Fast Hyperspectral Unmixing via Reweighted Sparse Regression

Hongwei HAN∗,††††, Ke GUO†b, Maozhi WANG†c, Tingbin ZHANG††††d, Nonmembers, and Shuang ZHANG††††e, Member

SUMMARY The sparse unmixing of hyperspectral data has attracted much attention in recent years because it does not need to estimate the number of endmembers nor consider the lack of pure pixels in a given hyperspectral scene. However, the high mutual coherence of spectral libraries strongly affects the practicality of sparse unmixing. The collaborative sparse unmixing via variable splitting and augmented Lagrangian (CLSUnSAL) algorithm is a classic sparse unmixing algorithm that performs better than other sparse unmixing methods. In this paper, we propose a CLSUnSAL-based hyperspectral unmixing method based on dictionary pruning and reweighted sparse regression. First, the algorithm identifies a subset of the original library elements using a dictionary pruning strategy. Second, we present a weighted sparse regression algorithm based on CLSUnSAL to further enhance the sparsity of endmember spectra in a given library. Third, we apply the weighted sparse regression algorithm on the pruned spectral library. The effectiveness of the proposed algorithm is demonstrated on both simulated and real hyperspectral datasets. For simulated data cubes (DC1, DC2 and DC3), the number of the pruned spectral library elements is reduced by at least 94% and the runtime of the proposed algorithm is less than 10% of that of CLSUnSAL. For simulated DC4 and DC5, the runtime of the proposed algorithm is less than 15% of that of CLSUnSAL. For the real hyperspectral datasets, the pruned spectral library successfully reduces the original dictionary size by 76% and the runtime of the proposed algorithm is 11.21% of that of CLSUnSAL. These experimental results show that our proposed algorithm not only substantially improves the accuracy of unmixing solutions but is also much faster than some other state-of-the-art sparse unmixing algorithms.

key words: sparse unmixing, dictionary pruning, hyperspectral imaging, iterative reweighting

1. Introduction

With the development of imaging spectroscopy, hyperspectral remote sensors have evolved to collect multiple images of a scene using a range of spectra from ultraviolet and visible to infrared [1]–[4]. Hyperspectral imaging has increased significantly over the past decade, mainly because of a large amount of reference information, including the complete spectra of ground objects, which enables precise material identification using spectroscopic analysis [5]–[7]. There is a wide range of hyperspectral imaging applications such as terrain classification [8], remote surveillance [2], mineral detection and exploration, environmental monitoring, and military surveillance [9], and pharmaceutical process monitoring and quality control [10]. The main advantage of using hyperspectral imagery is that the spectral signature of each pixel can help identify the materials in the scene [11]. However, mixed pixels often occur in real hyperspectral data because of the relatively low spatial resolution of the sensor and the varying ground surface [12], [13]. Thus, to make full use of these data, the spectral unmixing technique, which decomposes a mixed pixel into a collection of constituent spectra (called endmembers) and their corresponding fractional abundances, has become an essential procedure [12], [14]–[16]. Two basic models are used to analyze the mixed-pixel problem: the linear mixture model (LMM) [17]–[19] and the nonlinear mixture model [20]–[22]. Compared with the nonlinear mixture model, LMM has been widely employed for many different applications because of its computational tractability and flexibility [23]. Despite the fact that the LMM is not always accurate, especially under certain scenarios that exhibit strong nonlinearity, it is generally recognized as an acceptable model for many real-world scenarios [24]. In the past few decades, a large number of algorithms have been proposed for hyperspectral unmixing, most of which are based on the LMM [12], [25].

In a linear spectral unmixing scenario, a mixed pixel is modeled by a linear combination of endmember signatures weighted by the corresponding abundances [5], [26]. In general, conventional algorithms for spectral unmixing with LMM involves two steps: endmember extraction and abundance estimation [12], [22]. A variety of endmember identification and abundance algorithms have been proposed [10]. However, these methods assume that pure pixels can be found in the original hyperspectral image, which is a very difficult task [13]. The sparse unmixing approach, which aims to find the optimal subset of endmembers from a given spectral library to model each pixel in the hyperspectral scene, has recently been introduced [4], [27]–[30]. This sparse unmixing formulation sidesteps two common limitations of classic spectral unmixing approaches, namely, the
lack of pure pixels in hyperspectral scenes and the need to estimate the number of endmembers in a given scene; therefore, sparse unmixing has attracted much attention [13].

Various sparse unmixing algorithms have been developed from different perspectives [31], [32]. Zare et al. proposed a method of sparsity that promotes iterated constrained endmember detection in hyperspectral imagery [33]. This algorithm attempts to autonomously determine the number of endmembers, which is found by adding a sparsity-promoting term to the iterated constrained endmember’s objective function. Sparse unmixing by variable splitting and augmented Lagrangian (SUnSAL) was developed in [27]. The algorithms proposed in that study are based on the alternating direction method of multipliers, which decomposes a difficult problem into a sequence of simpler ones. Hence, these algorithms achieve higher accuracy and shorter time than earlier ones.

Although sparsity-based unmixing methods are simple and fast, it is not always true that they can obtain better spectral unmixing results than traditional unmixing techniques [22], [34]. There are two main reasons for this. On one hand, in real applications, the high mutual coherence of the hyperspectral libraries imposes limits on the performance of sparse unmixing techniques. In other words, signatures that are similar to each other are more difficult to unmix [22], [28], [35]. On the other hand, many sparse unmixing techniques only consider the spectral information while ignoring the possible spatial correlations between pixels. Intuitively, the utilization of spatial correlation integrated with spectral information should improve the performance of spectral unmixing algorithms [36], so this is an important factor to consider.

