performance of the L_1 minimization framework is improved by weighting the L_1 parameterization and iteratively updating the weights. The weighted L_1 minimization problem can be expressed as follows:

$$\min_s \|\omega * s\|_1 \text{ s.t. } \mathbf{y} = \mathbf{A}s \quad (5)$$

where $*$ represents the element-wise multiplication. $\omega \in \mathbf{R}^{K \times 1}$ is the weight vector. By proving that the weighted L_1 regularizer can obtain a sparser solution than the L_1 regularizer with a suitable vector of weights. However, how to set the weights is a crucial problem. Candes et al. [22] proposed an iterative reweighting algorithm to solve a series of weighted L_1 minimization problems, in which the weights for the next iteration are calculated based on the current abundance matrix, i.e.,

$$W_{i,j}^{(k+1)} = 1./(|S_{i,j}^{(k)}| + \text{eps}) \quad (6)$$

where $S_{i,j}^{(k)}$ denotes the abundance matrix of the k th iteration, and eps is a positive constant, $./$ represents the element-wise division.

3.2. The Smoothness of the Abundance

The rows of the abundance matrix are smooth. This is due to the similar fractional abundances in adjacent pixels of the same end member. However, discontinuities may occur owing to the abrupt changing of end-member abundance in spatial space. Therefore, the spectral data are piecewise smooth in spatial domains. In this paper, Total Variation (TV) regularization is used to facilitate the piecewise smoothing property, and the process can be regarded as an abundance map denoising process. TV regularization was first proposed

by Rudin et al. [23] for solving the grayscale map denoising problem, and for a grayscale image \mathbf{y} of size $m \times n$, the TV parametrization is defined as

$$\|\mathbf{y}\|_{TV} = \sum_{i=1}^{m-1} \sum_{j=1}^{n-1} \{|y_{i,j} - y_{i+1,j}| + |y_{i,j} - y_{i,j+1}|\} + \sum_{i=1}^{m-1} |y_{i,n} - y_{i+1,n}| + \sum_{j=1}^{n-1} |y_{m,j} - y_{m,j+1}| \quad (7)$$

For an HSI, every band of the HSI can be treated as a grayscale image, so the hyper-spectral total variance regularization is defined as [14]

$$\|\mathbf{S}\|_{HTV} = \sum_{j=1}^K \|FS^j\|_{TV} \quad (8)$$

where K is the number of end members in the image and S^j represents the row vector form of the j th band of the HSI. If the number of pixels in the HSI is N , F represents the operator that reshapes the row vector (with a total of N pixels) as a matrix of $m \times n$, i.e., $N = m \times n$.

3.3. The Smoothness of the End member

HSI has high spectral resolution, and the adjacent bands have similar spectral reflection values, so the end-member spectra of HSI have a certain degree of smoothness. However, the removal of the noise band may cause abrupt changes in the reflection values of adjacent bands, so the end-member spectral curve has piecewise smoothness. First of all, this paper obtains the estimated value of the end-member signatures A_{est} through an unconstrained NMF, which is used as a priori information to determine the degree of smoothing of the end-member matrix A . That is, different smoothing levels are assigned according to the differences in the reflectance values of adjacent bands of A_{est} , where the smoothing levels are described by the weight matrix Q , which is defined as

$$Q_{i,j} = e^{-\frac{(A_{est_{i,j}} - A_{est_{i+1,j}})^2}{\sigma}} \quad (9)$$

where $A_{est_{i,j}}$ is the i th row and j th column element of the end-member matrix estimated using the unconstrained NMF algorithm, i.e., the spectral reflection value corresponding to the i th band in the j th end-member signature. σ is the parameter that controls the degree of smoothing, and obviously, the closer $A_{est_{i,j}}$ is to $A_{est_{i+1,j}}$, the larger the weight $Q_{i,j}$ is.

Figure 1 shows the comparison of the end-element signatures extracted by the unconstrained NMF method with the reference spectra. From the figure, it can be seen that the estimated end-member signatures are very similar to the reference signatures and their trends are basically the same. Therefore, the estimated signatures can be used as a priori information to measure the smoothness of the end-member spectra.

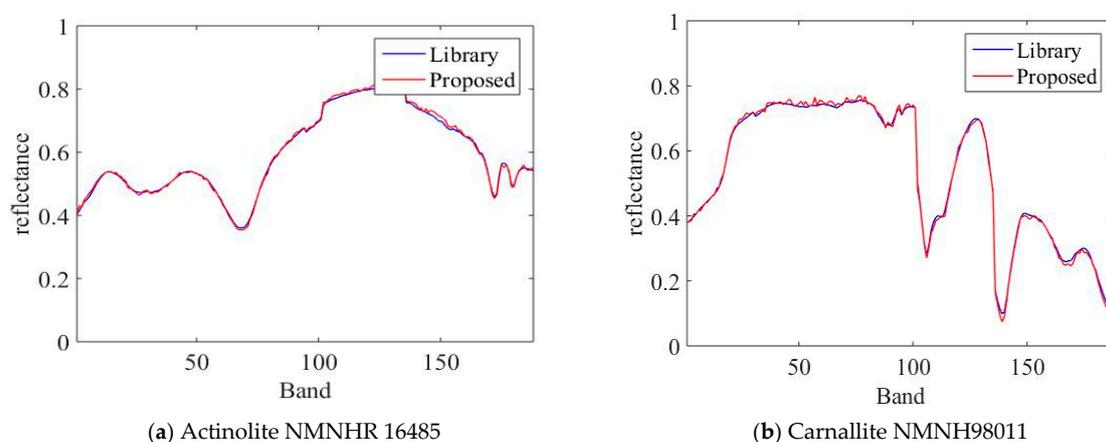


Figure 1. Comparison of the two end-members spectra obtained using NMF estimation in the simulated data with the real end-members spectra.

The end-member spectral piecewise smoothness regularization is defined as

$$\text{minimize } J_1(\mathbf{A}) = \sum_{j=1}^K \sum_{i=1}^{L-1} Q_{i,j} (A_{i,j} - A_{i+1,j})^2 \quad (10)$$

where K is the number of end member, L is the number of spectral bands, and $A_{i,j}$ is the spectral reflection value of the i th band corresponding to the j th end-member signature to be estimated.

