Research Article

Backtracking-Based Simultaneous Orthogonal Matching Pursuit for Sparse Unmixing of Hyperspectral Data

Fanqiang Kong,¹ Wenjun Guo,¹ Yunsong Li,² Qiu Shen,¹ and Xin Liu¹

¹College of Astronautics, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

²State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China

Correspondence should be addressed to Fanqiang Kong; kongfq@nuaa.edu.cn

Received 12 November 2014; Revised 3 April 2015; Accepted 3 April 2015

Academic Editor: Kishin Sadarangani

Copyright © 2015 Fanqiang Kong et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Sparse unmixing is a promising approach in a semisupervised fashion by assuming that the observed signatures of a hyperspectral image can be expressed in the form of linear combination of only a few spectral signatures (endmembers) in an available spectral library. Simultaneous orthogonal matching pursuit (SOMP) algorithm is a typical simultaneous greedy algorithm for sparse unmixing, which involves finding the optimal subset of signatures for the observed data from a spectral library. But the numbers of endmembers selected by SOMP are still larger than the actual number, and the nonexistent endmembers will have a negative effect on the estimation of the abundances corresponding to the actual endmembers. This paper presents a variant of SOMP, termed backtracking-based SOMP (BSOMP), for sparse unmixing of hyperspectral data. As an extension of SOMP, BSOMP incorporates a backtracking technique to detect the previous chosen endmembers' reliability and then deletes the unreliable endmembers. Through this modification, BSOMP can select the true endmembers more accurately than SOMP. Experimental results on both simulated and real data demonstrate the effectiveness of the proposed algorithm.

1. Introduction

With the rapid development of space technology, hyperspectral remote sensing image has gained more and more attention in many application domains, such as environmental monitoring, target detection, material identification, mineral exploration, and military surveillance. However, due to the low spatial resolution of the hyperspectral imaging sensor, each pixel in the hyperspectral image often contains a mixture of several different materials. In order to deal with the problem of spectral mixing, hyperspectral unmixing is used to decompose each pixel's spectrum to identify and quantify the fractional abundances of the pure spectral signatures or endmembers in each mixed pixel [1, 2]. There are two basic models of hyperspectral unmixing used to analyze the mixed pixel problem: the linear mixture model and the nonlinear mixture model. The linear mixture model assumes that each mixed pixel can be expressed as a linear combination of endmembers weighted by their corresponding abundances [3]. The linear mixture model has been widely applied for spectral unmixing, due to its computational tractability and

flexibility. Under the linear mixture model, the traditional linear spectral unmixing algorithms based on geometry [4–8], statistics [5, 9], and nonnegative matrix factorization [10–12] have been proposed. However, some of these methods [10–12] are unsupervised and could extract virtual endmembers with no physical meaning. In [5], the presence in the data of at least one pure pixel per endmember is assumed. If the pure pixel assumption is not fulfilled because of the inadequate spatial resolution and the microscopic mixture of distinct materials, the unmixing results will not be accurate and the unmixing process is a rather challenging task.

Sparse unmixing, as a semisupervised method, has been proposed to overcome this challenge. It assumes that the observed image can be expressed as a linear combinations of spectral signatures from a spectral library that is known in advance [3, 13]. But the number of spectral signatures in the spectral library is much larger than the number of endmembers in the hyperspectral image; the sparse unmixing model is combinatorial and difficult to find a unique, stable, and optimal solution. Fortunately, sparse linear regression techniques can be used [14, 15] to solve it.

