Double Reweighted Sparse Regression and Graph Regularization for Hyperspectral Unmixing

Si Wang 1,*, Ting-Zhu Huang 2,*, Xi-Le Zhao 2, Gang Liu 3 and Yougan Cheng 4

1 School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China
2 Institute of Computational Science, School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, China; xilezhao@uestc.edu.cn
3 College of Applied Mathematics, Chengdu University of Information Technology, Chengdu 610225, China; liugang@cuit.edu.cn
4 School of Mathematics, University of Minnesota Twin Cities, Minneapolis, MN 55455, USA; yougan.cheng@case.edu

* Correspondence: uuestcswang@uestc.edu.cn (S.W.); tzhuang@uestc.edu.cn (T.-Z.H.)

Abstract: Hyperspectral unmixing, aiming to estimate the fractional abundances of pure spectral signatures in each mixed pixel, has attracted considerable attention in analyzing hyperspectral images. Plenty of sparse unmixing methods have been proposed in the literature that achieved promising performance. However, many of these methods overlook the latent geometrical structure of the hyperspectral data which limit their performance to some extent. To address this issue, a double reweighted sparse and graph regularized unmixing method is proposed in this paper. Specifically, a graph regularizer is employed to capture the correlation information between abundance vectors, which makes use of the property that similar pixels in a spectral neighborhood have higher probability to share similar abundances. In this way, the latent geometrical structure of the hyperspectral data can be transferred to the abundance space. In addition, a double weighted sparse regularizer is used to enhance the sparsity of endmembers and the fractional abundance maps, where one weight is introduced to promote the sparsity of endmembers as a hyperspectral image typically contains fewer endmembers compared to the overcomplete spectral library and the other weight is exploited to improve the sparsity of the abundance matrix. The weights of the double weighted sparse regularizer used for the next iteration are adaptively computed from the current abundance matrix. The experimental results on synthetic and real hyperspectral data demonstrate the superiority of our method compared with some state-of-the-art approaches.

Keywords: graph regularization; sparse; hyperspectral imaging; reweighted; unmixing; abundance

1. Introduction

Hyperspectral imaging has the capacity to record the same scene over many contiguous spectral bands including the visible, nearinfrared and shortwave infrared spectral bands, and it has been widely used in practical applications such as terrain classification, mineral detection and exploration [1,2], military target discrimination, environmental monitoring and pharmaceutical counterfeiting [3]. Mixed pixels are commonly present in hyperspectral data due to the low spatial resolution of the sensors and the complexity of the terrain. The occurrence of mixed pixels severely degrades the application of hyperspectral data [4]. Thus, hyperspectral unmixing (HU) is an important process for hyperspectral data exploitation.

The task of HU is to identify a set of endmembers, and to estimate their corresponding abundances in the formation of each hyperspectral image (HSI) pixel [5–7]. HU algorithms mainly
depend on the expected type of mixing, which can be classified as either a nonlinear or linear model [8]. Nonlinear mixing model assumes that part of the source radiation is affected by multiple scattering effects through several endmembers in the scene before being collected at the sensor. Therefore, the observed mixed pixel spectrum is generated from nonlinear function of the abundance associated with the endmembers. On the other hand, the linear mixture model (LMM) assumes that the mixing only occurs within the measuring instrument itself and interactions among distinct endmembers are negligible. Hence, the observed mixed pixel spectrum can be expressed as a linear combination of a collection of endmember signatures, which is weighted by their corresponding fractional abundances. Even though LMM is not always a precise model to characterize many real scenarios, it is generally recognized as an acceptable approximation of the light scattering and instrument mechanism. It has been widely used for HU due to its simplicity, effectiveness in different applications and computational tractability.

Based on the LMM, many methods have been proposed, such as geometry and statistics based approaches [9]. Geometry based methods usually require that pixel observations of hyperspectral data are within a simplex and their vertices correspond to the endmember [10]. This class of methods includes N-FINDR [11], vertex component analysis [12], minimum-volume simplex analysis [13], simplex growing algorithm [14], etc. In general, these methods require the estimation of the number of the endmembers or the presence of pure pixels. These assumptions may not hold in many hyperspectral data owing to the relatively low spatial resolution of imaging spectrometers and the presence of the mixture phenomena at different scales.

Statistical approaches utilize a statistical characteristic of the hyperspectral data, such as independent component analysis (ICA) [15], and nonnegative matrix factorization (NMF) [16]. NMF can learn a parts-based representation of the data and decompose the high-dimensional data into two nonnegative matrices. Consequently, one can get the endmember signatures and their corresponding abundances simultaneously. This method does not require the pure pixel assumption. Nevertheless, NMF could lead to many local minima owing to the nonconvexity of the objective function. To alleviate this disadvantage, many methods have been proposed by adding some extra constraints under the NMF framework. For example, spatial and spectral regularization is considered in [4,17], the sparsity-constrained of abundances is considered in [18], and the manifold structure of the hyperspectral data is considered in [19,20]. All these methods are unsupervised methods, i.e., both the endmember signatures and the abundances are unknown.

As more spectral libraries become available to the public, many semisupervised methods have been proposed recently. In [21], Bioucas-Dias et al. proposed a sparse unmixing method, named SUnSAL, which is based on variable splitting and augmented Lagrangian. This method assumes that the observed image signatures can be expressed by a linear combination of a small number of pure spectral signature. The high mutual coherence of the hyperspectral library limits the performance of SunSAL. To mitigate the impact of the mutual coherence, Iordache et al. proposed a collaborative sparse regression method for hyperspectral unmixing (CLSUnSAL) [22]. To further improve the performance of the CLSUnSAL, a HU method based on weighted sparse regression with weighted $l_{d,1}$ ($d = 1$ or 2) mixed norm was proposed by Zheng et al. [23]. In [24], Tang et al. assumed that some materials in the spectral library are known to exist in the hyperspectral scene in many situations and they propose the SunSPI algorithm which employs the spectral a priori information. This method improves the performance of SunSAL and CLSUnSAL. However, these methods do not exploit the spatial-contextual information of the hyperspectral images. In [5], Iordache et al. proposed the SUnSAL-TV method, which utilizes the spatial information of the first-order pixel neighborhood system of the abundance map through total variation (TV) on sparse unmixing formulation. In [25], Zhao et al. studied total variation regularization in deblurring and sparse unmixing of hyperspectral images and their method gets better performance than SUnSAL-TV. Taking the non-local spatial information of whole abundance image into account, the non-local mean unmixing method which utilizes the similar patterns and structures in the abundance map is proposed in [26]. Allowing for the
low rankness and sparsity property of abundance matrices, Giampouras et al. proposed a method which combines the weighted $l_1$-norm and the weighted nuclear norm of the abundance matrix as a penalty term [27]. These methods ignore the underlying structure of the hyperspectral images. To take advantage of this property, some graph-based methods are proposed [28,29], which employ the graph topology and sparse group lasso regularization.

To fully utilize the underlying structure of hyperspectral images (HSIs) and sparsity property of abundance matrix, we propose a double reweighted sparse and graph regularized hyperspectral unmixing model. The motivation is based on two priors. First, the similar pixels in a spectral neighborhood are more likely to share the same abundances according to the fundamental spatial–spectral correlation property of HSI. In our method, this latent structure is modeled by a graph regularizer, which groups the highly related neighboring pixels on the graph. In this way, it transfers the structure of hyperspectral data to the unmixing results. Second, a hyperspectral image typically contains fewer endmembers comparing to the spectral library and the abundance maps are naturally sparse. Thus, we employ a double reweighted sparse regularizer to capture this prior. Particularly, we introduce one weight to promote the column sparsity of fractional abundance, while another weight is employed to enhance the sparsity of the abundance along the abundance vector corresponding to each endmember. Moreover, we design a symmetric alternating direction method of multipliers based algorithm to solve the proposed model.