3.4. Smoothing and Sparse Constraints-NMF(SSC-NMF) HU Model

According to the above discussion about the characteristics of HSI, the SSC-NMF model proposed in this paper is obtained by integrating the piecewise smoothness constraint and sparsity constraint into the NMF model. The SSC-NMF model is expressed as

$$\begin{aligned} \text{minimize } & J(\mathbf{A}, \mathbf{S}) = \frac{1}{2} \|\mathbf{Y} - \mathbf{A}\mathbf{S}\|_F^2 + \lambda \|\mathbf{W} * \mathbf{S}\|_1 + \tau \|\mathbf{S}\|_{HTV} + \beta J_1(\mathbf{A}) \\ \text{subject to } & \mathbf{A} \geq 0, \mathbf{S} \geq 0, \mathbf{1}_K^T \mathbf{S} = \mathbf{1}_N^T \end{aligned} \quad (11)$$

where the first term is the standard fidelity term, parameter λ controls the sparsity of the abundance matrix, and parameter τ controls the piecewise smoothness of the abundance matrix. The last term promotes the piecewise smoothness of the end-member spectra, and parameter β controls the piecewise smoothness of the end-member matrix.

3.5. Model Optimization

This paper introduces an auxiliary variable \mathbf{L} to make better use of the multiplicative iteration rule. The objective function is then transformed into

$$\begin{aligned} \text{minimize } & J(\mathbf{A}, \mathbf{S}, \mathbf{L}) = \frac{1}{2} \|\mathbf{Y} - \mathbf{A}\mathbf{S}\|_F^2 + \lambda \|\mathbf{W} * \mathbf{S}\|_1 + \tau \|\mathbf{L}\|_{HTV} + \beta J_1(\mathbf{A}) \\ \text{subject to } & \mathbf{A} \geq 0, \mathbf{S} \geq 0, \mathbf{1}_K^T \mathbf{S} = \mathbf{1}_N^T, \mathbf{L} = \mathbf{S} \end{aligned} \quad (12)$$

If we take \mathbf{S} as a noisy version of the auxiliary variable \mathbf{L} , then the constraint $\mathbf{L} = \mathbf{S}$ can be integrated into the objective function thus we obtain the following problem

$$\begin{aligned} \text{minimize } & J(\mathbf{A}, \mathbf{S}, \mathbf{L}) = \frac{1}{2} \|\mathbf{Y} - \mathbf{A}\mathbf{S}\|_F^2 + \lambda \|\mathbf{W} * \mathbf{S}\|_1 + \frac{\mu}{2} \|\mathbf{L} - \mathbf{S}\|_F^2 + \tau \|\mathbf{L}\|_{HTV} + \beta J_1(\mathbf{A}) \\ \text{subject to } & \mathbf{A} \geq 0, \mathbf{S} \geq 0, \mathbf{1}_K^T \mathbf{S} = \mathbf{1}_N^T \end{aligned} \quad (13)$$

where parameter μ is used as the penalty parameter to control the similarity of \mathbf{L} and \mathbf{S} . For the optimization problem of Equation (13), it can be decomposed into three sub-problems to optimize \mathbf{A} , \mathbf{S} , \mathbf{L} respectively.

$$\begin{aligned} \mathbf{A} &= \underset{\mathbf{A}}{\text{argmin}} J(\mathbf{A}, \mathbf{S}, \mathbf{L}) \\ \mathbf{S} &= \underset{\mathbf{S}}{\text{argmin}} J(\mathbf{A}, \mathbf{S}, \mathbf{L}) \\ \mathbf{L} &= \underset{\mathbf{L}}{\text{argmin}} J(\mathbf{A}, \mathbf{S}, \mathbf{L}) \end{aligned} \quad (14)$$

The proposed method consists of three steps: (1) an end-member estimation step, (2) an abundance estimation step, and (3) an abundance denoising step. In each step, the value of one variable is updated according to the current values of the other variables so that the value of the objective function iteratively decreases. The following is a more detailed description.

3.6. Update Rules

3.6.1. End-Members Estimation

The optimization problem of \mathbf{A} can be formulated as

$$\text{minimize } J(\mathbf{A}) = \frac{1}{2} \|\mathbf{Y} - \mathbf{A}\mathbf{S}\|_F^2 + \beta J_1(\mathbf{A}) \quad \text{s.t. } \mathbf{A} \geq 0 \quad (15)$$

In Equation (15), the constrained problem is transformed into an unconstrained problem using the Lagrange multiplier method

$$\text{minimize } J(\mathbf{A}) = \frac{1}{2} \|\mathbf{Y} - \mathbf{A}\mathbf{S}\|_F^2 + \beta J_1(\mathbf{A}) + \text{Tr}(\boldsymbol{\psi} \mathbf{A}^T) \quad (16)$$

where $\boldsymbol{\psi} \in \mathbf{R}^{B \times K}$ is the Lagrange multiplier in matrix format. Based on the Karush–Kuhn–Tucker (KKT) conditions, we obtain the following linear equations:

$$\frac{\partial J(\mathbf{A})}{\partial \mathbf{A}} = \mathbf{A}\mathbf{S}\mathbf{S}^T - \mathbf{Y}\mathbf{S}^T + \beta \frac{\partial J_1(\mathbf{A})}{\partial \mathbf{A}} + \boldsymbol{\psi} = 0 \quad (17)$$

$$\mathbf{A} * \boldsymbol{\psi} = 0 \quad (18)$$

If both sides of Equation (17) are simultaneously multiplied by matrix \mathbf{A} , we can obtain the following equation:

$$\mathbf{A} * (\mathbf{A}\mathbf{S}\mathbf{S}^T) - \mathbf{A} * (\mathbf{Y}\mathbf{S}^T) + \mathbf{A} * \left(\beta \frac{\partial J_1(\mathbf{A})}{\partial \mathbf{A}}\right) + \mathbf{A} * \boldsymbol{\psi} = 0 \quad (19)$$

We obtain the following update rule for \mathbf{A} by combining Equation (18) with Equation (19)

$$\mathbf{A} \leftarrow \mathbf{A} * (\mathbf{Y}\mathbf{S}^T - \beta \frac{\partial J_1(\mathbf{A})}{\partial \mathbf{A}}) ./ \mathbf{A}\mathbf{S}\mathbf{S}^T \quad (20)$$

where the derivative of $J_1(\mathbf{A})$ with respect to each element in \mathbf{A} is

$$\frac{\partial J_1(\mathbf{A})}{\partial A_{i,j}} = 2(A_{i,j} - A_{i+1,j}) * Q_{i,j} \quad (21)$$