Several sparse regression techniques, such as greedy algorithms (GAs) [16–19] and convex relaxation methods [20–23], are usually adopted to solve the sparse unmixing problem. Convex relaxation methods, such as SUnSAL [20], SUnSAL-TV [21], and CLSUnSAL [22], use the alternating direction method of multipliers to efficiently solve the l_1 norm sparse regression problem which can decompose a complex problem into several simpler ones. Convex relaxation methods can obtain the global sparse optimization solution and are more sophisticated than the GAs; however they are far more complicated than the GAs. The GAs, such as OMP [16], SP [17], and CGP [18, 19], adopt one or more potential endmembers from the spectral library in each iteration that explains the largest correlation between the current residual of one input signal and the spectral library. The GAs can get an approximate solution for the l_0 norm problem without smoothing the penalty function and have low computational complexity. However the endmembers selection criterion of GAs is not optimal in the sense of minimizing the residual of the new approximation; it means that a nonexistent endmember once selected into the supporting set will never be deleted. So the GAs tend to be trapped into the local optimum and are likely to miss some of the actual endmembers. To solve the local optimal solutions problem of GAs, several representative simultaneous greedy algorithms (SGAs) are presented, including simultaneous subspace pursuit (SSP) [24], subspace matching pursuit (SMP) [24], and simultaneous orthogonal matching pursuit (SOMP) [25]. These SGAs divide the whole hyperspectral image into several blocks and pick some potential endmembers from the spectral library in each block. Then the endmembers picked in each block are associated as the endmembers sets of the whole hyperspectral data. Finally, the abundances are estimated using the whole hyperspectral data with the obtained endmembers sets. The SGAs have the same low computational complexity as the GAs and can find the actual endmembers far more accurately than the GAs. The SGAs adopt a block-processing strategy to efficiently solve the local optimal problem, but the endmembers picked in all the blocks are not all actual endmembers and the nonexistent endmembers will affect the estimation of the abundances corresponding to the actual endmembers. To solve the drawback of the SGAs, RSFoBa [26] combines a forward greedy step and a backward greedy step, which can select the actual endmembers more accurately than the SGAs.

Inspired by the existing SGA methods, we propose a sparse unmixing algorithm termed backtracking-based simultaneous orthogonal matching pursuit (BSOMP) in this paper. Similar to SOMP and SMP, it uses a block-processing strategy to select some potential endmembers and adds them to the estimated endmembers set, which divides the whole hyperspectral image into several blocks and picks some potential endmembers from the spectral library in each block. Furthermore, BSOMP incorporates a backtracking process to detect the previous chosen endmembers' reliability and then deletes the unreliable endmember from the estimated endmembers set in each iteration. Through this modification, BSOMP can identify the true endmembers set more accurately than the other considered SGA methods.

The remainder of the paper is organized as follows. Section 2 introduces the simultaneous sparse unmixing model. In Section 3, we present the proposed BSOMP algorithm and give out the theoretical analysis for the algorithm. Section 4 provides a quantitative comparison between BSOMP and previous sparse unmixing algorithms, using both simulated hyperspectral and real hyperspectral data. Finally, we conclude in Section 5.

2. Simultaneous Sparse Unmixing Model

The linear sparse unmixing model assumes that the observed spectrum of a mixed pixel is a linear combination of a few spectral signatures presented in a known spectral library. Let $y \in R^L$ denote the measured spectrum vector of a mixed pixel with L bands, $A \in R^{L \times m}$, where m is the number of signatures in the library A ; denote a $L \times m$ spectral library; then the linear sparse unmixing model can be expressed as follows [23]:

$$y = Ax + n, \quad (1)$$

where $x \in R^m$ denotes the fractional abundance vector with regard to the library A and $n \in R^L$ is the noise. Considering physical constraints, abundance nonnegativity constraint (ANC) and abundance sum-to-one constraint (ASC) are imposed on the linear sparse model as follows:

$$\begin{aligned} x &\geq 0, \\ \sum_{i=1}^m x_i &= 1, \end{aligned} \quad (2)$$

where x_i is the i th element of x .

The sparse unmixing problem can be expressed as follows:

$$\begin{aligned} \min_x \quad & \|x\|_0 \\ \text{s.t.} \quad & \|y - Ax\|_2^2 \leq \delta, \\ & x \geq 0, \end{aligned} \quad (3)$$

where $\|x\|_0$ (called the l_0 norm) denotes the number of nonzero atoms in x and $\delta \geq 0$ is the tolerated error due to the noise and model error. It is worth mentioning that we do not explicitly add the ASC in (3), because the hyperspectral libraries generally contain only nonnegative components, and the nonnegativity of the sources automatically imposes a generalized ASC [23].