There are some works related to us (e.g., [4,29]). In [4], Wang et al. introduced a double reweighted sparse unmixing and TV (DRSU-TV) algorithm for hyperspectral unmixing. The difference is that they employ the total variation regularization which utilizes local neighbourhood pixels on the abundances rather than graph regularization which can exploit the structures hidden in the hyperspectral data and allow connecting a pixel with as many other pixels in the image as long as they are similar. In [29], the authors incorporated a graph-based total variation framework within the unmixing problem, which takes the strength of graph TV regularization and $l_{2,1}$-norm regularization. Unlike [29], in which the graph TV regularization is imposed on the reconstructed spectra, we exploit the graph Laplacian directly on abundances. In addition, we adopt a double reweighted $l_1$-norm rather than $l_{2,1}$-norm regularization employed in work [29].

The rest of this paper is organized as follows. In Section 2, we briefly review the linear mixture model and symmetric ADMM algorithm for subsequent discussion. The former is the basic model of the linear hyperspectral unmixing and the latter is one of the representative optimization methods for unmixing.

2. Preliminary

In this section, we briefly review the LMM and symmetric ADMM algorithm for subsequent discussion. The former is the basic model of the linear hyperspectral unmixing and the latter is one of the representative optimization methods for unmixing.

2.1. Linear Mixture Model (LMM)

LMM is a popular model used in hyperspectral image unmixing, which assumes that the spectral response of a pixel is a linear combination of spectral signatures (called endmembers) [19,20,30–32]. Given an observed spectrum vector of a mixed pixel $y \in \mathbb{R}^{l \times 1}$, it can be approximated by a nonnegative linear combination of $m$ endmembers, i.e.,

$$y = Ax + e,$$  \hspace{1cm} (1)
where \( A \in \mathbb{R}^{l \times m} \) represents a spectral library containing \( l \) spectral bands and \( m \) spectral signatures, \( x \in \mathbb{R}^{m \times 1} \) is the abundance vector, and \( e \in \mathbb{R}^{l \times 1} \) is the error term (i.e., the noise affecting the measurements). More generally, if the observed data set contains \( n \) mixed pixels organized in a matrix \( Y = [y_1, y_2, \cdots, y_n] \in \mathbb{R}^{l \times n} \), the linear mixture model can be written as

\[
Y = AX + E,
\]

where \( X = [x_1, x_2, \cdots, x_n] \in \mathbb{R}^{m \times n} \) is the abundance fractional matrix (each column \( x_j \in \mathbb{R}^{m \times 1} \) corresponds with the abundance fractions of a mixed pixel \( y_j \in \mathbb{R}^{1 \times l} \)), and \( E = [e_1, e_2, \cdots, e_n] \in \mathbb{R}^{l \times n} \) is the noise matrix. In general, two restrictions are imposed on the fractional abundances \( X \) in LMM due to physical constraints \([5,21,24,26,33]\). On the one hand, we often assume \( X_{ij} \geq 0 \) where \( X_{ij} \) denotes the \((i,j)\)-th entry of \( X \), termed as the abundance non-negative constraint (ANC). On the other hand, it is common to assume \( \mathbf{1}^T x_j = 1 \) where \( \mathbf{1}^T \in \mathbb{R}^{1 \times m} \) is a line vector of 1’s compatible with \( x_j \), termed as the abundance sum-to-one constraint (ASC). Under some circumstances, the sum-to-one constraint cannot be exactly satisfied because of the variability of the signature \([24]\). Hence, we do not consider the sum-to-one constraint.

### 2.2. Symmetric ADMM

The symmetric alternating direction method with multipliers (symmetric ADMM) is an acceleration method of ADMM, which can be used to solve the constraint optimization formulation in image processing \([34–43]\).

Consider the following linearly separable convex minimization problem of the form:

\[
\min \{ \theta_1(x) + \theta_2(y) | Ax + By = b, x \in \Omega_1, y \in \Omega_2 \},
\]

where \( \theta_1 : \mathbb{R}^{n_1} \rightarrow \mathbb{R}, \theta_2 : \mathbb{R}^{n_2} \rightarrow \mathbb{R} \) are convex functions, \( b \in \mathbb{R}^m, \Omega_1 \subset \mathbb{R}^{n_1} \) and \( \Omega_2 \subset \mathbb{R}^{n_2} \) are closed convex sets, \( A \in \mathbb{R}^{m \times n_1} \) and \( B \in \mathbb{R}^{m \times n_2} \) are given matrices. The augmented Lagrangian function of Equation (3) can be written as:

\[
\mathcal{L}_\beta(x, y, \lambda) = \theta_1(x) + \theta_2(y) - \lambda(Ax + By - b) + \frac{\beta}{2} \| Ax + By - b \|^2,
\]

where \( \lambda \) is the Lagrangian multipliers and \( \beta > 0 \) is a penalty parameter. Then, the symmetric ADMM scheme is:

\[
\begin{align*}
    x^{k+1} &= \arg \min_x \{ \mathcal{L}_\beta(x, y^k, \lambda^k) | x \in \Omega_1 \}, \\
    \lambda^{k+\frac{1}{2}} &= \lambda^k - \beta (Ax^{k+1} + By^k - b), \\
    y^{k+1} &= \arg \min_y \{ \mathcal{L}_\beta(x^{k+1}, y, \lambda^{k+\frac{1}{2}}) | y \in \Omega_2 \}, \\
    \lambda^{k+1} &= \lambda^{k+\frac{1}{2}} - \beta (Ax^{k+1} + By^{k+1} - b).
\end{align*}
\]

### 3. Double Reweighted Sparse and Graph Regularized Unmixing Method

In this section, we present the details of the proposed DRSGUH method, including the model formulation and its optimization by symmetric ADMM.

#### 3.1. Double Reweighted \( l_1 \) Sparse Prior

As each pixel can be expressed as the linear combinations of a few number of endmembers, the abundance vector of a pixel in HSI is usually assumed to be sparse. In other words, most of the entries of the abundance matrix are supposed to be zero. Many methods have been proposed to utilize the sparse property of abundances in recent years. A very straightforward method is to introduce \( l_0 \) quasi-norm regularizer to count the nonzero entries in the abundance matrix. However, this is a NP-hard problem. To circumvent this disadvantage, a common surrogate regularizer is \( l_1 \)-norm,
where \( \| \cdot \| \) is a convex substitution for \( l_0 \) regularizer. To further improve the estimation performance over the \( l_1 \)-norm and promote the sparsity of the solution, some other regularizers, such as \( l_p \)-norm regularizer, collaborative sparse regularizer [22], reweighted collaborative sparse regularizer [23] are proposed to solve unmixing problems. In this paper, we propose to use a weighted formulation of \( l_1 \) minimization to improve the estimation performance over the \( l_1 \)-norm, inspired by [4,44,45]. Considering the fact that a hyperspectral image usually includes fewer endmembers compared with the overcomplete spectral library and the abundance map is inherently sparse, we employ the double reweighted \( l_1 \) sparse constraint which defined as

\[
\min_X R_1(X) = \| W_2 \odot W_1 \odot X \|_{1,1}, \quad \text{s.t. } X \geq 0. \tag{6}
\]