The third condition is satisfied in the manuscript by ensuring the non-negativity of \mathbf{A} during initialization. According to Equations (4) and (5), the update rules for SSC-NMF can be formulated as follows:

$$\mathbf{A} \leftarrow \mathbf{A} - \mu_A (\mathbf{A}\mathbf{S}\mathbf{S}^T - \mathbf{Y}\mathbf{S}^T + \beta \frac{\partial J_1(\mathbf{A})}{\partial \mathbf{A}})$$

$$\mathbf{S} \leftarrow \mathbf{S} - \mu_S * (\mathbf{A}^T \mathbf{A} \mathbf{S} - \mathbf{A}^T \mathbf{Y} + \lambda \mathbf{W} + \mu(\mathbf{S} - \mathbf{L}))$$

where μ_A and μ_S are the step sizes. They are set as $\mu_A = \mathbf{A} ./ \mathbf{A}\mathbf{S}\mathbf{S}$ and $\mu_S = \mathbf{S} ./ \mathbf{A}\mathbf{A}\mathbf{S}$ to meet the non-negative constraints. Thus, the update rules can be obtained as follows:

$$\mathbf{A} \leftarrow \mathbf{A} * (\mathbf{Y}\mathbf{S}^T - \beta \frac{\partial J_1(\mathbf{A})}{\partial \mathbf{A}}) ./ \mathbf{A}\mathbf{S}\mathbf{S}^T$$

$$\mathbf{S} \leftarrow \mathbf{S} * (\mathbf{A}^T \mathbf{Y} + \mu \mathbf{L}) ./ (\mathbf{A}^T \mathbf{A} \mathbf{S} + \lambda \mathbf{W} + \mu \mathbf{S})$$

The initialization of \mathbf{A} and \mathbf{S} should be non-negative to ensure their non-negativity during the iteration under the rules presented by Equations (20) and (24). The cost function Equation (15) is non-increasing under the update rules, and it will be convergent to a stationary point [24].

3.6.2. Abundance Estimation

Similarly, the optimization problem of \mathbf{S} can be formulated as

$$\text{minimize } J(\mathbf{S}) = \frac{1}{2} \|\mathbf{Y} - \mathbf{A}\mathbf{S}\|_F^2 + \lambda \|\mathbf{W} * \mathbf{S}\|_1 + \frac{\mu}{2} \|\mathbf{L} - \mathbf{S}\|_F^2 \text{ s.t. } \mathbf{S} \geq 0 \quad (22)$$

The augmentation matrix Y and A in each iteration is given by the following equation to satisfy the abundance sum-to-one constraint (ASC).

$$\bar{Y} = \begin{bmatrix} Y \\ \delta \mathbf{1} \end{bmatrix}, \quad \bar{A} = \begin{bmatrix} A \\ \delta \mathbf{1} \end{bmatrix} \quad (23)$$

where δ controls the effect of ASC constraint, $\mathbf{1}$ denotes the all-1 row vector, Equation (22) is an inequality constrained optimization problem, and the iterative formula for the abundance matrix S is obtained by basing the KKT condition on

$$S \leftarrow S * (\bar{A}^{-T} Y + \mu L) ./ (\bar{A}^{-T} AS + \lambda W + \mu S) \quad (24)$$

For the validity of the update rule shown in Equation (23), please refer to [14].

3.6.3. Abundance Denoising

The optimization problem for L can be formulated as

$$\text{minimize} \quad L = \frac{\mu}{2} \|L - S\|_F^2 + \tau \|L\|_{HTV} \quad (25)$$

Equation (25) is equated to solve the following problem

$$\text{minimize} \quad J(L) = \sum_{j=1}^K \left(\frac{\mu}{2} \|FL^j - FS^j\|_F^2 + \tau \|FL^j\|_{TV} \right) \quad (26)$$

Further translating Equation (26) into K standard TV smoothing problems

$$\hat{L}^j = \min_{L^j} \frac{\mu}{2} \|FL^j - FS^j\|_F^2 + \tau \|FL^j\|_{TV}, j = 1, \dots, K \quad (27)$$

In this paper, the fast gradient projection algorithm (FGP) [25] is used to solve Equation (27), such as the following Algorithm 1.

Algorithm 1 Smoothing and Sparse Constraints NMF for HU

1. **Input:** The observed mixture data $Y \in \mathbf{R}^{K \times N}$, the number of end-members K , the maximum number of iterations c , the parameters $\lambda, \beta, \mu, \tau, \sigma$.
 2. **Output:** End-member signature matrix A and abundance matrix S .
 3. **Initialize** A, S, L , and weighted matrix W .
 4. **Repeat** until convergence:
 5. Update the weight matrix W with Equation (6);
 6. Using Equation (19) to update A ;
 7. Obtain the augmentation matrix of A and S respectively using Equation (23)
 8. Update S by Equation (24);
 9. Update L with Equation (27).
-

4. Experimental Results and Discussion

4.1. Simulated Data Experiments

In the simulated data experiment, four end-member spectra from the United States Geological Survey (USGS) digital spectral library are selected to apply to the simulation data experiment. They are linearly mixed in a certain proportion to obtain simulation data. In this research, the mixture proportions are listed in Table 1:

Table 1. Mixture proportions of the simulated data.

Mixture Proportion (%)	End Member 1 (Carnallite)	End Member 2 (Ammonioalunite)	End Member 3 (Biotite)	End Member 4 (Actinolite)
Case 1	20	20	20	40
Case 2	33	33	33	0
Case 3	0	33	33	33
Case 4	25	25	25	25

As shown in Figure 2a, each spectral curve includes the spectral reflectance corresponding to 224 spectral bands in the wavelength range of 0.4–2.5 μm . By removing the noise and water absorption bands, only the remaining 188 low-noise bands are selected to synthesize the simulated data. After linear mixing, 48×48 pixels are generated, and each pixel has 188 bands. The simulated image is shown in Figure 2b; the pure pixel area is displayed in the first row and areas with different levels of mixing are displayed in 2–4 lines. In addition, the background pixels are also mixed with four types of end members in different proportions, where different colors indicate different degrees of mixing. In Figure 2c–f, it is found that the more yellow color means the higher proportion of a certain end member in the region.

4.1.1. Parameter Selection

The selection of parameter λ is related to the sparsity of the dataset. This paper refers to a method proposed by Qian [12] for estimating the sparse regularization parameter, which is defined as

$$\lambda_e = \frac{1}{\sqrt{L}} \sum_{l=1}^L \frac{\sqrt{N} - \|Y_l\|_1 / \|Y_l\|_2}{\sqrt{N} - 1} \quad (28)$$

where L denotes the number of bands, N denotes the number of pixels, and Y_l denotes the vector corresponding to the l th band.