To mitigate these drawbacks, many improved unmixing algorithms based on sparsity have been proposed recently. For instance, enhancing spectral unmixing using local neighborhood weights [36] was developed to utilize both spectral information and spatial information. Iordache et al. proposed the sparse unmixing via variable splitting augmented Lagrangian and total variation (SUnSAL-TV) [37] by exploiting the spatial contextual information present in the hyperspectral images. In [14], Qian et al. extended the nonnegative matrix factorization (NMF) method by incorporating the L1/2 sparsity constraint (the L1/2-NMF method). Lu et al. proposed a method for sparse unmixing by incorporating manifold regularization into sparsity-constrained NMF in [12]. Because this additional term can maintain a close link between the original image and the material abundance maps, Lu et al.’s approach leads to a more desired unmixing performance. Further, Iordache et al. proposed the collaborative sparse unmixing via variable splitting and augmented Lagrangian (CLSUnSAL) algorithm [28]. This method improves the unmixing results by solving a joint sparse regression problem because the sparsity is simultaneously imposed on all pixels in a given hyperspectral dataset. Zhong et al. proposed a sparse unmixing algorithm based on non-local means by exploiting similar patterns and structures in the abundance image [23]. Sparse unmixing using spectral a priori information (SUnSPI), which incorporates the spectral a priori information into sparse unmixing model was proposed in [4]. These algorithms are able to improve the spectral unmixing accuracy because they consider both spectral information and spatial information.

However, the high mutual coherence of spectral libraries, jointly with their ever-growing dimensionality, strongly limits the practical applicability of sparse unmixing [13], [44]. This limitation has been partially mitigated by the dictionary pruning strategy. For example, there is a sparse unmixing methodology for obtaining such a dictionary pruning [38], the multiple signal classification and collaborative sparse regression method (MUSIC-CSR) [13], a robust subspace solution for dictionary pruning [32], sparse unmixing with dictionary pruning for hyperspectral change detection [39], and an unmixing algorithm via low-rank representation based on space consistency constraint and spectral library pruning [22]. These algorithms improve unmixing performance and considerably decrease the runtime of sparse unmixing algorithms.

As mentioned above, a large number of sparse unmixing algorithms have been developed in the last ten years. However, sparse unmixing still faces huge challenges. For example, how to reduce the running time and further improve the performance of the sparse unmixing. To mitigate the above drawbacks, in this paper, we propose a hyperspectral unmixing method based on dictionary pruning and reweighted sparse regression.

In this paper, we assume that the hyperspectral dataset to be unmixes is well approximated by the LMM. We exploit the fact that most hyperspectral datasets exist in a lower dimensional subspace and the number of endmembers present in a given scene is often much less than the number of library signatures [13], [22], [32], [38]. Based on this fact, we propose a model aimed at mitigating the abovementioned limitations of hyperspectral sparse unmixing. First, we identify the signal subspace, which is estimated using the hyperspectral subspace identification by the minimum error (HySime) algorithm [40]. Second, we prune the spectral library by computing the projection error from each library member to the estimated signal subspace. Only the members from the original spectral library that have a projection error below a preset threshold T are retained. Thus, we obtain a pruned spectral library, the dimensionality of which is usually much smaller than that of the original library. Then, we propose a methodology based on weighted sparse regression to further improve the unmixing performance of the CLSUnSAL algorithm [28]. Finally, the proposed modified CLSUnSAL algorithm operates on a pruned spectral library. Because of the lower spectral library dimensionality and enhancement of the proposed algorithm, the performance of the sparse unmixing is naturally improved. The obtained unmixing results are demonstrated for both simulated data and real hyperspectral data.

The remainder of this paper is organized as follows. Section 2 reviews the LMM and some related work. Section 3 describes the proposed method. Section 4 analyses
the performance of the proposed approach with simulated data and real hyperspectral data. Section 5 is discussion of the Parameters Setting. Section 6 concludes the paper with some remarks.

2. Related Work

In this section, we first outline the LMM. Then, two basic methods for our algorithm, the CLSUnSAL algorithm and dictionary pruning using the subspace approach, are introduced.

2.1 Sparse Unmixing under the LMM

Because of its simplicity, the LMM is widely used in hyperspectral images [41], [42]. Let \( y \) denote a given hyperspectral observed column vector with \( L \) spectral signatures, a pixel \( x \) can be expressed as a linear combination of spectral signatures in an \( L \times m \) spectral library \( A \), as follows [18], [28]:

\[
y = Ax + n
\]

where the elements of vector \( x \in \mathbb{R}^m \) are the fractional abundances of each material in the pixel and vector \( n \in \mathbb{R}^L \) holds the errors affecting the measurements at each spectral band.

As mentioned in [17], the abundance fractions must be nonnegative and sum to one. The constraints \( x \geq 0 \) and \( \|x\|_1 = 1 \) are called, in the hyperspectral community, the abundance non-negativity constraint and abundance sum-to-one constraint, respectively, and are often imposed on model (1) [13].

Assuming that the dataset contains \( p \) pixels organized in matrix \( Y = [y_1, y_2, \ldots, y_p] \) and \( X = [x_1, x_2, \ldots, x_p] \) is the abundance fraction matrix. Then, (1) can be rewritten as follows:

\[
Y = AX + N
\]

where \( N = [n_1, n_2, \ldots, n_p] \) is the noise and error matrix.