The simultaneous sparse unmixing model assumes that several input signals can be expressed in the form of different linear combinations of the same elementary signals. This means that all the pixels in the hyperspectral image are constrained to share the same subset of endmembers selected from the spectral library. Then we can use SGA methods for sparse unmixing; the sparse unmixing model in (1) becomes

$$Y = AX + N, \quad (4)$$

where $Y \in R^{L \times K}$ denote the hyperspectral data matrix with L bands and K mixed pixels, $A \in R^{L \times m}$ denote a $L \times m$ spectral

Part 1 (SOMP):

- (1) Initialize hyperspectral data Y and spectral library A
- (2) Divide hyperspectral data Y into several blocks: $Y = [Y_1, Y_2, \dots, Y_B]$, and initialize the index set $S_{\text{SOMP}} = \emptyset$
- (3) **For** each block **do**
- (4) Set index set $S_b = \emptyset$ and iteration counter $k = 1$. Initialize the residual data of block b : $R_0 = Y_b$;
- (5) **While** stopping criterion 1 has not been met **do**
- (6) Compute the index of the best correlated member of A to the actual residual: $j = \arg \max_i \|(R^{l-1})^T A_i\|_2$, where A_i is the i th column of A
- (7) Update support set: $S_b = S_b \cup j$
- (8) Compute $X^l = (A_{S_b}^* A_{S_b})^{-1} A_{S_b}^* Y_b$ $\{A_{S_b}$ is the matrix containing the columns of A having the indexes from $S_b\}$
- (9) Update residual: $R^l = Y_b - A_{S_b} X^l$
- (10) $l \leftarrow l + 1$
- (11) **End while**
- (12) Set $S_{\text{SOMP}} = S_{\text{SOMP}} \cup S_b$
- (13) **End for**

Part 2 (Backtracking process):

- (14) Initialize hyperspectral data Y , the index set $S_t = S_{\text{SOMP}}$ and $A_t = A_{S_{\text{SOMP}}}$ $\{A_t$ is the matrix containing the columns of A having the indexes from $S_t\}$
- (15) **While** stopping criterion 2 has not been met **do**
- (16) Compute solution: $X^t = (A_t^* A_t)^{-1} A_t^* Y$
- (17) Compute the member of A_t having the lowest abundance: $\text{index} \leftarrow \min_j (X_t^j)$
- (18) Remove the member having the lowest fractional abundance from the index set $S_t \leftarrow S_t \setminus S_{\text{index}}$ and the endmember set: $A_t \leftarrow A_t \setminus A_{\text{index}}$
- (19) $t \leftarrow t + 1$
- (20) **End while**

Part 3 (Abundance estimation):

- (21) Estimate abundances using the original hyperspectral data matrix and the endmember set under the constraint of nonnegativity.
 $X \leftarrow \arg \min_X \|A_t X - Y\|_F$, subject to $X \geq 0$.

ALGORITHM 1: BSOMP for hyperspectral sparse unmixing.

library, $X \in R^{m \times K}$ denotes the fractional abundance matrix, each column of which corresponds with the abundance fractions of a mixed pixel, and $N \in R^{L \times K}$ is the noise matrix.

Under the simultaneous sparse unmixing model, the simultaneous sparse unmixing problem can be expressed as follows:

$$\begin{aligned}
 \min_X \quad & \|X\|_{\text{row}=0} \\
 \text{s.t.} \quad & \|Y - AX\|_F^2 \leq \delta, \\
 & X \geq 0,
 \end{aligned} \tag{5}$$

where $\|X\|_F$ denotes the Frobenius norm of X , $\|X\|_{\text{row}=0}$ is the number of nonzero rows in matrix X [25], and δ is the tolerated error due to the noise and model error.

It should be noted that the model in (5) is reasonable because there should be only a few rows with nonzero entries in the abundance matrix in light of only a small number of endmembers in the hyperspectral image, compared with the dimensionality of the spectral library [22].

3. Backtracking-Based SOMP

In this section, we first present our new algorithm, BSOMP, for sparse unmixing of hyperspectral data. Then, a theoretical

analysis of the algorithm will be given. Then, we give a convergence theorem for the proposed algorithm.

3.1. Statement of Algorithm. The whole process of using BSOMP for sparse unmixing of hyperspectral data is summarized in Algorithm 1. The algorithm includes three main parts: SOMP for endmember selection, backtracking processing, and abundance estimation. In the first part, we adopt a block-processing strategy for SOMP to efficiently select endmembers. This strategy divides the whole hyperspectral image into several blocks. Then, in each block, SOMP will pick several potential endmembers from the spectral library and add them to the estimated endmembers set. In the second part, we utilize a backtracking strategy to remove some endmembers chosen wrongly from the estimated endmembers set in previous processing and identify the true endmembers set more accurately. The backtracking processing is halted when the maximum total correlation between an endmember in the spectral library and the residual drops below threshold δ . Finally, the abundances are estimated using the obtained endmembers set under the constraint of nonnegativity.