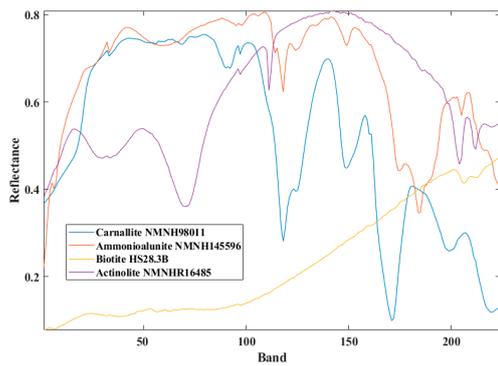
Parameter τ is chosen to be related to the abundance smoothing of the dataset. Therefore, the similarity of the neighboring pixels is evaluated based on the similarity of their spectral values, and then the smoothness of the abundance map is estimated, which is defined as

$$\tau_e = \frac{1}{N} \sum_{i=1}^{N-1} \|x_i - x_{i+1}\|^2 \quad (29)$$

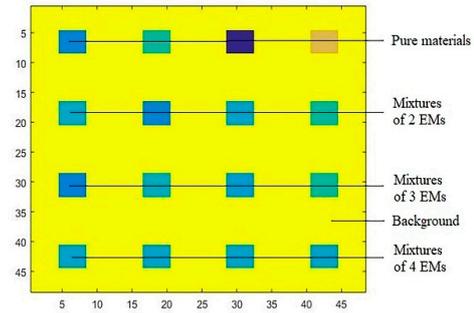
where N denotes the number of pixels, and x_i denotes the spectral vector corresponding to the i th pixel.

When the other parameters are fixed, we discuss the influence of the sparsity parameter λ and the TV regularization parameter τ on the experimental results of the simulation data based on Equations (28) and (29). The results are shown in Figure 3. From this figure, it can be seen that better results can be obtained when λ and τ take smaller values, so choosing the right parameters will have a positive impact.

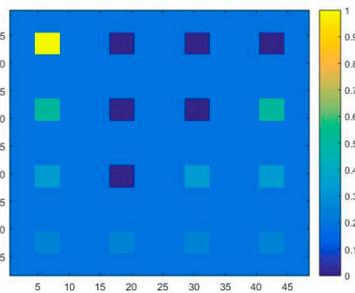
From Figure 3, it can be obtained that the optimal parameter of λ in the proposed method is less than λ_e , and the optimal parameter is within $[\lambda_e/10, \lambda_e]$. The optimal parameter of τ is optimized within the next order of magnitude of $\tau_e/10$ to find the optimal value. Considering both RMSE (Root Mean Square Error) and SAD (Spectral Angle Distance), the λ parameter of the proposed method in this paper is set to 0.005 and τ is set to 0.01 in the simulated data experiments.



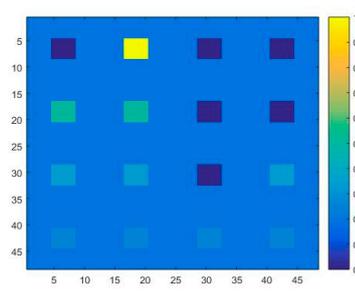
(a) Spectral signatures used to generate the simulated data



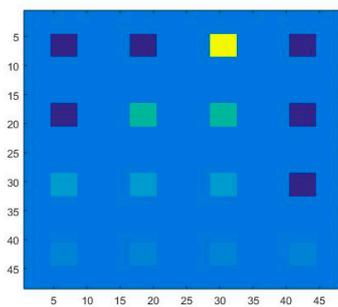
(b) Simulated image



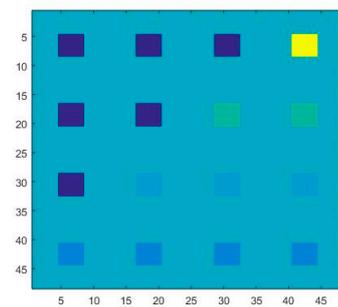
(c) The abundance map of end member 1



(d) The abundance map of end member 2

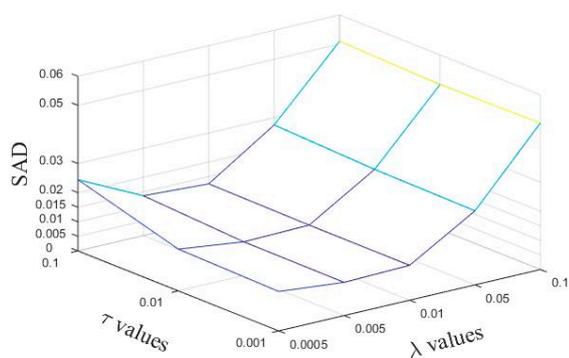


(e) The abundance map of end member 3

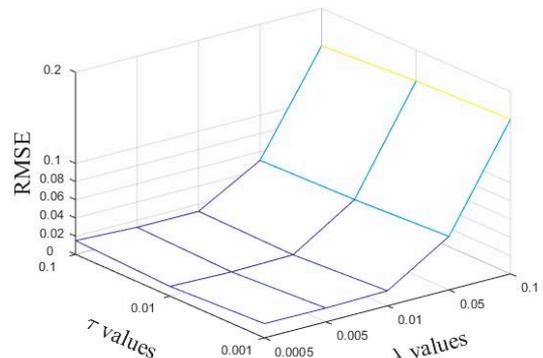


(f) The abundance map of end member 4

Figure 2. Spectral signatures used to generate the simulated data and abundance map of end member, 1–4.



(a)



(b)

Figure 3. Performance of SSC-NMF with respect to parameters λ , τ in terms of (a) SAD (Spectral Angle Distance) and (b) RMSE (Root Mean Square Error).

Secondly, parameter μ is the penalty parameter for violating the linear constraint $L = S$. The larger it is, the closer Equations (12) and (13) are. In this paper, the value of μ varies between 1, 10, 100, and 1000, and the results are shown in Figure 4. RMSE and SAD are two metrics that show the average distance between the predicted values from the model and the actual values in the dataset. The lower the RMSE and SAD, the better an algorithm is able to unmix the data. It can be seen from Figure 4 that RMSE first decreases and then increases when the value of μ becomes larger, and the SAD value basically remains the same in the simulated dataset. However, for each dataset, the choice of μ is different, and the best value is detected by searching for parameters within an order of magnitude above 100.

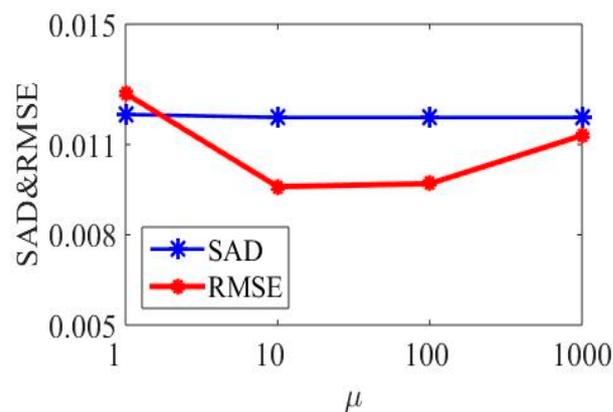


Figure 4. Performance of SSC-NMF with respect to parameter μ in terms of SAD and RMSE.