2.2 CLSUnSAL

The CLSUnSAL algorithm [28] adopts a collaborative sparse regression framework that improves the unmixing results by solving a joint sparse regression problem. Reference [28] solved the following optimization problem:

\[
\min_x \|AX - Y\|_F^2 + \lambda \sum_{k=1}^m \|x_k\|_2 \quad \text{subject to: } X \geq 0
\]

where \( X \) is an \( L \times p \) matrix that denotes a hyperspectral image with \( L \) spectral bands and \( p \) pixels. Matrix \( A \) is an \( L \times m \) spectral library that has \( m \) spectral signatures, \( \|X\|_F = \sqrt{\text{trace}(XX^T)} \) denotes the Frobenius norm, and \( \lambda > 0 \) is a regularization parameter. Matrix \( X \) is an \( m \times p \) abundance fraction matrix and \( X \geq 0 \), the convex term \( \sum_{k=1}^m \|x_k\|_2 \) (where \( x_k \) denotes the \( k \)-th line of \( X \)) is the \( \ell_{2,1} \) mixed norm, which promotes sparsity among the lines of \( X \). Using the notation \( \|X\|_{2,1} = \sum_{k=1}^m \|x_k\|_2 \) to denote the \( \ell_{2,1} \) norm, optimization problem (3) can be written in the following equivalent form:

\[
\min_x \|AX - Y\|_F^2 + \lambda \|X\|_{2,1} \quad \text{subject to: } X \geq 0
\]

The optimization problem in (4) was solved via the variable splitting and alternating direction method of multipliers (ADMM) method [27], [28]. The great advantage of CLSUnSAL, which has an objective function composed of only two terms, is that it needs only one regularization parameter, which strongly alleviates the computational load and simplifies the parameter setting process [28].

2.3 Dictionary Pruning Using the Subspace Approach

As discussed in Sect. 1, sparse unmixing aims to find the optimal subset of endmembers from a given spectral library. However, the large dictionary size and high mutual coherence of the library limits the operational applicability of sparse unmixing [13], [32], [38]. These two difficulties may be circumvented by applying dictionary pruning. In [13], the authors proposed a dictionary pruning strategy based on the subspace approach. Let us assume that the noise matrix \( N \) is equal to zero. Given the hyperspectral dataset \( Y = AX \in \mathbb{R}^{L \times p} \), with \( X \in \mathbb{R}^{m \times p} \), this subspace method may be best described as follows.

When \( \text{rank}(Y) = \|X\|_0 = k < L \), in this case, we write \( Y = A_S X^S \), where \( A_S \in \mathbb{R}^{L \times k} \) and \( X^S \) are the matrices holding, respectively, the columns of \( A \) and rows of \( X \), whose indices are in \( S \). Assuming that \( \text{spark}(A) > k + 1 \) and \( X \) has full row rank, we have \( \text{range}(Y) = \text{range}(A_S) \). Let \( U_S \in \mathbb{R}^{L \times k} \) denote a matrix that contains the first \( k \) left singular vectors of \( Y \). We have

\[
P^*_S a_j = 0 \quad \text{if} \quad a_j = a_{j_S} \quad \text{for some} \quad S \in \{1, \ldots, k\}
\]

where \( P^*_S a_j = I - U_S U_S^T \) is the projector on \( \text{range}(A_S)^\perp \). The physical meaning of (5) is that if a spectral sample \( a_j \) in the dictionary is also one of the spectral signatures in the scene, then it must be perpendicular to the orthogonal complement signal subspace. From an algorithm viewpoint, we can correctly identify the index set \( \{j_1, \ldots, j_k\} \) using the equations on the left-hand side of (5), at least, in the noiseless case.

However, noise does exist in real applications. Thus, the left-hand side of (5) cannot always be true. Under such circumstances, for each spectral signature \( a_j \), \( j = 1, \ldots, m \), we compute the norm of the projection onto \( \text{range}(A_S) \), normalized by the norm of \( a_j \) as follows:

\[
epsilon(j) = \frac{\|P^*_S a_j\|_2}{\|a_j\|_2}
\]

We determine \( \hat{A} = \{j_1, \ldots, j_k\} \) such that for \( S = 1, \ldots, k \) when \( \epsilon(j_S) < \epsilon(j) \) for all \( j \notin \hat{A} \). This dictionary pruning procedure has been found to be able to improve the performance for sparse unmixing and decrease the runtime of the process [13].
3. Proposed Algorithm

3.1 Reweighted Sparse Regression

CLSUnSAL has performed better than traditional sparse-based unmixing methods. To further improve the unmixing performance of CLSUnSAL, we propose a weighted sparse regression algorithm in this paper.

3.1.1 Weighted $\ell_1$ Minimization

Let $y$ denote a column vector with $L$ spectral bands and $x$ be an $m \times 1$ vector. Sparse regression optimization can be represented as follows:

$$\min_x \|x\|_0 \quad \text{subject to} \quad Ax = y$$

where $A$ is an $L \times m$ matrix. Its relaxation problem is

$$\min_x \|x\|_1 \quad \text{subject to} \quad Ax = y$$

To enhance the sparsity of the $\ell_1$ norm in (8), [43] proposed a weighted formulation of the $\ell_1$ minimization that can be described as follows:

$$\min_x \|Wx\|_1 \quad \text{subject to} \quad Ax = y$$

where $W$ is a diagonal matrix with $\omega_1, \cdots, \omega_m$ on the diagonal, whose entry is $\omega_i^{k+1} = 1/(\|x_i^k\| + \epsilon)$, where $\epsilon$ is a small positive value, which means the weights used for the next iteration were computed from the value of the current solution. This formulation can efficiently enhance the sparsity of the solution and improve the estimation performance over the $\ell_1$ norm [25].