Then, we introduce the double reweighted sparse regularizer to unmixing problem, and the new objective function is

\[
(P1): \quad \min_{X \geq 0} \frac{1}{2} \| AX - Y \|_F^2 + \alpha_2 \| W_2 \odot W_1 \odot X \|_{1,1}, \tag{7}
\]

where the operator \( \odot \) denotes the Hadamard product of two variables, and \( W_1 \) and \( W_2 \) are weighted matrices. The first weight \( W_1 \) is introduced to promote the sparsity of endmembers in the spectral library as a hyperspectral image usually contains fewer endmembers compared with the overcomplete spectral library. Due to the unavailability of \( X \), we use the abundance matrix of the current solution to compute the weight matrix \( W_1 \) for the next iteration

\[
(W_1)_i^k = \frac{1}{\| X(i,:) \|_1 + \epsilon}, \tag{8}
\]

where \((W_1)_i^k \in \mathbb{R}^{m \times n}\) denotes the \((i,j)\)-th element of \( W_1 \) estimated in the \( k \)-th iteration and \( \| X(i,:) \|_1 \) denotes the \( l_1 \)-norm of \( X(i,:) \) (the \( i \)-th row of \( X \) estimated in the \( k \) iteration). \( \epsilon > 0 \) is a small positive constant to avoid singularity. Then,

\[
\| W_1 \odot X \|_{1,1} = \sum_{i=1}^m \| X(i,:) \|_1 \approx \| X \|_{1,0},
\]

where \( \| X \|_{1,0} \) is the \( l_0 \) quasi-norm of vector \( \| X(1,:) \|_1, \| X(2,:) \|_1, \ldots, \| X(m,:) \|_1 \), equivalent to the number of nonzero rows of \( X \) [23]. In this way, \( \| W_1 \odot X \|_{1,1} \) has the similar sparsity-promoting property as \( \| X \|_{1,0} \), which is used to measure the number of nonzero row of abundance matrix \( X \). \( \| W_1 \odot X \|_{1,1} \) equally treats all of the entries in each abundance vector. In other words, the same material appears in every pixel. However, the fraction abundance matrix is intrinsically sparse. Hence, we introduce another weight \( W_2 \) to improve sparsity along the abundance vector corresponding to each pixel. Similar to \( W_1 \), we compute \( W_2 \) iteratively, i.e., the iteration weighting coefficients of \( W_2 \) are based on the current estimation of \( X^k \):

\[
(W_2)_i^k = \frac{1}{| X_i^k | + \epsilon}. \tag{9}
\]

It suggests that the small weights of \( W_2 \) encourage nonzero entries in the estimated abundance matrix, and vice versa. It can be seen that when weighted matrices \( W_1 \) and \( W_2 \) are all-ones matrices, formulation (P1) become to sparse unmixing model proposed in [21].

3.2. Graph Regularizer

The sparse constraint only studies sparsity characteristic of the abundance matrix. However, it ignores the latent geometrical structure of the HSI data. As we know, similar pixels in a spatial or spectral neighborhood have higher probability to share similar abundances [46–48]. Therefore, fully making use of this property would enhance the accuracy of the unmixing. We consider the pixel similarity relationship

...
among the neighboring pixels in the spectral space, that is, the abundance vector data set maintains the manifold structure of the spectral feature data set, and incorporates this information into the unmixing problem by the graph Laplacian regularizer. Given a hyperspectral data \( Y \in \mathbb{R}^{l \times n} \) with \( n \) pixels, we view each pixel \( y_i \) as a node, then we find its \( p \) nearest neighbors for each data point \( y_i \) and put a edge between \( y_i \) and its neighbors to construct a weighted graph \( W \) on the \( n \) data points. There are many different methods to define the weighted graph matrix. Here, we exploit heat kernel weighting, i.e., if nodes \( i \) and \( j \) are connected, set

\[
W_{ij} = e^{-\frac{(y_i-y_j)^2}{\sigma^2}},
\]

(10)

where \( \sigma \) is the kernel’s bandwidth \( W \) [48]. Otherwise, if two nodes \( y_i \) and \( y_j \) are not connected, we set the weight \( W_{ij} = 0 \). Obviously, when \( y_i \) and \( y_j \) are very close, the weight \( W_{ij} \) tends to be 1, i.e., a larger \( W_{ij} \) value indicates that pixels \( y_i \) and \( y_j \) are similar.

Through above process, we can get the weight matrix \( W \) of size \( n \times n \) that contains the latent structures in hyperspectral data. We aim to transfer the structure of \( Y \) to the abundance space. As the corresponding abundances of any two points \( y_i \) and \( y_j \) are expected to maintain the same structure, we adopt the following graph Laplacian regularizer:

\[
\min_{X} R_2(X) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} ||x_i - x_j||^2 W_{ij} = \sum_{i=1}^{n} x_i^T x_i D_{ii} - \sum_{i=1}^{n} \sum_{j=1}^{n} x_i^T x_j W_{ij}
\]

(11)

\[
= \text{Tr}(XDX^T) - \text{Tr}(XWX^T) = \text{Tr}(XLX^T), \text{ s.t. } X \geq 0,
\]

where \( \text{Tr}(\cdot) \) represents the trace of a matrix and \( D \) is a diagonal matrix whose entries are column (or row, since \( W \) is symmetric) sums of \( W, D_{ii} = \sum_{j} W_{ij} \). \( L = D - W \) is called graph Laplacian [46]. By minimizing \( R_2(X) \), we expect that if two data points \( y_i \) and \( y_j \) in the HSI data space are close, then \( x_i \) and \( x_j \) in the abundance space also should be close.

3.3. The Double Reweighted Sparse and Graph Regularized Unmixing Model

To simultaneously utilize structure information of HSI data and the sparse property of abundance matrix, we propose the double reweighted sparse and graph regularized unmixing model which can be expressed as follows:

\[
(P2): \quad \min_{X \geq 0} \frac{1}{2} ||AX - Y||_F^2 + \frac{\alpha_1}{2} \text{Tr}(XLX^T) + \alpha_2 ||W_2 \odot W_1 \odot X||_{1,1},
\]

(12)

where \( \alpha_1 > 0 \) and \( \alpha_2 > 0 \) are regularization parameters. The first term is the reconstruction error, the second term is introduced to capture the correlation information between the abundance vectors that is hidden in the HSI data, and the third term is exploited to promote the sparseness of the abundance matrix.

3.4. Optimization Algorithm

Many optimization methods can be applied to solve the proposed formulation \( (P2) \), such as the split Bregman method, alternating direction method with multipliers [35–40,42,43,49,50]. Here, we employ symmetric ADMM to solve \( (P2) \) due to its simplicity and efficiency [34].

The optimization problem \( (P2) \) can be reformulated as

\[
\min_{X,V_1,V_2} \frac{1}{2} ||AX - Y||_F^2 + \frac{\alpha_1}{2} \text{Tr}(V_1 LV_1^T) + \alpha_2 ||W_2 \odot W_1 \odot V_2||_{1,1},
\]

subject to \( V_1 = X, \quad V_2 = X, \quad V_2 \in \Omega := \{ V_2 \in \mathbb{R}^{m \times n}, V_2 \geq 0 \} \).