Finally, the effect of parameter β, σ on the experimental results is discussed, and the results are shown in Figure 5. It can be seen from the figure that SAD and RMSE decrease continuously as β decreases, and finally it reaches a stable state. The value of β is generally taken as around 10 for different datasets due to the different spectral resolutions. For parameter σ , the searching parameter is usually performed by an order of magnitude above and below 0.005. The parameters of other reference methods in this paper are determined according to the analysis in the corresponding references.

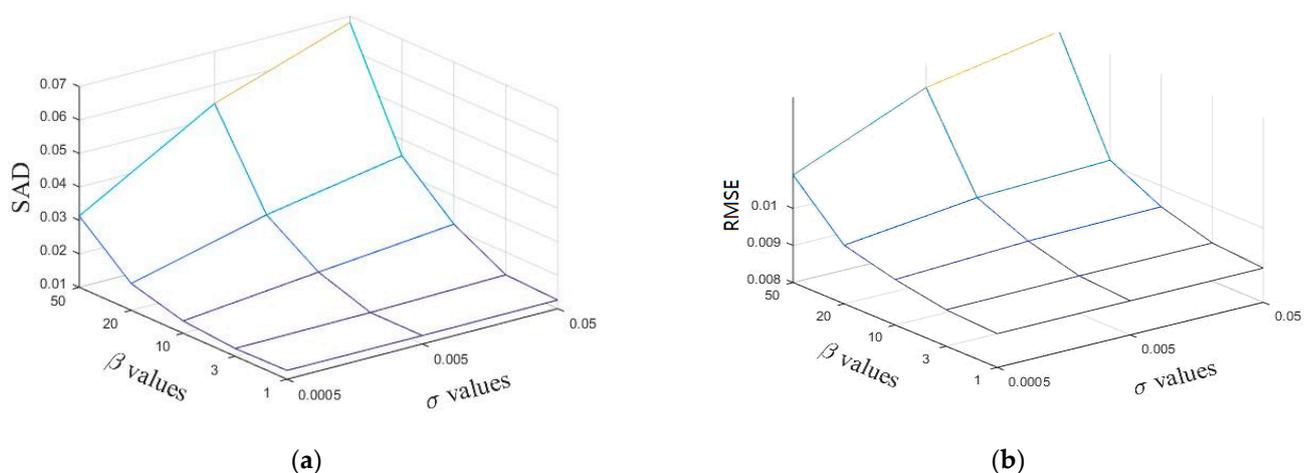


Figure 5. Performance of SSC-NMF with respect to parameters β, σ in terms of (a) SAD and (b) RMSE.

4.1.2. Robustness Analysis

To study the robustness of the proposed method to noise, this paper verifies the performance of the proposed method under different noise levels by adding noise to the simulated data. By setting three different signal-to-noise ratios, three different levels of

noise data are obtained, which are 15, 25, and 35 dB, respectively, where the signal-to-noise ratio is defined as

$$\text{SNR} = 10 \log_{10} \frac{E[x^T x]}{E[n^T n]} \quad (30)$$

where x, n is the observed value of the image and the noise, respectively, and $E[\cdot]$ denotes the expectation operator.

The SAD and RMSE obtained by the four methods with Gaussian noise at different signal-to-noise ratios are listed in Tables 2 and 3, respectively. The benchmarked methods include TV-RSNMF [14], $L_{1/2}$ NMF, and VCA-FCLS, where VCA-FCLS is a supervised-based unmixing method that first uses VCA for end element extraction and then uses the FCLS method for abundance estimation. $L_{1/2}$ NMF and TV-RSNMF methods are unsupervised-based unmixing methods, where the $L_{1/2}$ NMF method enhances the sparsity of the unmixing by imposing sparsity constraints in the NMF objective function, and the TV-RSNMF method considers both the sparsity of abundance and segmental smoothing. When the signal-to-noise ratio is increased to a higher level, such as 35 dB, the majority of the methods, including SSC-NMF, TV-RSNMF, and $L_{1/2}$ NMF, achieved similar results. In this case, the proposed method is not the best one, but the difference is insignificant. This reveals that the proposed method is more suitable for those harder applications where the signal-to-noise ratio is lower than 35 dB.

Table 2. SAD values of different methods with the simulated data.

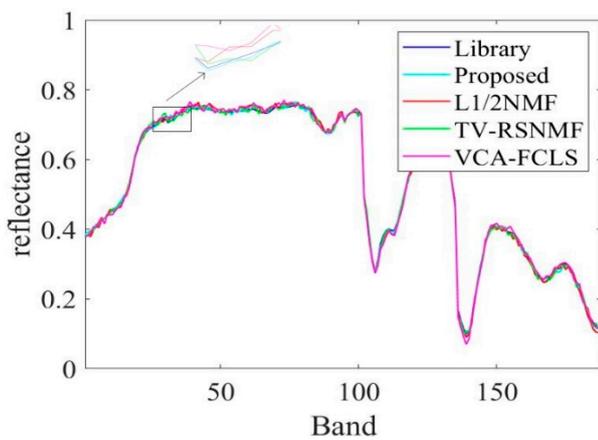
SNR/dB	SSC-NMF	TV-RSNMF	$L_{1/2}$ NMF	VCA-FCLS
15	0.0389	0.0397	0.0440	0.0564
25	0.0116	0.0125	0.0149	0.0167
35	0.0046	0.0047	0.0046	0.0049

Table 3. RMSE values of different methods with the simulated data.

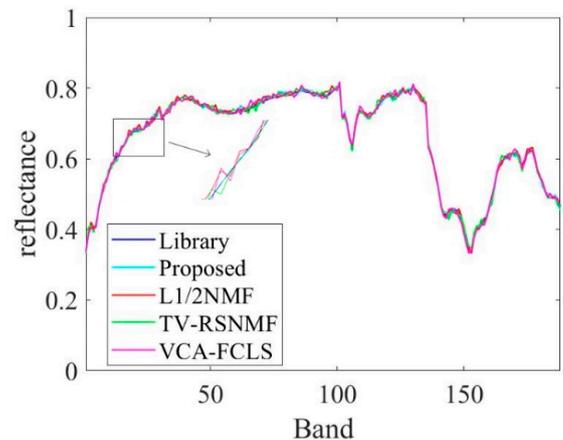
SNR/dB	SSC-NMF	TV-RSNMF	$L_{1/2}$ NMF	VCA-FCLS
15	0.0378	0.0418	0.0492	0.0473
25	0.0092	0.0104	0.0158	0.0171
35	0.0057	0.0058	0.0059	0.0074