3.1.2 Weighted Sparse Regression Based on CLSUnSAL

Inspired by [43], we propose a weighted sparse regression formulation based on CLSUnSAL (W-CLSUnSAL) as follows:

$$\min_X \|X - Y\|_F^2 + \lambda \|WX\|_{2,1} \quad \text{subject to} \quad X \succeq 0$$

where $W$ is also a diagonal matrix with $\omega_1, \cdots, \omega_m$ on the diagonal. How to set the values of $W$ is an immediate question. Because of the unavailability of $X$, we adopt an iterative reweighted approach to design $W$.

$$W^{(k+1)} = \text{diag}(1/(\|X^{(k)}(1,:)\|_2 + \epsilon), \cdots, 1/(\|X^{(k)}(m,:)\|_2 + \epsilon))$$

Where $X^{(k)}(j,:)$ denotes the jth row of $X$ estimated at the kth iteration and $\epsilon$ is a small positive value. The introduction of parameter $\epsilon$ in (11) is adopted to provide stability and ensure that in the next step a nonzero estimate is not strictly prohibited by a zero-valued component in $X^{(k)}(j,:)$.

As shown in [43], a cautious choice of $\epsilon$ provides the stability necessary to correct for inaccurate coefficient estimates. Matrix $W$ enhances the sparsity of the endmember spectra in library $A$ and improves the unmixing performance, as shown in Sect. 4.

To solve problem (10), we adopt variable splitting and ADMM.

3.1.3 W-CLSUnSAL Algorithm

The optimization problem (10) can be rewritten as the following equivalent form:

$$\min_X \frac{1}{2} \|AX - Y\|_F^2 + \lambda \|WX\|_{2,1} + I_{R_+}(X)$$

where $I_{R_+}(X)$ is the indicator function and $I_{R_+}(X)$ is zero if $X$ is nonnegative and $+\infty$ otherwise.

Optimization problem (12) has the following equivalent formulation:

$$\min_X I_{V_1, V_2, V_3} \frac{1}{2} \|V_1 - Y\|_F^2 + \lambda \|V_2\|_{2,1} + I_{R_+}(V_3) \quad s.t. \quad V_1 = AX, V_2 = WX, V_3 = X$$

By introducing scaled Lagrangian multipliers $D = (D_1, D_2, D_3)$, the augmented Lagrangian function of (13) can be defined as

$$L(X, V_1, V_2, V_3, D_1, D_2, D_3) = \frac{1}{2} \|V_1 - Y\|_F^2 + \lambda \|V_2\|_{2,1} + I_{R_+}(V_3) + \frac{\mu}{2} \|AX - V_1 - D_1\|_F^2 + \frac{\mu}{2} \|WX - V_2 - D_2\|_F^2 + \frac{\mu}{2} \|X - V_3 - D_3\|_F^2$$

Algorithm 1 W-CLSUnSAL

1. Initialization: set $k = 0$, $W^{(0)} = I_*$, $\mu \geq 0$, $X^{(0)}$, $V_1^{(0)}, V_2^{(0)}, V_3^{(0)}, D_1^{(0)}, D_2^{(0)}, D_3^{(0)}$.

2. Repeat:
   3. $X^{(k+1)} = \mathbf{arg} \min_X \frac{\mu}{2} \|AX - V_1^{(k)} - D_1^{(k)}\|_F^2 + \frac{\mu}{2} \|WX - V_2^{(k)} - D_2^{(k)}\|_F^2$
   4. $V_1^{(k+1)} = \mathbf{arg} \min_{V_1} \frac{1}{2} \|V_1 - Y\|_F^2 + \frac{\mu}{2} \|AX - V_1 - D_1^{(k)}\|_F^2 + \frac{\mu}{2} \|WX - V_2^{(k)} - D_2^{(k)}\|_F^2$
   5. $V_2^{(k+1)} = \mathbf{arg} \min_{V_2} \|V_2\|_{2,1} + \frac{\mu}{2} \|AX - V_1^{(k+1)} - D_1^{(k+1)}\|_F^2 + \frac{\mu}{2} \|WX - V_2^{(k+1)} - D_2^{(k+1)}\|_F^2$
   6. $V_3^{(k+1)} = \mathbf{arg} \min_{V_3} I_{R_+}(V_3) + \frac{\mu}{2} \|AX - V_1^{(k+1)} - D_1^{(k+1)}\|_F^2 + \frac{\mu}{2} \|WX - V_2^{(k+1)} - D_2^{(k+1)}\|_F^2$
   7. $W^{(k+1)} = \text{diag}(\frac{1}{\|X^{(k)}(1,:)\|_2 + \epsilon}, \cdots, \frac{1}{\|X^{(k)}(m,:)\|_2 + \epsilon})$
   8. Update Lagrangian multipliers:
      $D_1^{(k+1)} \leftarrow D_1^{(k)} - \mu (AX - V_1^{(k+1)} - D_1^{(k+1)})$  
      $D_2^{(k+1)} \leftarrow D_2^{(k)} - \mu (WX - V_2^{(k+1)} - D_2^{(k+1)})$  
      $D_3^{(k+1)} \leftarrow D_3^{(k)} - \mu (X - V_3^{(k+1)} - D_3^{(k+1)})$
   9. Update iteration: $k \leftarrow k + 1$

10. Until the stopping criterion is satisfied.
The pseudocode of W-CLSUnSAL is shown in Algorithm 1.

3.2 Dictionary Pruning and the W-CLSUnSAL Algorithm

To mitigate the high mutual coherence of spectral libraries, we propose an algorithm that combines dictionary pruning and the W-CLSUnSAL algorithm, which is called the DPW-CLSUnSAL algorithm.

A pruned library $A_Q$ is detected by dictionary pruning using the subspace approach in Sect. 2.3. In real hyperspectral applications, however, the subspace identification is affected by a few degradation mechanisms, which include nonlinearities, calibration errors among the signatures available in the spectral library, and spectral variability. These degradation mechanisms may lead to the incorrect detection of support $S$, described in Sect. 2.3. To avoid missing an active signature, we introduce a preset threshold $T$, which is larger than the number of endmembers resulting from the HySime estimations but nevertheless much smaller than $m$, the number of signatures in library $A$. We denote the reduced library as $A_Q$. In addition, we have, with high probability, $S \subset Q$, that is, we do not often miss active signatures.