(13)

Then, Equation (13) can be further expressed as
where \( V = (V_1; V_2) \), \( \theta_1(X) = \frac{1}{2} \| AX - Y \|_F^2 \), \( \theta_2(Z) = \frac{\alpha_2}{2} \text{Tr}(LV_1^T) + \alpha_2 \| W_2 \odot W_1 \odot V_2 \|_{1,1} \), and \( B \) and \( C \) are, respectively, given by

\[
B = \begin{bmatrix} I & 0 \\ I & -I \end{bmatrix}, \quad C = \begin{bmatrix} -I & 0 \\ 0 & -I \end{bmatrix}.
\]

By introducing the Lagrangian multiplier \( \Lambda = (\Lambda_1; \Lambda_2) \in \mathbb{R}^{2m \times n} \), the augmented Lagrangian function of Equation (13) can be written as

\[
L(X, Z, \Lambda) = \theta_1(X) + \theta_2(Z) + \langle \Lambda, BX + CZ \rangle + \frac{\beta}{2} \| BX + CZ \|_F^2
\]

\[
= \frac{1}{2} \| AX - Y \|_F^2 + \frac{\alpha_1}{2} \text{Tr}(LV_1^T) + \alpha_2 \| W_2 \odot W_1 \odot V_2 \|_{1,1} + \frac{\beta}{2} \| X - V_1 - \Lambda_1 \|_F^2 + \frac{\beta}{2} \| X - V_2 - \Lambda_2 \|_F^2
\]

where \( \beta > 0 \) is a penalty parameter. Note that for any fixed value of \( X \), the variables \( V_1 \) and \( V_2 \) are decoupled. Therefore, we can solve \( V_1 \) and \( V_2 \) on their corresponding subproblems separately. In the framework of symmetric ADMM, we use the following iteration scheme to approach the solution of Equation (15):

\[
\begin{align*}
X^{k+1} &\in \arg\min_X \frac{1}{2} \| AX - Y \|_F^2 + \frac{\beta}{2} \| X - V_1^k \|_F^2 + \frac{1}{\beta} \Lambda_1^k \|_F^2 + \frac{\beta}{2} \| X - V_2^k \|_F^2 + \frac{1}{\beta} \Lambda_2^k \|_F^2, \\
(W_1)_i^k &\in \arg\min \frac{1}{\| X^{k+1}(i,:) \|_1 + \epsilon}, \\
(W_2)_i^k &\in \arg\min \frac{1}{|X_i^{k+1}| + \epsilon}, \\
\Lambda_1^{k+1} &\in \Lambda_1^k + r \beta (X^{k+1} - V_1^k), \\
\Lambda_2^{k+1} &\in \Lambda_2^k + r \beta (X^{k+1} - V_2^k), \\
V_1^k &\in \arg\min \frac{\alpha_1}{2} \text{Tr}(LV_1^T) + \frac{\beta}{2} \| X^{k+1} - V_1 \|_F^2 + \frac{1}{\beta} \Lambda_1^{k+1} \|_F^2, \\
V_2^k &\in \arg\min \frac{\alpha_2}{2} \| W_2 + W_1 \odot V_2 \|_{1,1} + \frac{\beta}{2} \| X^{k+1} - V_2 \|_F^2 + \frac{1}{\beta} \Lambda_2^{k+1} \|_F^2.
\end{align*}
\]

For the \( X \)-subproblem, we get that

\[
(A^T A + 2 \beta I) X^{k+1} = A^T Y + \beta (V_1^k + V_2^k) - \Lambda_1^k - \Lambda_2^k.
\]

The solution of \( X^{k+1} \) is explicitly given by:

\[
X^{k+1} = (A^T A + 2 \beta I)^{-1} (A^T Y + \beta (V_1^k + V_2^k) - \Lambda_1^k - \Lambda_2^k).
\]

For the \( V_1 \)-subproblem, we solve the following equation:

\[
V_1^{k+1}(a_1 L + \beta I) = (\beta X^{k+1} + \Lambda_1^{k+\frac{1}{2}}),
\]

the solution of which is
\[ V^{k+1}_1 = (\beta X^{k+1} + A^{k+1}_1)(a_1 L + \beta I)^{-1}. \]  

Denoting \( \tilde{W} = W_2 \odot W_1 \), then the \( V_2 \)-subproblem can be written as

\[ V^{k+1}_2 \in \arg\min_{V_2 \geq 0} a_2 \| \tilde{W}^{k+1}_1 \odot V_2 \|_{1,1} + \frac{\beta}{2} \| X^{k+1} - V_2 + \frac{1}{\beta} A^{k+1}_2 \|_F^2, \]  

As the \( V_2 \) subproblem is component-wise separable, the solution of \( V_2 \) is

\[ (V_2^{k+1})_{ij} = \max \left\{ \left| (X^{k+1} + \frac{1}{\beta} A^{k+1}_2)_{ij} - \frac{a_2}{\beta} \tilde{W}^{k+1}_1, 0 \right| \right\} \circ \text{sign}((X^{k+1} + \frac{1}{\beta} A^{k+1}_2)_{ij}), \]

where \( \tilde{W}^{k+1} = \frac{1}{((X^{k+1} + \beta)_{ii} + \epsilon) ((X^{k+1} + \epsilon))_{ii + \epsilon}} \), “\( \circ \)” represents the point-wise product and \( \text{sign}(x) \) is the signum function.

In summary, the proposed double reweighted sparse graph regularized hyperspectral unmixing algorithm (DRSGHU) is described in Algorithm 1.

**Algorithm 1** Pseudocode of DRSGHU

1. Initialization: \( V^0_1, V^0_2, \Lambda^0_1, \Lambda^0_2 \);
2. set \( k = 0 \), choose \( a_1, a_2, \beta, r, s \);
3. repeat
   \( X^{k+1} = (A^T A + 2 \beta I)^{-1} (A^T Y + \beta (V^k_1 + V^k_2) - \Lambda^k_1 - \Lambda^k_2) \),
   \( \tilde{W}^{k+1}_1 = \frac{1}{((X^{k+1} + r) ||(X^{k+1})||_1 + \epsilon}) \),
   \( \Lambda^{k+1}_1 = \Lambda^k_1 + r \beta (X^{k+1} - V^k_1) \),
   \( \Lambda^{k+1}_2 = \Lambda^k_2 + r \beta (X^{k+1} - V^k_2) \),
   \( V^{k+1}_2 = (\beta X^{k+1} + \Lambda^{k+1}_2) (a_1 L + \beta I)^{-1} \),
   \( (V^{k+1}_2)_{ij} = \max \left\{ \left| ((X^{k+1} + \frac{1}{\beta} \Lambda^{k+1}_2)_{ij}) - \frac{a_2}{\beta} \tilde{W}^{k+1}_1, 0 \right| \right\} \circ \text{sign}(((X^{k+1} + \frac{1}{\beta} \Lambda^{k+1}_2)_{ij}) \),
   \( \Lambda^{k+1}_1 = \Lambda^{k+1}_1 + s \beta (X^{k+1} - V^{k+1}_1) \),
   \( \Lambda^{k+1}_2 = \Lambda^{k+1}_2 + s \beta (X^{k+1} - V^{k+1}_2) \);
4. update iteration \( k \leftarrow k + 1 \);
5. until the stopping criterion is satisfied.