For the convenience of comparison, the optimal results in this paper are labeled in bold. It can be seen from Table 2 that the SAD and RMSE values obtained by the proposed method are lower compared to TV-RSNMF [14], $L_{1/2}$ NMF [12], and VCA-FCLS [6], indicating that the algorithm is better at extracting the end-member signatures and inversion abundance. When SNR = 15 dB, the $L_{1/2}$ NMF method obtains the highest RMSE value, mainly due to the fact that the $L_{1/2}$ NMF method cannot handle the low SNR case. The VCA-FCLS method tends to treat noise as an end member when the noise level is high because this method treats monomorphic vertices as an end member, which largely limits the performance of HU. However, the disadvantage of VCA-FCLS is not obvious when the signal-to-noise ratio is low. In addition, the TV-RSNMF method achieves better performance than the $L_{1/2}$ NMF method due to the $L_{1/2}$ NMF method only considering the sparsity constraint of abundance and ignoring the smoothness constraint of abundance, thus indicating the superiority of TV-RSNMF. Further, the end-member piecewise smoothness constraint with TV-RSNMF further improves the accuracy of HU, illustrating the effectiveness of the method proposed in this paper. Figure 6 shows a comparison of the end-member signatures extracted by the different methods with the reference signature. It is seen in Figure 6 that there is a significant deviation between the results obtained by the VCA-FCLS method and the reference value. On the other hand, the end-member signatures extracted by TV-RSNMF and SSC-NMF are in good agreement with the reference signature. Overall, the proposed method achieved a smaller difference with the reference spectrum compared with other benchmarked methods. Compared with different methods for estimating abundance maps

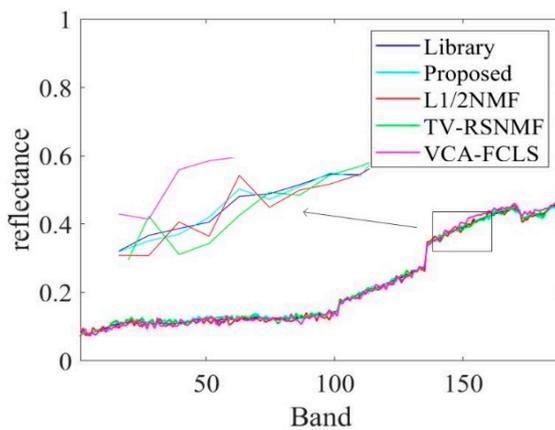
of the fourth end member, the most accurate abundance estimation is obtained from the method of SSC-NMF in Figure 7.



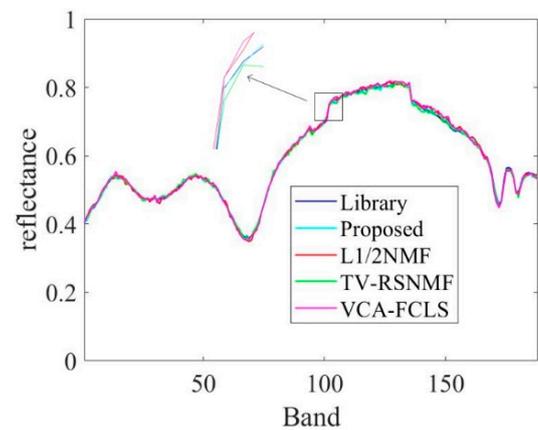
(a) Carnallite NMNH98011



(b) Ammonioalunite NMNH145596



(c) Biotite HS28.3B



(d) Actinolite NMNHR 16485

Figure 6. Comparison of the end-member signatures extracted by the SSC-NMF method with the reference spectra.

4.2. Real Data Experiments

Cuprite dataset: The Cuprite real dataset is used for experiments, which is the benchmark dataset in the HU study. The original cuprite dataset contains 224 bands that cover the wavelength range of 0.37–2.48 μm . In the experiment, noisy bands and water-absorption bands were removed before the unmixing, leaving 188 bands in total. We choose 250×190 regions for the experiment, and there is no accurate answer for the most terminal end-member number due to the strong spectral variability of the AVIRIS Cuprite scene and the presence of a large amount of mineral alteration in the scene. In this paper, the number of end members is set to 12 by referring to other articles [14], where the end-member signatures are shown in Figure 8. It is worth noting that only the end-member signatures are known as a priori for this dataset; the abundance map is unknown.