In the following paragraphs, we provide some definitions and then give two stopping conditions for the BSOMP algorithm.

Definition 1. If x is a vector, where x_i is the i th element of x , the l_p norm of x is defined as

$$\|x\|_p = \left(\sum_i |x_i|^p \right)^{1/p} \quad \text{for } 1 \leq p < \infty, \quad (6)$$

$$\|x\|_\infty = \max_i |x_i|.$$

Definition 2. The (p, q) operator norms of the matrix A is defined as

$$\|A\|_{p,q} = \max_{x \neq 0} \frac{\|Ax\|_q}{\|x\|_p}. \quad (7)$$

Several of the (p, q) operator norms can be computed easily using the following lemma [25, 27].

Lemma 3. (1) The $(1, q)$ operator norm is the maximum l_p norm of any column of A .

(2) The $(2, 2)$ operator norm is the maximum singular value of A .

(3) The (p, ∞) operator norm is the maximum l_p norm of any row of A .

We give a stopping condition for the first part of the BSOMP algorithm:

$$\frac{\|R_l\|_F - \|R_{l-1}\|_F}{\|R_{l-1}\|_F} \leq \sigma, \quad (8)$$

where $\|R_l\|_F$ is the Frobenius norm of the residual corresponding to the l th iteration and σ is a threshold.

For the second part of the BSOMP algorithm, we give a stopping condition for deciding when to halt the iteration:

$$\frac{\|A^* R_t\|_{\infty, \infty} - \|A^* R_{t-1}\|_{\infty, \infty}}{\|A^* R_t\|_{\infty, \infty}} \leq \delta, \quad (9)$$

where Y_t is the residual corresponding to the t th iteration, A^* denotes the conjugate transpose of A , $\|A^* R_t\|_{\infty, \infty}$ is the maximum total correlation between an endmember in the spectral library A and the residual R_t , and δ is a threshold.

3.2. Theoretical Analysis. In this section, we will introduce some definitions and then give a theoretical analysis for the BSOMP algorithm.

Definition 4. The coherence μ of a spectral library A equals the maximum correlation between two distinct spectral signatures [28]:

$$\mu = \max_{k \neq h} |\langle A_k, A_h \rangle|, \quad (10)$$

where A_k is the k th column of A . If the coherence parameter is small, each pair of spectral signatures is nearly orthogonal.

Definition 5. The cumulative coherence [29, 30] is defined as

$$\mu(k) = \max_{|\Lambda| \leq k} \max_{j \notin \Lambda} \sum_{i \in \Lambda} |\langle A_i, A_j \rangle|, \quad (11)$$

where the index set $\Lambda \in S_{\text{all}}$, S_{all} is the support of all the spectral signatures in the spectral library A . The cumulative coherence measures the maximum total correlation between a fixed spectral signature and t distinct spectral signatures. Particularly, $\mu(0) = 0$.

Definition 6. Given hyperspectral image data matrix $Y \in R^{L \times K}$ and available spectral library matrix $A \in R^{L \times m}$,

S_{opt} is an index set which lists at most T endmembers presented in the hyperspectral image scene;

A_{opt} is an optimal submatrix of the spectral library A which contains the T columns indexed in S_{opt} ;

X_{opt} is the optimal abundance matrix corresponding to A_{opt} ;

Y_{opt} is a T -term best approximation of Y , and $Y_{\text{opt}} = A_{\text{opt}} X_{\text{opt}}$;

S_t is an index set which lists k endmembers after t iterations and assumes $S_{\text{opt}} \in S_t$;

A_t is a submatrix of the spectral library A which contains the k columns indexed in S_t ;

X_t is the abundance matrix corresponding to A_t ;

Y_t is the reconstruction of Y by BSOMP algorithm after t iterations (assume that BSOMP algorithm selects a nonexisting endmember to remove from index set S_t in each iteration), and $Y_t = A_t X_t$;

R_t is the residual error of Y after t iterations, and $R_t = Y - Y_t$.