4. Experimental Results

In this section, we test DRSGHU for hyperspectral unmixing using two simulated hyperspectral data and one real data. Both simulated and real hyperspectral data are generated to evaluate the effectiveness of the proposed algorithms. In the synthetic data experiments, we compare our method with some representative methods: (1) SUnSAL [21]; (2) CLSUnSAL [22]; (3) DRSU-TV [4]; (4) GLUP-Lap [28]; (5) Graph-regularized unmixing method (without sparse constraint on abundances), abbreviated as GraphHU; and (6) Sparse graph-regularized unmixing method (set \( W_1 = W_2 \) as all ones matrices in problem (P2)), abbreviated as SGHU. For quantitative comparison, the signal-to-reconstruction error (SRE) is introduced to measure the quality of the estimated fractional abundances of endmembers. SRE is defined as:

\[ \text{SRE} := \frac{\mathbb{E} [|| X_{\text{true}} ||^2_F]}{\mathbb{E} [|| X_{\text{true}} - \hat{X} ||^2_F]}, \]

measured in dB: \( \text{SRE}(\text{dB}) \equiv 10 \log_{10}(\text{SRE}) \), where \( \mathbb{E} \) denotes the expected value, see [22]. Here, \( X_{\text{true}} \) is the true fractional abundance of endmembers and \( \hat{X} \) is the estimated abundance of endmembers. Generally speaking, when the SRE becomes bigger, the estimation approximates the truth better. For initialization, we set \( V^0_1 = 0, V^0_2 = 0, \Lambda^0_1 = 0, \Lambda^0_2 = 0 \), where \( 0 \) is a zero matrix with size \( m \times n \).
The parameters \( r \) and \( s \) are set to be \( r = 0.9 \), and \( s = 0.9 \), the same as [51]. The stopping criteria is set as \(|\|X_k - X_{k-1}\||_F < 10^{-3}\) or the maximum iteration number is reached, which is set to be 500 in our experiments.

To achieve the best performance, we test the proposed method for different combinations of \( \alpha_1, \alpha_2 \) and \( \beta \) according to different noise levels. Specifically, \( \alpha_1 \) is selected from \{0.001, 0.01, 0.05, 0.1, 0.5, 1, 5, 10, 100\}, \( \alpha_2 \) is selected from \{0.01, 0.05, 0.1, 0.5, 1\}, and the penalty parameter \( \beta \) is selected from \{0.01, 0.05, 0.1, 0.5, 1\}. For a fair comparison, we also carefully tune the parameters in all competing methods to get the best possible performance.

All experiments are performed under windows 10 and Matlab R2016b running on a desktop with an Intel(R) Core(TM) i7-7700 CPU at 3.60 GHz and 16.0 GB of memory.

4.1. Experiments on Simulated Data

In the synthetic data experiments, we evaluate the unmixing performance of DRSGHU in situations of different signal-to-noise ratios (SNR \( \equiv 10 \log(\frac{\|AX\|_F^2}{\|N\|_F^2}) \)), different spectral libraries, and different simulated data. The simulated data are contaminated by Gaussian noise with different noise levels of the SNR: 20, 30 and 40 dB. We consider the following two spectral libraries in our simulated data experiments:

- \( A_1 \in \mathbb{R}^{100 \times 120} \) is generated from a library of 262 spectral library signatures generally found on satellites from the National Aeronautics and Space Administration Johnson Space Center (NASA JSC) Spacecraft Materials Spectral Database [52], with 100 spectral bands.
- \( A_2 \in \mathbb{R}^{224 \times 240} \) is generated from a random selection of 240 materials from the U.S. Geological Survey (USGS) digital spectral library (splib06) [24]. The reflectance values are measured for 224 spectral bands uniformly distributed in the interval 0.4–2.5 \( \mu \)m.

In our simulation experiments, we test two different data sets. The details are presented in the following examples.

**Example 1.** In this example, we test the performance on simulated data cube 1 (DC1). This simulated data cube contains 100 \( \times \) 100 pixels generated from nine randomly selected signature from \( A_1 \), and each pixel has 100 bands. Then, a Gaussian white noise with different SNR (20 dB, 30 dB or 40 dB) is added to DC1. The true fractional abundances of nine piecewise smooth with sharp transitions endmembers are shown in Figure 1. Table 1 lists the SRE results for all simulated data sets. From this table, we observe that DRSGHU gets higher SRE than GLUP-Lap, GraphHU, SGHU, SUnSAL for all noise levels. It indicates that both the double reweighted \( l_1 \) sparse prior and graph regularizer mutually enhance the abundance estimation accuracy. As the level of noise decreases, the rate at which DRSGHU improves comparing to SUnSAL increases. This is because that the observed HSI data contains less noise, the graph Laplacian matrix becomes more reliable. In most cases, DRSGHU achieves the best performance and DRSU-TV achieves second best results except when the level of noise is 20 dB. One possible reason is that the graph Laplacian \( L \) is obtained from the very noisy hyperspectral image rather than the clean image, making the graph Laplacian less reliable.

Figure 2 shows the estimated abundance maps by DRSGHU for a noise level of 30 dB. It can be seen that the unmixing results by DRSGHU are very close to the true fractional abundances presented in Figure 1. We further present the relative errors between the estimated fractional abundances and the true fractional abundances of all the endmembers in Table 2 for detailed performance evaluation of our method. In Table 2, the relative errors for the each endmember obtained by DRSGHU is smallest in most cases, and the average relative error of all the endmembers is also smallest. It indicates that the estimation by DRSGHU approximates the truth best.

Furthermore, we present the estimated fractional abundances of all endmembers in the library \( A_1 \) for the first 150 pixels under a noise level of 30dB in Figure 3 to appreciate the sparsity of the fractional abundances. We observe that DRSUTV and DRSGHU produce less outliers than other
methods. As both methods employ the double reweighted sparse prior, they can get relative sparser solution. This suggests that the double weighted matrices promote the sparsity of abundance maps. We also see that the unmixing results by SGHU are much better than SUnSAL. This is because the graph regularizer is imposed onto the sparsity constraint, where the latent structure of HSI data can be utilized. It indicates that incorporating graph regularizer in the unmixing problem improves the abundance estimation accuracy.

Table 1. SRE (dB) values achieved by different methods for noise levels 20 dB, 30 dB and 40 dB on all simulation data.

| Data | SNR | SUnSAL | CLSUnSAL | DRSU-TV | GLUP-Lap | GraphHU | SGHU | DRSGHU |
|------|-----|--------|----------|---------|----------|---------|------|--------|
|      |     |        |          |         |          |         |      |        |
| DC1  | 20  | 3.89   | 4.84     | 10.69   | 5.23     | 4.19    | 5.52 | 7.72   |
|      | 30  | 8.54   | 9.47     | 19.81   | 12.46    | 11.19   | 11.96| 20.26  |
|      | 40  | 14.13  | 15.84    | 28.64   | 19.28    | 16.49   | 17.40| 28.85  |
| DC2  | 20  | 2.98   | 6.20     | 10.72   | 7.58     | 5.75    | 6.82 | 12.27  |
|      | 30  | 6.20   | 7.03     | 25.86   | 18.07    | 11.63   | 20.27| 26.03  |
|      | 40  | 11.30  | 11.94    | 30.42   | 27.58    | 18.95   | 28.99| 49.19  |

Figure 1. True fractional abundances of endmembers of DC1.
Figure 2. The data is contaminated with Gaussian white noise with SNR of 30 dB. Abundances estimated by DRSGHU of DC1 ($\alpha_1 = 0.01$, $\alpha_2 = 1$, $\beta = 0.5$).