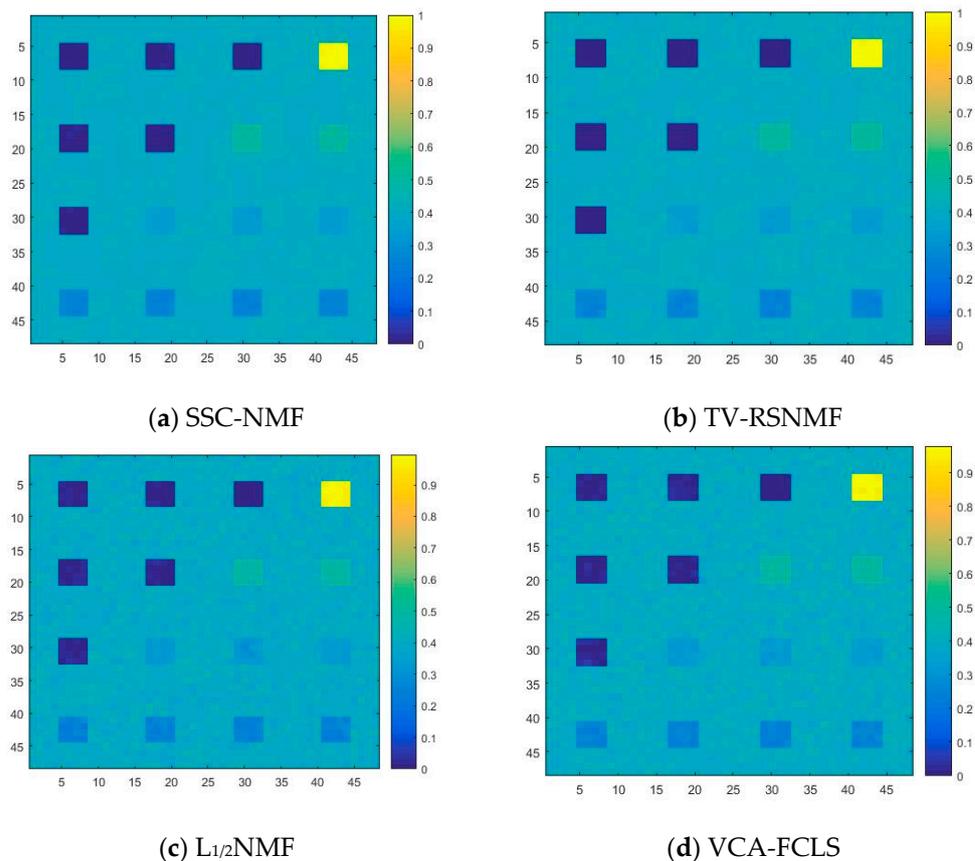
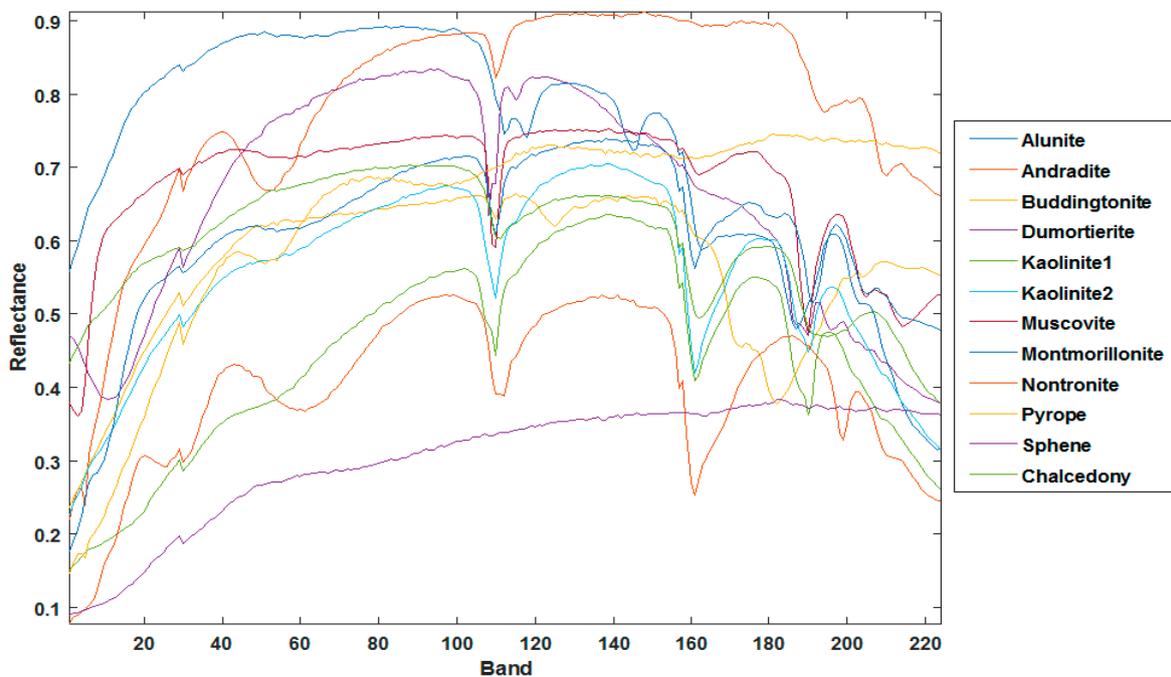


Figure 7. Comparison of abundance maps of the fourth end member estimated by different methods.



Twelve kinds of end-member spectral signatures

Figure 8. Twelve kinds of end-member signature from Cuprite hyperspectral dat. (Courtesy of NASA AVIRIS dataset at <http://aviris.jpl.nasa.gov/> (accessed on 10 December 2021)).

Considering that the spectra in the USGS spectral library are obtained under ideal conditions, and the spectra in real scenes are affected by atmospheric interference and other environmental factors, there are inevitably some differences between the end-member signatures extracted by SSC-NMF and the reference spectra. Table 4 presents the SAD values achieved by the different HU methods, including TV-RSNMF [14], $L_{1/2}$ NMF [12], VCA-FCLS [6], and the recently developed ULTRA-V [19]. From the table, it can be seen that most of the end members obtained by SSC-NMF have optimal or sub-optimal SAD. The average SAD of the proposed method is significantly lower than those of the TV-RSNMF, $L_{1/2}$ NMF [12], VCA-FCLS [6], and is slightly lower than the ULTRA-V. This can indicate the superiority of this algorithm.

Table 4. SAD values of different methods with the real Cuprite dataset.

Method	SSC-NMF	TV-RSNMF	$L_{1/2}$ NMF	VCA-FCLS	ULTRA-V
Alunite	0.1049	0.1064	0.0921	0.0859	0.0842
Andradite	0.0872	0.0878	0.0652	0.0582	0.0511
Buddingtonite	0.0972	0.0964	0.0648	0.0724	0.0571
Dumortierite	0.1086	0.1112	0.0972	0.0978	0.0991
Kaolinite1	0.1316	0.1316	0.1268	0.1222	0.1778
Kaolinite2	0.0450	0.0449	0.0440	0.0458	0.0481
Muscovite	0.1278	0.1279	1.1667	1.1522	0.8819
Montmorillonite	0.0696	0.0698	0.0720	0.0717	0.0919
Nontronite	0.0902	0.0904	0.1173	0.1070	0.1379
Pyrope	0.0879	0.0881	0.1897	0.1783	0.1421
Sphene	1.0130	1.0130	0.0826	0.0876	0.0897
Chalcedony	0.0695	0.0761	0.1919	0.1675	0.1771
Mean	0.1694	0.1703	0.1925	0.1872	0.1698

Figure 9 shows the convergence curve of the algorithm, where the x -axis stands for the iteration number and the y -axis denotes the objective value discussed in Equation (12). As seen in Figure 9, the objective value decreases monotonically with the algorithm iteration. Therefore, it can be expected that the unmixing results will approach true values step by step.

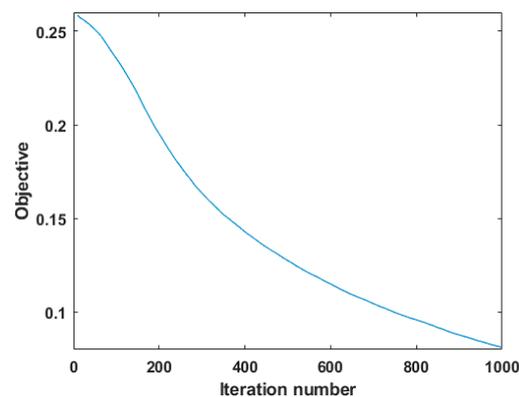


Figure 9. Illustration of convergence of the optimization.

5. Discussion and Conclusions

SSC-NMF, a new HU model, is proposed in this paper to improve the problems of traditional NMF-based models in unmixing applications. This method has the advantages of fully considering the characteristics of HSI, integrating the end-member piecewise smoothness, abundance smoothness and sparsity constraints into the NMF model to further limit the solution space. Especially in the case of a low signal-to-noise ratio, the method proposed in this paper has better robustness, as proven by the simulation experiments. The SSC-NMF model is solved by multiplication and iteration rules. The results from experiments of the

simulated data and Cuprite real data show that this method has improved performance compared with other methods, which demonstrates that the integration of abundance and end-member features in SSC-NMF is effective. However, the method proposed in this paper still has some shortcomings. For example, the computing complexity of the proposed methods is similar to the TV-RSNMF method and $L_{1/2}$ NMF method, but it is relatively higher than the VCA-FCLS method. Moreover, selecting parameters is an important step in this method, which requires a sufficient training dataset and becomes a disadvantage for real-time applications. In this research, the parameters are optimized within a certain range. Therefore, the next focus will be on how to realize the adaptive problem of smoothing parameters. Moreover, only the class labels for each pixel are given in the ground truth associated with the Cuprite dataset, leaving us without the true end-member spectral data and abundance data. In future research, we will test the proposed method on other datasets where the abundance and end-member information are included.