Let $A$ be a given spectral library. Then, the steps of our method are as follows: (1) We identify the signal subspace using the HySime algorithm. (2) A pruned library $A_Q$ is detected by dictionary pruning using the subspace approach in Sect. 2.3 with a preset threshold $T$. (3) To estimate the corresponding fractional abundances from $A_Q$, we replace $A$ with $A_Q$ in (10). Thus, we solve a following weighted collaborative sparse regression problem utilizing the W-CLSUnSAL algorithm described in Sect. 3.1.3 as follows:

$$\min_{E} \|A_Q E - Y\|_F^2 + \lambda \|WE\|_{2,1} \quad \text{subject to: } E \geq 0$$

The pseudocode for the resulting DPW-CLSUnSAL algorithm is shown in Algorithm 2.

In step 5, we sort the normalized projection errors in increasing order. In step 6, we retain the indices of the first $q$ in set $\pi$ when a preset threshold $T$ is satisfied. Step 8 solves the weighted collaborative sparse regression optimization problem using the pruned library $A_Q$.

It is worth noting that our improved algorithm is similar to MUSIC-CSR [13], but the difference is also obvious. The main difference is that our improved algorithm calls W-CLSUnSAL algorithm described in Sect. 2.3 in step 7, while the MUSIC-CSR method calls the classical CLSUnSAL algorithm.

The numerical and visual comparisons between them will be shown in Sect. 4.

4. Experiments with Simulated Data and Real Hyperspectral Data

In this section, we evaluate the effectiveness of the proposed dictionary pruning and weighted sparse regression method in various simulated and real scenarios. In Sect. 4.1, we use five simulated datasets to analyze the performance of the proposed approach, and the estimation accuracy and computational performance of the results are discussed. In Sect. 4.2, we qualitatively evaluate the proposed method with respect to other traditional methods using a real hyperspectral dataset.

4.1 Experimental Results Using Simulated Datasets

In this section, we use synthetic hyperspectral images to demonstrate the effectiveness of the proposed approach. We have considered three spectral libraries that we use in our experiments: A1, A2, and A3. A1 and A2 are two different dictionaries of minerals which were randomly extracted from the USGS library splib06 [28], which was released in September 2007. It comprises spectral signatures with 224 spectral bands, which are distributed uniformly in the interval of 0.4–2.5 $\mu$m. The mutual coherence of the library is very close to one. A1 has 342 spectral signatures, and the angle between any two different signatures is larger than 3$^\circ$. A2 has 240 spectral signatures, and the angle between any two elements of A2 is larger than 4.44$^\circ$. A3 was obtained from a random selection of 120 materials from the NASA JSC Spacecraft Materials Spectral Database [45]. It comprises 262 spectral library signatures with 100 bands.

Using library A1, we generate three data cubes of 5,000 pixels, each containing a different number of endmembers, namely, $d = \{2, 5, 8\}$. The three data cubes, DC1, DC2, and DC3, were generated using endmembers that were randomly chosen from library A1. In each simulated pixel, the abundances are generated following the uniform Dirichlet distribution. The obtained data cubes were contaminated with zero-mean independent and identically distributed Gaussian noise at three signal-to-noise ratios (SNRs), namely, 30, 40, and 50 dB, which are common levels in hyperspectral applications.

According to the methodology of [37], we generated simulated data cube DC4 using library A2. DC4 has 75 x 75 pixels, using 224 bands per pixel. Each simulated pixel
was generated with five randomly selected signatures from A2 as the endmembers, using an LMM and imposing the ASC in each pixel. Figure 1 (a) shows the simulated image, in which there are pure regions, as well as mixed regions constructed using mixtures of two to five endmembers, distributed spatially in the form of distinct square regions. The true fractional abundances of each of the five endmembers are shown in Figs. 1 (b)–(f). The background pixels consist of mixtures of the same five endmembers, but their fractional abundance values were randomly fixed to values 0.1149, 0.0741, 0.2003, 0.2055, and 0.4051. The scene was again contaminated with white noise using the same SNR value adopted for DC1, DC2, and DC3.

Using library A3, we generated simulated data cube DC5 by using nine randomly selected signatures from A3. DC5 contains 100 × 100 pixels, and the fractional abundances satisfy the ANC and the ASC. Figure 2 shows the true abundances of the nine endmembers. The scene was again contaminated with white noise, using the same SNR values adopted for DC4.

We measured the unmixing performance using the signal-to-reconstruction error (SRE) [38]:

\[
SRE = \frac{E[\|x\|^2]}{E[\|x - \hat{x}\|^2]}\]

expressed in dB: \(SRE(dB) = 10 \log_{10}(SRE)\). The higher the SRE, the better the quality of the unmixing.

We also computed the spectral angle distance (SAD), which is used to compare the similarity of the \(j\)th true endmember signature \(x_j\) and its estimate \(\hat{x}_j\). The SAD is formally defined as follows [14]:

\[
SAD_j = \arccos\left(\frac{x_j^T \hat{x}_j}{\|x_j\| \|\hat{x}_j\|}\right)
\]

Parameter \(\varepsilon\) plays an important role in the proposed DPW-CLSUnSAL algorithm and we empirically set it to 0.0001 in our simulated experiment. In addition, regularization parameter \(\lambda\) affects the performance of unmixing methods. To obtain the best settings for this parameter, we tested the values in

\[
C = \{10^{-5}, 5 \times 10^{-5}, 10^{-4}, 5 \times 10^{-4}, 10^{-3}, 5 \times 10^{-3}, 10^{-2}, 5 \times 10^{-2}, 0.1, 0.2, 0.5, 1, 2\}
\]

to determine the best value of \(\lambda\) for SUnSAL, CLSUnSAL, MUSIC-CSR, and the proposed algorithm. Table 1 shows the relationship between SRE and \(\lambda\) in DC1 at SNR = 30dB for the proposed algorithm. Table 1 shows that the obtained performance does not vary substantially, which indicates the robustness of the algorithm to this parameter. For all the algorithms in this paper, the parameters were carefully tuned for optimal performance. For the sake of comparison, the MUSIC-CSR algorithm uses the same numbers of the pruned library as

Table 1 Relation between SRE and \(\lambda\) in simulated DC1 with SNR = 30 dB by the proposed method.