Table 2. Relative errors of the estimated fractional abundances of chosen endmembers for a noise level of 30 dB of DC1.

| No.   | SUnSAL  | CLSUnSAL | DRSU-TV | GLU-Lap | GraphHU | SGHU   | DRSGHU |
|-------|---------|----------|---------|---------|---------|--------|--------|
| Endmember1 | 0.2123  | 0.1808   | 0.0488  | 0.0768  | 0.1089  | 0.1084 | 0.0388 |
| Endmember2 | 0.2400  | 0.1205   | 0.1006  | 0.1391  | 0.2306  | 0.1773 | 0.1232 |
| Endmember3 | 0.0505  | 0.0658   | 0.0717  | 0.0861  | 0.1392  | 0.0690 | 0.0702 |
| Endmember4 | 0.2140  | 0.2152   | 0.1533  | 0.2118  | 0.3095  | 0.2631 | 0.1328 |
| Endmember5 | 0.3633  | 0.3972   | 0.0682  | 0.2328  | 0.2168  | 0.2015 | 0.0498 |
| Endmember6 | 0.4986  | 0.4864   | 0.1154  | 0.2341  | 0.3873  | 0.3574 | 0.1079 |
| Endmember7 | 0.4645  | 0.3654   | 0.1390  | 0.2049  | 0.3592  | 0.3574 | 0.1531 |
| Endmember8 | 0.1715  | 0.1115   | 0.0625  | 0.0810  | 0.1042  | 0.1124 | 0.0496 |
| Endmember9 | 0.0797  | 0.0290   | 0.1060  | 0.1086  | 0.1551  | 0.0517 | 0.0940 |
| average  | 0.2549  | 0.2191   | 0.0962  | 0.1528  | 0.2234  | 0.2111 | 0.0819 |
We can observe that, when the noise level is very high (SNR = 20 dB), all methods perform not very good and tend to produce piecewise smooth results. On the contrary, the unmixing results by DRSGHU is visually more satisfactory. Although the unmixing results by DRSU-TV are much better than SUnSAL and CLSUnSAL, some unexpected transitions occurred and some squares are not correctly estimated, especially when the noise level is high as shown in Figure 5. This is possibly because that DRSU-TV links a pixel with only four neighbors.

Example 2. In this example, we test our method on simulated data cube 2 (DC2). This simulated data cube contains 75 × 75 pixels generated from five randomly selected signature from A_2. DC2 has 100 bands per pixel [5]. The true abundances of five endmembers are shown in Figure 4. To illustrate the effectiveness of the proposed method, we present the estimated abundance of endmember 4 (EM4) by different methods in Figure 5. Note that the lines denote the abundance of a certain endmember in the first 150 pixels of the images.

Figure 3. Top panel, from left to right: Ground-truth fractional abundances in DC1 containing 150 pixels and k = 9 endmembers randomly extracted from library A_1, and fractional abundances estimated by SUnSAL, CLSUnSAL, and DRSU-TV. Bottom panel: Fractional abundances estimated by GLUP-Lap, GraphHU, SGHU and DRSGHU.

Figure 4. True fractional abundances of endmembers of DC2.
Figure 5. Abundance maps obtained by different methods for the fourth endmember (EM4) in DC2. From left to right: the unmixing results for the noise levels 20 dB, 30 dB and 40 dB. From top to bottom: The unmixing results by SUnSAL, CLSUnSAL, DRSU-TV, GLUP-Lap, GraphHU, SGHU and DRSGHU.

Table 3. Relative errors of the estimated fractional abundances of EM4 for different noise levels of DC2.

| SNR   | SUnSAL | CLSUnSAL | DRSU-TV | GLUP-Lap | GraphHU | SGHU | DRSGHU |
|-------|--------|----------|---------|----------|---------|------|--------|
| 20 dB | 0.5397 | 0.4010   | 0.1190  | 0.1986   | 0.1153  | 0.1594| 0.0722 |
| 30 dB | 0.2451 | 0.2033   | 0.0453  | 0.0531   | 0.0170  | 0.0446| 0.0133 |
| 40 dB | 0.1214 | 0.0976   | 0.0267  | 0.0152   | 0.0060  | 0.0104| 0.0028 |
4.2. Experiments on Real Data

In this subsection, we use the well-known Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) Cuprite data (http://aviris.jpl.nasa.gov/html/aviris.freedata.html) set in our real data experiments. We use the portion corresponding to a 350 × 350-pixel subset with 188 spectral bands in this experiment (see [5] for more details). A mineral map produced in 1995 by USGS, in which the Tricorder 3.3 software product was used to map different minerals present in the Cuprite mining district (http://speclab.cr.usgs.gov/spectral.lib06). Essential calibration was performed to mitigate the mismatches between the real image spectra and the spectra available in the library before unmixing the AVIRIS Cuprite hyperspectral data [5]. The regularization parameters $\alpha_1$ and $\alpha_2$ are set to be 10, 0.1, and the penalty parameter $\beta$ is set to 0.01 for the proposed method.

The estimated fractional abundance maps for Alunite, Muscovite, and Chalcedony by DRSU-TV, GraphHU, SGHU and DRSGHU are shown in Figure 6. In this figure, we also present the spatial distribution of these materials extracted from the Tricorder software product providing a reference for visual assessment. All compared algorithms produce consistent unmixing results. GraphHU, SGHU and DRSGHU reveal more spatial details for Alunite and Muscovite, and produce fewer outliers for Chalcedony comparing to DRSU-TV.

![Figure 6. Fractional abundance maps estimated for the AVIRIS Cuprite subscene with the USGS library. From top to bottom: Software product, the estimated fractional abundance maps by DRSU-TV, GraphHU, SGHU and DRSGHU.](image-url)
5. Discussion

5.1. The Graph Laplacian Matrix $L$

From our numerical experiments, we notice that it is very important to choose appropriate graph Laplacian matrix $L$ to achieve good performance, as $L$ reflects the similarity relationship among the spectral pixels. There are many methods to construct the graph Laplacian matrix $L$, most of which used the observed data directly to define the graph Laplacian, such as the ones mentioned in Section 2. However, when the observed data are highly corrupted by the noise, the graph Laplacian matrix directly constructed by observed data may not be reliable. Therefore, it is necessary to preprocess the observed data to get better estimation of $L$. For example, some denoising methods can be applied to observed data to get a good graph Laplacian matrix. Table 4 lists unmixing results by DRSGHU after applying the gaussian filter or bilateral filter [53] to the observed data when the SNR = 20 dB. The “reference” column represents the SREs of unmixing results using clean data (not contaminated by the noise) to construct the graph Laplacian matrix. It clearly demonstrates the advantage of pre-denoising before constructing $L$. Some other denoising methods could also be applied to get better performance, such as non-local total variation, BM3D, and sparse representation [2,26].

| Data  | Non-Filter | Gaussian Filter | Bilateral Filter | Reference |
|-------|------------|-----------------|-----------------|-----------|
| DC1   | 7.61       | 10.19           | 9.49            | 11.72     |
| DC2   | 12.27      | 12.84           | 13.00           | 14.38     |

5.2. Parameter Analysis

In this section, we analyze the influence of the parameters $\alpha_1$, $\alpha_2$ and $\beta$ for DRSGHU. For simplicity, we take the simulation data DC1 under different noise levels as an example to illustrate the influence of parameters. We separate these three parameters into two groups, one is the regularization parameters $\alpha_1$ and $\alpha_2$, and another group is penalty parameter $\beta$. To analyze the influence of the regularization parameters, we fix the penalty parameter $\beta$. Figure 7 presents SRE as a function of the regularization parameters of $\alpha_1$ and $\alpha_2$. A general trend observed is that the optimal regularization parameters tend to be bigger when the noise level is higher. In addition, one can see that the surface is not notably steep, which suggests that the algorithm is quite robust around the optimal parameters. To analyze the influence of the penalty parameter, the regularization parameters $\alpha_1$ and $\alpha_2$ are fixed. Figure 8 presents SRE as a function of the penalty parameter $\beta$. It can be observed that the optimal penalty parameter becomes bigger with higher noise level. Moreover, we can see that SREs are relatively stable around the optimal penalty parameter under different noise level, also suggesting the robustness of our algorithm.