Author Contributions: Conceptualization, X.J. and B.G.; methodology, X.J.; software, X.J.; validation, X.J. and B.G.; writing—original draft preparation, X.J.; supervision, B.G.; project administration, B.G.; funding acquisition, B.G. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been supported by National Natural Science Foundation of China, grant No. 61375011.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: A publicly archived dataset Cuprite can be downloaded from https://www.ehu.es/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes (accessed on 20 December 2021).

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

References

1. Lu, X.; Wang, B.; Zheng, X.; Li, X. Exploring Models and Data for Remote Sensing Image Caption Generation. *IEEE Trans. Geosci. Remote Sens.* **2017**, *56*, 2183–2195. [CrossRef]
2. Bioucas-dias, J.; Plaza, A.; Dobigeon, N.; Parente, M.; Du, Q.; Gader, P.; Chanussot, J. Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2012**, *5*, 354–379. [CrossRef]
3. Dobigeon, N.; Tourneret, J.-Y.; Richard, C.; Bermudez, J.C.M.; McLaughlin, S.; Hero, A.O. Nonlinear Unmixing of Hyperspectral Images: Models and Algorithms. *IEEE Signal Process. Mag.* **2013**, *31*, 82–94. [CrossRef]
4. Wang, W.; Cai, G. Endmember Extraction by Pure Pixel Index Algorithm from Hyperspectral Image. In Proceedings of the 2008 International Conference on Optical Instruments and Technology: Advanced Sensor Technologies and Applications, Beijing, China, 16 November 2008; p. 7157. [CrossRef]
5. Winter, M.E. N-FINDR: An Algorithm for Fast Autonomous Spectral End_Member Determination in Hyperspectral Data. In *Imaging Spectrom V*; International Society for Optics and Photonics: Bellingham, WA, USA, 1999; Volume 3753, pp. 266–275. [CrossRef]
6. Nascimento, J.; Dias, J. Vertex Component Analysis: A Fast Algorithm to Unmix Hyperspectral Data. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 898–910. [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. Heinz, D.; Chang, C.I. Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 529–545. [CrossRef]
9. Jiang, X.; Jiang, Y.; Wu, F.; Wu, F. Quantitative Interpretation of Mineral Hyperspectral Images Based on Principal Component Analysis and Independent Component Analysis Methods. *Appl. Spectrosc.* **2014**, *68*, 502–509. [CrossRef]
10. Lee, D.D.; Seung, H.S. Algorithms for non-negative matrix factorization. In *Advances in Neural Information Processing Systems*; MIT Press: Cambridge, MA, USA, 2001; pp. 556–562.
11. Zare, A.; Gader, P. Hyperspectral Band Selection and Endmember Detection Using Sparsity Promoting Priors. *IEEE Geosci. Remote Sens. Lett.* **2008**, *5*, 256–260. [CrossRef]
12. Qian, Y.; Jia, S.; Zhou, J.; Robleskelly, A. Hyperspectral Unmixing Via $L_{1/2}$ Sparsity-Constrained Nonnegative Matrix Factorization. *IEEE Trans Geosci. Remote Sens.* **2011**, *49*, 4282–4297. [CrossRef]

13. Hua, Z.; Li, X.; Qiu, Q.; Zhao, L. Autoencoder Network for Hyperspectral Unmixing with Adaptive Abundance Smoothing. *IEEE Geosci. Remote Sens. Lett.* **2020**, *18*, 1640–1644. [[CrossRef](#)]
14. He, W.; Zhang, H.; Zhang, L. Total Variation Regularized Reweighted Sparse Nonnegative Matrix Factorization for Hyperspectral Unmixing. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 3909–3921. [[CrossRef](#)]
15. Gu, Y.; Zhang, Y.; Zhang, J. Integration of Spatial–Spectral Information for Resolution Enhancement in Hyperspectral Images. *IEEE Trans. Geosci. Remote Sens.* **2008**, *46*, 1347–1358. [[CrossRef](#)]
16. Jia, S.; Qian, Y.-T.; Ji, Z. Nonnegative Matrix Factorization with Piecewise Smoothness Constraint for Hyperspectral Unmixing. In Proceedings of the International Conference on Wavelet Analysis & Pattern Recognition, Hong Kong, China, 30–31 August 2008. [[CrossRef](#)]
17. Casalino, G.; Gillis, N. Sequential Dimensionality Reduction for Extracting Localized Features. *Pattern Recognit.* **2017**, *63*, 15–29. [[CrossRef](#)]
18. Yuan, Y.; Dong, L.; Li, X. Hyperspectral Unmixing Using Nonlocal Similarity-Regularized Low-Rank Tensor Factorization. *IEEE Trans. Geosci. Remote Sens.* **2021**, *60*, 5507614. [[CrossRef](#)]
19. 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.* **2019**, *58*, 1833–1842. [[CrossRef](#)]
20. 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.* **2020**, *59*, 629–647. [[CrossRef](#)]
21. Huang, R.; Li, X.; Zhao, L. Nonnegative Matrix Factorization with Data-Guided Constraints for Hyperspectral Unmixing. *Remote Sens.* **2017**, *9*, 1074. [[CrossRef](#)]
22. Candes, E.; Wakin, M.; Boyd, S. Enhancing sparsity by reweighted ℓ_1 minimization. *J. Fourier Anal. Appl.* **2008**, *14*, 877–905. [[CrossRef](#)]
23. Rudin, L.I.; Osher, S.; Fatemi, E. Nonlinear total Variation Based Noise Removal Algorithms. *Phys. D Nonlinear Phenom.* **1992**, *60*, 259–268. [[CrossRef](#)]
24. Kimura, T.; Takahashi, N. Gauss-Seidel HALS Algorithm for Nonnegative Matrix Factorization with Sparseness and Smoothness Constraints. *IEICE Trans. Fundam. Electron.* **2017**, *E100-A*, 2925–2935. [[CrossRef](#)]
25. Beck, A.; Teboulle, M. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. *IEEE Trans. Image Process.* **2009**, *18*, 2419–2434. [[CrossRef](#)] [[PubMed](#)]