<table>
<thead>
<tr>
<th>(\lambda)</th>
<th>10^{-5}</th>
<th>5 \times 10^{-5}</th>
<th>10^{-4}</th>
<th>5 \times 10^{-4}</th>
<th>10^{-3}</th>
<th>5 \times 10^{-3}</th>
<th>0.01</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRE</td>
<td>21.81</td>
<td>21.79</td>
<td>20.10</td>
<td>22.57</td>
<td>21.43</td>
<td>22.84</td>
<td>22.94</td>
</tr>
<tr>
<td>SAD</td>
<td>0.05</td>
<td>0.1</td>
<td>0.2</td>
<td>0.5</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>
the proposed algorithm.

In the following, we illustrate the DPW-CLSUnSAL algorithm using DC2, which is contaminated with noise at SNR = 30 dB. The performance of the proposed method is compared with SUnSAL, CLSUnSAL, and MUSIC-CSR.

In Algorithm 2, we first estimate the data subspace using the HySime algorithm, and library A1 members are projected onto the estimated subspace. Second, we compute the Euclidean distance to the subspace for each member in library A1. Figure 3 shows the obtained projection errors for all members. The errors corresponding to the true endmembers are highlighted indicated by red circles.

From Fig. 3, we can see that the errors corresponding to true endmembers have the lowest projection errors of all the library members. To avoid missing an active signature, we set parameter $T$ to 10, and then the reduced library $A_Q$ was built by retaining the members corresponding to the lowest $q = 10$ projection errors, which correspond to a maximum allowed error of 0.0365. Finally, the unmixing was performed using the W-CLSUnSAL algorithm with reduced library $A_Q$.

Figure 4 shows the unmixing results obtained by different unmixing algorithms for DC2. Figure 4 (a) shows the true fractional abundances of the five endmembers in DC2, and Fig. 4 (b), 4 (c), 4 (d), and 4 (e) show the fractional abundances estimated by the SUnSAL algorithm, CLSUnSAL algorithm, the MUSIC-CSR algorithm, and the proposed algorithm, respectively.

Figure 4 shows that the SUnSAL algorithm can only correctly identify one of the five true endmembers, whereas the CLSUnSAL algorithm can identify four of them. The CLSUnSAL algorithm performs substantially better than the SUnSAL algorithm because it enforces joint sparsity among all the pixels. Due to the high mutual coherence of the library signatures, this unmixing problem is quite difficult. However, as shown in Fig. 4 (e), the method proposed in this paper is able to identify all the true endmembers because we consider the pruning strategy and sparsity enhancement strategy simultaneously. We can also see that the MUSIC-CSR algorithm is more accurate than the CLSUnSAL algorithm, and less accurate than the proposed algorithm. The reason is that the MUSIC-CSR is focused on the pruning strategy. Hence, we conclude that the fractional abundances inferred by the proposed are indeed closer to the true ones than those inferred by the SUnSAL, CLSUnSAL, and MUSIC-CSR algorithms.

Table 2 shows the SRE and SAD (per endmember) obtained by applying SUnSAL, CLSUnSAL, MUSIC-CSR, and the proposed algorithm to simulated DC1, DC2 and DC3 with noise. The parameter $T$ was set as follows: 5 for DC1, 10 for DC2, and 20 for DC3 in our proposed algorithm.

Table 2 shows that the proposed algorithm outperforms the other three methods. The proposed method attains the highest SRE and the lowest SAD values in all cases. The accuracy of all algorithms decreases when the observations.
are affected by high levels of noise. Another important observation from Table 2 is that the accuracy of the considered unmixing algorithms decreases as the number of endmembers in the data cube increases. This is because the sparsity of the solution mitigates the difficulty of the unmixing caused by the high mutual coherence of the libraries.

Table 3 reports the computation times for simulated DC1, DC2 and DC3. We set the regularization parameter $\lambda = 10^{-2}$ for all algorithms. The maximum number of iterations was set to 1,000 in all cases. In the proposed algorithm, we set $T$ to 20 and then the pruned dictionary $A_Q$ retained 20 members of original library $A_1$ for all simulated data cubes. The methods were performed on a PC with an AMD Athlon™ II X4 645 Processor @3.10 GHz and 4.00 GB of RAM memory.

Table 3 shows that the runtime of the proposed algorithm is significantly shorter than that of SUnSAL and CLSUnSAL methods. The reason is that our algorithm converges faster when the library is pruned. Using the same pruning strategy, the proposed algorithm runs a little longer than the MUSIC-CSR algorithm because our proposed method considers the sparsity enhancement strategy. However, the performance of the proposed algorithm is much better than that of the MUSIC-CSR algorithm in Table 2. We can also see that the computation times of all algorithms were affected by high noise, namely, the higher the value of SNR, the lower the runtime of the algorithm.

We set $T$ to 10 in our proposed algorithm, and then
Fig. 6  Fractional abundances maps estimated by different unmixing algorithms for endmember 9 in DC5 and, from top to bottom, SNR is 30dB, 40dB, and 50dB: (a) abundances map for SUnSAL; (b) abundances map for CLSUnSAL; (c) abundances map for the MUSIC-CSR algorithm; (d) abundances map for the proposed algorithm.