Figure 7. SRE (dB) as a function of parameters $\alpha_1$ and $\alpha_2$ for DC1 with different SNR levels. From left to right: SNR = 20 dB, $\beta = 1$; SNR = 30 dB, $\beta = 0.5$; and SNR = 40 dB, $\beta = 0.5$. 
Figure 8. SRE (dB) as a function of parameter $\beta$ for DC1 with different noise levels. From left to right: SNR = 20 dB, $a_1 = 0.05, a_2 = 1$; SNR = 30 dB, $a_1 = 0.01, a_2 = 1$; and SNR = 40 dB, $a_1 = 0.001, a_2 = 0.1$.

5.3. Convergence Analysis

The theoretical convergence analysis for the reweighted norm minimization problem is difficult to establish [44]. Nevertheless, we can observe that the relative error between restored abundance $X$ and the truth abundance $X_{true}$ becomes smaller with the increasing of iteration through the experiments. We take simulated data DC1 as an example to illustrate the convergence of our method. In this test, we set the maximum iteration number as 1000. In Figure 9, one can see that the relative error changes little after 200 iterations, which indicates the stability of our method.

Figure 9. Relative error as a function of iteration number for DC1 with different noise levels. From left to right: SNR = 20 dB, SNR = 30 dB, and SNR = 40 dB.

6. Conclusions

In this paper, we propose a double reweighted sparse and graph regularization based formulation for hyperspectral unmixing. Then, we solve the proposed optimization problem by symmetric alternating direction method with multipliers. The proposed method takes both the structure information of HSI data and sparse property of abundance matrix into consideration. The graph regularization term is introduced to capture the structure information of HSI data, and the double reweighted $l_1$-norm is used to describe the sparsity property of endmembers and fractional abundance maps. The experiments indicate that graph regularizer and weighted sparse regularizer both improve the unmixing performance. It is worth pointing out that the graph Laplacian matrix obtained from the observed data is important to the performance of unmixing result. When the noise level is very high, we could first pre-denoise the hyperspectral data before computing the graph Laplacian matrix to get better performance.

Some issues are worth being investigated in the future. We only consider the Gaussian noise in this work. In real applications, HSIs are usually degraded by different types of noise, such as hybrid Gaussian and stripe noise. We will study how to simultaneously remove these types of noise and unmixing. In addition, we utilize the graph regularizer to capture the structure information of hyperspectral images. Some other regularizers, such as nonlocal TV and low rank tensor, may also be
well-suited for revealing spatial and spectral information of hyperspectral images. It also needs to be studied in further.

**Author Contributions:** All authors contributed to the design of the methodology and the validation of experiments. S.W. performed the numerical simulations, analyzed the results and wrote the manuscript. T.-Z.H., X.-L.Z., G.L. and Y.C. reviewed and revised the paper.

**Funding:** This research was funded by NSFC (61772003) and the Fundamental Research Funds for the Central Universities (ZYGX2016) [132].

**Acknowledgments:** The authors would like to thank the supports by NSFC (61772003) and the Fundamental Research Funds for the Central Universities (ZYGX2016) [132].

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

**References**

1. Manolakis, D.; Siracusa, C.; Shaw, G. Hyperspectral subpixel target detection using the linear mixing model. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 1392–1409. [CrossRef]
2. Zhao, Y.Q.; Yang, J. Hyperspectral image denoising via sparse representation and low-rank constraint. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 296–308. [CrossRef]
3. Li, C.; Sun, T.; Kelly, K.F.; Zhang, Y. A compressive sensing and unmixing scheme for hyperspectral data processing. *IEEE Trans. Image Process.* **2012**, *21*, 1200–1210. [PubMed]
4. Wang, R.; Li, H.C.; Pizurica, A.; Li, J.; Plaza, A.; Emery, W.J. Hyperspectral Unmixing Using Double Reweighted Sparse Regression and Total Variation. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1146–1150. [CrossRef]
5. Iordache, M.D.; Bioucas-Dias, J.M.; Plaza, A. Total variation spatial regularization for sparse hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 4484–4502. [CrossRef]
6. Du, Q.; Raksuntorn, N.; Younan, N.H.; King, R.L. End-member extraction for hyperspectral image analysis. *Appl. Opt.* **2008**, *47*, F77–F84. [CrossRef] [PubMed]
7. Heinz, D.; Chang, C.I. Fully constrained least squares linear mixture analysis for material quantification in hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 529–545. [CrossRef]
8. He, W.; Zhang, H.; Zhang, L. Sparsity-Regularized Robust Non-Negative Matrix Factorization for Hyperspectral Unmixing. *IEEE J. Sel. Top. Appl. Earth Observ.* **2016**, *9*, 4267–4279. [CrossRef]
9. Nascimento, J.M.; Dias, J.M.B. Vertex component analysis: A fast algorithm to unmix hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 898–910. [CrossRef]
10. Liu, X.; Xia, W.; Wang, B.; Zhang, L. An approach based on constrained nonnegative matrix factorization to unmix hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 757–772. [CrossRef]
11. Winter, M.E. N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data. In Proceedings of the SPIE’s International Symposium on Optical Science, Engineering, and Instrumentation, Denver, CO, USA, 27 October 1999; pp. 266–275.
12. Nascimento, J.M.; Dias, J.M.B. Does independent component analysis play a role in unmixing hyperspectral data? *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 175–187. [CrossRef]
13. Li, J.; Bioucas-Dias, J.M. Minimum volume simplex analysis: A fast algorithm to unmix hyperspectral data. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, IGARSS, Boston, MA, USA, 7–11 July 2008; Volume 3, pp. 250–253.
14. Chang, C.I.; Wu, C.C.; Liu, W.m.; Ouyang, Y.C. A new growing method for simplex-based endmember extraction algorithm. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2804–2819. [CrossRef]
15. Hyvärinen, A. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans. Neural Netw.* **1999**, *10*, 626–634. [CrossRef] [PubMed]
16. Lee, D.D.; Seung, H.S. Learning the parts of objects by non-negative matrix factorization. *Nature* **1999**, *401*, 788–791. [PubMed]
17. Huck, A.; Guillaume, M. Robust hyperspectral data unmixing with spatial and spectral regularized NMF. In Proceedings of the 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Reykjavik, Iceland, 14–16 June 2010; pp. 1–4.
18. Qian, Y.; Jia, S.; Zhou, J.; Robles-Kelly, A. Hyperspectral unmixing via sparsity-constrained nonnegative matrix factorization. *IEEE Trans. Geosci. Remote Sens.* 2011, 49, 4282–4297. [CrossRef]

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. Zhu, F.; Wang, Y.; Xiang, S.; Fan, B.; Pan, C. Structured sparse method for hyperspectral unmixing. *ISPRS J. Photogramm.* 2014, 88, 101–118. [CrossRef]

21. Bioucas-Dias, J.M.; Figueiredo, M.A. Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing. In Proceedings of the 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Reykjavik, Iceland, 14–16 June 2010; pp. 1–4.