Theorem 7. Suppose that the best approximation Y_{opt} of the hyperspectral image data matrix Y over the index set S_{opt} satisfies the error bound: $\|A^*(Y - Y_{\text{opt}})\|_{\infty, \infty} \leq \tau$, the first part of BSOMP has selected k potential endmembers from the spectral library to constitute the initial index set S_t , and $S_{\text{opt}} \in S_t$. After iteration t of the second part of BSOMP, halt the second part of the algorithm if

$$\|A^* R_t\|_{\infty, \infty} \leq \frac{1 - \mu(T)}{1 - \mu(T) - \mu(k - T)} \tau, \quad (12)$$

where $k \geq T$. If this inequality fails, then BSOMP must remove another nonexisting endmember in iteration $(t + 1)$.

The proof of Theorem 7 is shown in the Appendix. If the second part of the BSOMP algorithm terminates at the end of iteration t , we may conclude that the BSOMP algorithm removes t nonexisting endmembers from S_t and has chosen k optimal endmembers.

4. Experimental Results

In this section, we use both simulated hyperspectral data and real hyperspectral data to demonstrate the performance of the proposed algorithm. The BSOMP algorithm is compared with three SGAs (SOMP [25], SMP [24], and RSFoBa [26]) and three convex relaxation methods (i.e., SUNSAL [20], SUNSAL-TV [21], and CLSUNSAL [22]). All the considered

FIGURE 1: Five spectral signatures from the USGS used in our simulated data experiments. The title of each subimage denotes the mineral corresponding to the signature.

algorithms have taken into account the abundance nonnegativity constraint. The TV regularizer used in SUnSAL-TV is a nonisotropic one. For the RSFoBa algorithm, the l_∞ norm is considered in our experiment. As mentioned above, when the spectral library is overcomplete, the GAs behave worse than the SGAs and the convex relaxation methods; thus, we do not include the GAs in our experiment.

In the simulated data experiments, we evaluated the performances of the sparse unmixing algorithms in situations of different SNR level of noise; the signal-to-reconstruction error (SRE) [21] is used to measure the quality of the reconstruction abundances of spectral endmembers:

$$\text{SRE} = \frac{E[\|X\|_2^2]}{E[\|X - \hat{X}\|_2^2]}, \quad (13)$$

$$\text{SRE (dB)} = 10 \log_{10}(\text{SRE}),$$

where X denotes the true abundances of spectral endmembers and $\hat{X}$ denotes the reconstruction abundances of spectral endmembers.

4.1. Evaluation with Simulated Data. In the simulated experiments, the United States Geological Survey (USGS) digital

spectral library (splib06a) [31] is used to build the spectral library A . The reflectance values of 498 spectral signatures are measured for 224 spectral bands distributed uniformly in the interval of $0.4\text{--}2.5\ \mu\text{m}$ ($A \in \mathbb{R}^{224 \times 498}$). Fifteen spectral signatures are chosen from the spectral library to generate our simulated data. Figure 1 shows five spectral signatures used for all the simulated data experiments. The other ten spectral signatures that are not displayed in the figure include Monazite HS255.3B, Zoisite HS347.3B, Rhodochrosite HS67, Neodymium Oxide GDS34, Pigeonite HS199.3B, Meionite WS700.Hlsep, Spodumene HS210.3B, Labradorite HS17.3B, Grossular WS484, and Wollastonite HS348.3B.

In Simulated Data 1 experiment, we generated seven datacubes of 50×50 pixels and 224 bands per pixel, each containing a different number of endmembers: 3, 5, 7, 9, 11, 13, and 15. In each simulated pixel, the fractional abundances of the endmembers were randomly generated following a Dirichlet distribution [23]. Note that there was no pure pixel in Simulated Data 1, and the fractional abundances of the endmembers were less than 0.8. After Simulated Data 1 was generated, Gaussian white noise or correlated noise was added to Simulated Data 1, having different levels of the signal-to-noise ratio (SNR), that is, 20, 25, 30, 35, 40, 45, and 50 dB. The Gaussian white noise was generated using

FIGURE 2: Results on Simulated Data 1 with 30 dB white noise: SRE as function of endmember number.

the `awgn` function in MATLAB. The correlated noise was obtained by low-pass filtering i.i.d. Gaussian noise.