Fig. 5 shows the fractional abundance maps estimated by different unmixing algorithms for endmember 5 in DC4. From Fig. 5, we can find that the regions with high abundance of the considered endmember are able to depict better, while mixed regions with lower concentrations are substantially more uniform. In fact, DPW-CLSUnSAL obtained better results than the other algorithms, especially the abundance maps obtained by SUnSAL, CLSUnSAL, and MUSIC-CSR are full of noise points with high decibel values (30dB). The reason is that the spectral angle distance in A2 is quite small, which makes it difficult to separate the endmember signatures from noise. Therefore, our improved algorithm can restrain the noise.

We set $T$ to 20 in our proposed algorithm, then, Fig. 6 shows the fractional abundance maps estimated by different unmixing algorithms for endmember 9 in DC5. From Fig. 6, we can see that there are many noise points in the results of the SUnSAL, CLSUnSAL, and MUSIC-CSR algorithms, especially endmembers at the boundary that are not extracted completely, while our improved algorithm obtain better results. The fact is that the fractional abundances in DC5 are smooth, with sharp transitions. We also find that the SNR affects the performance of the algorithm, and the higher the noise value, the worse the result.

To further test the performance of the improved algorithm, we have calculated SRE and SAD by using different unmixing algorithms for simulated DC4 and DC5 with noise, which are shown in Table 4. In the proposed algorithm, we set the parameters $T$ to 10 and 20 for DC4 and DC5, respectively. In Table 4, we can observe that the improved algorithm obtains more accurate results than the SUnSAL, CLSUnSAL, and MUSIC-CSR algorithms. Especially when the simulated data set has a high level of SNR, the improved algorithm can accurately select true endmembers from the original spectral library.

Table 5 reports the computation times of different unmixing algorithms for simulated DC4 and DC5. For comparison, we set the parameter $T$ to 20 in the proposed algorithm for DC4 and DC5. It’s important to note that the values reported correspond to the average times. The running environment, as well as the settings of parameters and the maximum number of iterations, are identical to those in Table 3 above.

From Table 5, we can find that our improved algorithm has the shorter running time than the SUnSAL, and CLSUnSAL methods, while a little longer than the
**Table 4** SRE(dB) and SAD using the considered unmixing algorithms for simulated DC4 and DC5.

<table>
<thead>
<tr>
<th>Data Cubes name</th>
<th>SNR (dB)</th>
<th>Method</th>
<th>SRE (dB)</th>
<th>SAD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>SUnSAL</td>
<td>CLSUnSAL</td>
<td>MUSIC-CSR</td>
</tr>
<tr>
<td>DC4</td>
<td>30</td>
<td>5.4044</td>
<td>5.9704</td>
<td>6.1472</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>11.1003</td>
<td>10.5819</td>
<td>11.9351</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>11.5933</td>
<td>17.7704</td>
<td>18.8924</td>
</tr>
<tr>
<td>DC5</td>
<td>30</td>
<td>7.2992</td>
<td>8.6748</td>
<td>8.8459</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>13.5050</td>
<td>14.9491</td>
<td>15.0143</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>15.8910</td>
<td>18.9876</td>
<td>19.3067</td>
</tr>
</tbody>
</table>

**Table 5** Computation times(s) for simulated DC4 and DC5.

<table>
<thead>
<tr>
<th>Data Cubes name</th>
<th>SNR (dB)</th>
<th>Method</th>
<th>Subspace Dimension</th>
<th>Endmembers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>SUnSAL</td>
<td>CLSUnSAL</td>
<td>MUSIC-CSR</td>
</tr>
<tr>
<td>DC4</td>
<td>30</td>
<td>0.4352</td>
<td>0.4179</td>
<td>0.4037</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>0.2612</td>
<td>0.2794</td>
<td>0.2355</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.2519</td>
<td>0.1351</td>
<td>0.1089</td>
</tr>
<tr>
<td>DC5</td>
<td>30</td>
<td>0.6925</td>
<td>0.4123</td>
<td>0.3975</td>
</tr>
<tr>
<td></td>
<td>40</td>
<td>0.3879</td>
<td>0.2658</td>
<td>0.2613</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>0.1972</td>
<td>0.1203</td>
<td>0.1158</td>
</tr>
</tbody>
</table>

**Fig. 7** Subimage and projection errors of real data: (a) band 30 ($\lambda = 667.3$ nm) of a subimage from the AVIRIS Cuprite Nevada dataset; (b) projection errors for all members onto this real dataset. The projection errors corresponding to these four materials are indicated by red circles.

In all cases, the running time of the improved algorithm is less than 13% of that of the CLSUnSAL algorithm in DC4, and less than 15% for DC5. This is because our dictionary pruning strategy significantly improves the running speed of the algorithm.

Table 5 reports the average runtimes on the real hyperspectral dataset using the same computing environment as validate the performance of hyperspectral unmixing algorithms. The scene comprises 224 spectral bands ranging from 0.4 to 2.5 $\mu$m, with a nominal spectral resolution of 10 nm. After removing low SNR and water absorption bands, 188 bands remain. The portion used in the experiments consists of 250 x 191 pixels.

In summary, by quantitatively and qualitatively comparing the results obtained by different unmixing algorithms, we can conclude that our proposed algorithm significantly improves the unmixing results in all cases, and the running speed is faster than that of other most unmixing methods. Although the results of our simulation experiments are encouraging, further experiments should be conducted with real hyperspectral data.

**4.2 Experimental Results Using Real Datasets**

In this section, we present the results of tests on the well-known AVIRIS Cuprite dataset [37], which was captured in Nevada in 1997. This scene has been widely used to

MUSIC-CS algorithms. In all cases, the running time of the improved algorithm is less than 13% of that of the CLSUnSAL algorithm in DC4, and less than 15% for DC5. This is because our dictionary pruning strategy significantly improves the running speed of the algorithm.