22. Iordache, M.D.; Bioucas-Dias, J.M.; Plaza, A. Collaborative sparse regression for hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* 2014, 52, 341–354. [CrossRef]

23. Zheng, C.Y.; Li, H.; Wang, Q.; Chen, C.P. Reweighted sparse regression for hyperspectral unmixing. *IEEE Trans. Geosci. Remote Sens.* 2016, 54, 479–488. [CrossRef]

24. Tang, W.; Shi, Z.; Wu, Y.; Zhang, C. Sparse Unmixing of Hyperspectral Data Using Spectral A Priori Information. *IEEE Trans. Geosci. Remote Sens.* 2015, 53, 770–783. [CrossRef]

25. Zhao, X.L.; Wang, F.; Huang, T.Z.; Ng, M.K.; Plemmens, R.J. Deblurring and sparse unmixing for hyperspectral images. *IEEE Trans. Geosci. Remote Sens.* 2013, 51, 4045–4058. [CrossRef]

26. Zhong, Y.; Feng, R.; Zhang, L. Non-local sparse unmixing for hyperspectral remote sensing imagery. *IEEE J. Sel. Top. Appl. Earth Observ.* 2014, 7, 1889–1909. [CrossRef]

27. Giampouras, P.V.; Themelis, K.E.; Rontogiannis, A.A.; Koutroumbas, K.D. Hyperspectral image unmixing via simultaneously sparse and low rank abundance matrix estimation. In Proceedings of the 7th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Tokyo, Japan, 2–5 June 2015; pp. 1–4.

28. Ammanouil, R.; Ferrari, A.; Richard, C. A graph laplacian regularization for hyperspectral data unmixing. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brisbane, Australia, 19–24 April 2015; pp. 1637–1641.

29. Ammanouil, R.; Ferrari, A.; Richard, C. Hyperspectral data unmixing with graph-based regularization. In Proceedings of the 7th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), Tokyo, Japan, 2–5 June 2015; pp. 1–4.

30. Rajabi, R.; Ghassemian, H. Sparsity Constrained Graph Regularized NMF for Spectral Unmixing of Hyperspectral Data. *J. Indian Soc. Remote.* 2015, 43, 269–278. [CrossRef]

31. Bioucas-Dias, J.M.; 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 Observ.* 2012, 5, 354–379. [CrossRef]

32. Heylen, R.; Parente, M.; Gader, P. A review of nonlinear hyperspectral unmixing methods. *IEEE J. Sel. Top. Appl. Earth Observ.* 2014, 7, 1844–1868. [CrossRef]

33. Iordache, M.D.; Bioucas-Dias, J.M.; Plaza, A. Sparse unmixing of hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* 2011, 49, 2014–2039. [CrossRef]

34. He, B.; Ma, F.; Yuan, X. On the Step Size of Symmetric Alternating Directions Method of Multipliers. Available online: http://www.optimization-online.org (accessed on 26 May 2015).

35. Wang, S.; Huang, T.Z.; Zhao, X.L.; Mei, J.J.; Huang, J. Speckle noise removal in ultrasound images by first-and second-order total variation. *Numer. Algorithms* 2018, 78, 513–533. [CrossRef]

36. Liu, J.; Huang, T.Z.; Selesnick, I.W.; Lv, X.G.; Chen, P.Y. Image restoration using total variation with overlapping group sparsity. *Inf. Sci.* 2015, 295, 232–246. [CrossRef]

37. Ji, T.Y.; Huang, T.Z.; Zhao, X.L.; Ma, T.H.; Liu, G. Tensor completion using total variation and low-rank matrix factorization. *Inf. Sci.* 2016, 326, 243–257. [CrossRef]

38. Liu, J.; Huang, T.Z.; Lv, X.G.; Wang, S. High-order total variation-based Poissonian image deconvolution with spatially adapted regularization parameter. *Appl. Math. Modell.* 2017, 45, 516–529. [CrossRef]

39. Mei, J.J.; Dong, Y.; Huang, T.Z.; Yin, W. Cauchy noise removal by nonconvex admm with convergence guarantees. *J. Sci. Comput.* 2018, 74, 743–766. [CrossRef]

40. Huang, Y.M.; Ng, M.K.; Wen, Y.W. A new total variation method for multiplicative noise removal. *SIAM J. Imag. Sci.* 2009, 2, 20–40. [CrossRef]
41. Boyd, S.; Parikh, N.; Chu, E.; Peleato, B.; Eckstein, J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.* 2011, 3, 1–122. [CrossRef]

42. Chen, Y.; Huang, T.Z.; Zhao, X.L.; Deng, L.J.; Huang, J. Stripe noise removal of remote sensing images by total variation regularization and group sparsity constraint. *Remote Sens.* 2017, 9, 559. [CrossRef]

43. Zhao, X.L.; Wang, F.; Ng, M.K. A new convex optimization model for multiplicative noise and blur removal. *SIAM J. Imaging Sci.* 2014, 7, 456–475. [CrossRef]

44. Candes, E.J.; Wakin, M.B.; Boyd, S.P. Enhancing sparsity by reweighted $l_1$ minimization. *J. Fourier Anal. Appl.* 2008, 14, 877–905. [CrossRef]

45. 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]

46. Cai, D.; He, X.; Han, J.; Huang, T.S. Graph regularized nonnegative matrix factorization for data representation. *IEEE Trans. Pattern Anal. Mach. Intell.* 2011, 33, 1548–1560. [PubMed]

47. Chung, F.R. *Spectral Graph Theory*; American Mathematical Soc.: Providence, RI, USA, 1997; Volume 92.

48. Xiang, S.; Nie, F.; Zhang, C. Semi-supervised classification via local spline regression. *IEEE Trans. Pattern Anal. Mach. Intell.* 2010, 32, 2039–2053. [CrossRef] [PubMed]

49. Afonso, M.V.; Bioucas-Dias, J.M.; Figueiredo, M.A. An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems. *IEEE Trans. Image Process.* 2011, 20, 681–695. [CrossRef] [PubMed]

50. He, B.; Liu, H.; Wang, Z.; Yuan, X. A Strictly Contractive Peaceman—Rachford Splitting Method for Convex Programming. *SIAM J. Optim.* 2014, 24, 1011–1040. [CrossRef] [PubMed]

51. Schildknecht, T.; Vannanti, A.; Krag, H.; Erd, C. Reflectance spectra of space debris in geo. In Proceedings of the 2009 AMOS Technical Conference, Maui, HI, USA, 1–4 September 2009.

52. Tomasi, C.; Manduchi, R. Bilateral filtering for gray and color images. In Proceedings of the Sixth International Conference on Computer Vision, Bombay, India, 7 January 1998; pp. 839–846.

© 2018 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 (http://creativecommons.org/licenses/by/4.0/).