In summary, by quantitatively and qualitatively comparing the results obtained by different unmixing algorithms, we can conclude that our proposed algorithm significantly improves the unmixing results in all cases, and the running speed is faster than that of other most unmixing methods. Although the results of our simulation experiments are encouraging, further experiments should be conducted with real hyperspectral data.

**Table 6** reports the average runtimes on the real hyperspectral dataset using the same computing environment as
used for the simulated data. The runtime of the proposed algorithm is 39.02% of that of SUnSAL and is 11.21% of that of CLSUnSAL. Overall, our proposed algorithm is much faster than the SUnSAL and CLSUnSAL algorithms, and is a little longer than MUSIC-CSR.
Table 6  Computation times for real data.

<table>
<thead>
<tr>
<th>Methods</th>
<th>SUUnSAL</th>
<th>CLSUnSAL</th>
<th>MUSIC-CSR</th>
<th>Proposed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>41.0652</td>
<td>142.9953</td>
<td>14.7183</td>
<td>16.0246</td>
</tr>
</tbody>
</table>

Fig. 9  Relationship between SRE and \( T \) in DC1 with SNR = 30 dB for the proposed method.

5. Discussion

The greatest advantage of the proposed algorithm is the adoption of a dictionary pruning strategy, which can identify a subset of the library for unmixing. Threshold \( T \) controls the size of the subset of the original spectral library. Figure 9 shows the relationship between SRE and \( T \) in DC1 at SNR = 30dB. Higher values of \( T \) lead to lower SRE values. However, when the dataset contains high levels of noise and the number of true signatures is large, threshold \( T \) should be set higher to avoid missing active signatures.

In the proposed algorithm, parameter \( \lambda \) must still be determined manually, which is a disadvantage. Hence, our future work will focus on how to automatically determine this parameter.

Although the results obtained using the proposed algorithm are very encouraging, the algorithm proposed in this paper is based on the LMM. In real hyperspectral images, nonlinearities do exist. Therefore, the nonlinear problem of mixed pixels should also be considered appropriately in the future design of the unmixing algorithm.

6. Conclusions

In this paper, we developed a hyperspectral unmixing method based on dictionary pruning and reweighted sparse regression. To mitigate the problems of high mutual coherence in spectral libraries, the algorithm identifies a subset of the original library elements using dictionary pruning strategy. The size of this subset is much smaller than that of the original library. To further improve the performance of unmixing method, we propose a weighted sparse regression algorithm called the W-CLSUnSAL algorithm. Thus, our proposed algorithm improves the accuracy of unmixing solutions while substantially reducing the average runtime.

Simulated and real hyperspectral data experiments showed that the proposed algorithm is more accurate and much faster than the SUUnSAL and CLSUnSAL algorithms.

Acknowledgments

This study is financially supported by the National Key R&D Program of China (2017YFC0601505), Chinese National Natural Science Foundation (41672325, 41602334), the Leading talent training project of Neijiang Normal University under 2017[Liu Yi-He]; the Innovative Team Program of the Neijiang Normal University under 17TD03; the Sichuan province academic and technical leader training funded projects under 13XSJS002; the Foundation of Ph. D. Scientific Research of Neijiang Normal University under 2019[zhang shuang] and 2019[wang jiujiang].

References


HAN et al.: FAST HYPERSPECTRAL UNMIXING VIA REWEIGHTED SPARSE REGRESSION


Hongwei Han received the B.S. degree in Mathematics education from Yunnan University, Yunnan, in 2002, and M.S. degree in electronics and communication engineering from Chengdu University of Technology, Chengdu, in 2009. He is currently completing a Ph.D in Geomatics in Chengdu University of Technology. His research interests include hyperspectral remote sensing image processing, Partial differential equation and its application.

Ke Guo is currently a professor and doctoral supervisor in the Chengdu University of Technology. He is academic and technical leader of sichuan province, and enjoys special subsidies from the state council government. He is currently the head of the teaching team of mathematics geology in sichuan province, and the head of the scientific research and innovation team of high-level resources and environment in sichuan province. He has long been engaged in the teaching and scientific research of quantitative evaluation and prediction of resources and environment.
Maozhi Wang received the B.S. degree in mathematics from Southwest Normal University, Chongqing, in 1996, and M.S. degree in computer software and theory from the University of Electronics Science and Technology of China, Chengdu, in 2003 and the Ph.D. degree in geo-detection and information technology from Chengdu University of Technology, Chengdu, in 2014. He is currently a researcher in Geomathematics Key Laboratory of Sichuan Province in Chengdu University of Technology. His research interests include hyperspectral remote sensing image processing and its application.

Tingbin Zhang received the B.S. in Surveying and Mapping from Chang’An University (Xian China), M.S. in Mineral Resources and Geological Engineering from Chengdu University of Technology (Chengdu, China) and Ph. D. in Geodesy and Survey Engineering from Southwest Jiaotong University (Chengdu, China). He is currently a professor in the Chengdu University of Technology. His current research interests include remote sensing of environment, remote sensing of resources, surveying and mapping, etc.

Shuang Zhang received the B.S. degree in Mathematics and Applied Mathematics from the Neijiang Normal University, Neijiang, China, in 2007, and the M.S. degree in Control Engineering from the Institute of Optics and Electronics, Chinese Academy of Sciences, Chengdu, in 2011. He is currently pursuing the Ph.D. degree in electrical and computer engineering with the University of Macau, Macau, China. He has been an Assistant Professor in College of Computer Science, Neijiang Normal University. His current research interests include human body communication, Digital signal processing, and body sensor network.