Simulated Data 2 was generated using nine randomly selected spectral signatures from A_1 , with 100×100 pixels and 224 bands per pixel, provided by Dr. M. D. Iordache and Prof. J. M. Bioucas-Dias. The fractional abundances satisfy the ANC and the ASC and are piecewise smoothed. The true abundances of three endmembers are illustrated in Figure 2. After Simulated Data 2 was generated, Gaussian white noise or correlated noise with seven levels of SNR (20, 25, 30, 35, 40, 45, and 50 dB) was added to Simulated Data 2 as well.

In the two simulated data experiments, all the SGA methods have adopted the block-processing strategy to get better performances. The block sizes of the SGA methods were empirically set to 10, and the stopping threshold parameter σ of SOMP and SMP was set to $1 \cdot 10^{-2}$. The stopping threshold parameters σ and δ of BSOMP were set to $1 \cdot 10^{-2}$ and 0.1. RSFoBa was tested using different values of the parameters: λ and σ : 0 and $5 \cdot 10^{-3}$ for Simulated Data 1 experiment; 1 and $1 \cdot 10^{-2}$ for Simulated Data 2 experiment. SUnSAL-TV was tested using different values of the parameters λ and λ_{TV} : $1 \cdot 10^{-2}$ and $1 \cdot 10^{-2}$, $1 \cdot 10^{-2}$ and $1 \cdot 10^{-2}$, $1 \cdot 10^{-2}$ and $5 \cdot 10^{-3}$, $5 \cdot 10^{-3}$ and $1 \cdot 10^{-3}$, $1 \cdot 10^{-3}$ and $5 \cdot 10^{-4}$, $5 \cdot 10^{-4}$ and $1 \cdot 10^{-4}$, and $5 \cdot 10^{-4}$ and $1 \cdot 10^{-4}$ for SNR = 20, 25, 30, 35, 40, 45, and 50 dB. SUnSAL and CLSUnSAL was tested using different values of the parameter λ : $1 \cdot 10^{-2}$, $5 \cdot 10^{-3}$, $1 \cdot 10^{-3}$, $5 \cdot 10^{-4}$, $5 \cdot 10^{-4}$, $5 \cdot 10^{-4}$, and $5 \cdot 10^{-4}$ for SNR = 20, 25, 30, 35, 40, 45, and 50 dB.

Figure 2 shows the SRE (dB) results obtained on Simulated Data 1 with white noise as a function of endmember number, using the different tested methods. The SREs of all the algorithms decrease as the endmember number increases.

FIGURE 3: Results on Simulated Data 1 with white noise when the endmember number is eleven: SRE as function of SNR.

We can find that BSOMP, SOMP, SMP, RSFoBa, and SUnSAL-TV perform better than SUnSAL and CLSUnSAL. All the SGA methods have comparable performances with SUnSAL-TV; BSOMP can always obtain the best estimation of the abundances when the number of endmembers is less than 15. Figure 3 shows the SRE (dB) results obtained on Simulated Data 1 with white noise as a function of SNR when the endmember number is eleven. The SREs of all the algorithms decrease as the SNR decreases. We can find that all the SGA methods perform better than SUnSAL-TV, SUnSAL, and CLSUnSAL. This result indicates that all the SGA methods adopt the block-processing strategy to effectively pick up all the actual endmembers. Among all the SGA methods, BSOMP and RSFoBa behave better than SOMP and SMP, and BSOMP can always obtain the best estimation of the abundances. Figures 4 and 5 show the SRE results obtained on Simulated Data 1 with correlated noise as a function of endmember number and SNR, respectively. As shown in Figures 4 and 5, BSOMP gets an overall better performance than the other algorithms.

Figure 6 shows the number of the potential endmembers obtained from the spectral library by all the SGA methods. The SGA methods adopt the block-processing strategy to effectively pick up all the actual endmembers. It can be observed that BSOMP and RSFoBa can select the actual endmembers more accurately than SOMP and SMP. It is worth mentioning that many endmembers in the estimated endmembers set of SOMP are nonexistent and affect the performance of SOMP. However, BSOMP incorporates a backtracking process to remove some endmembers chosen wrongly from the estimated endmembers set, which can select the actual endmembers more accurately than SOMP.

FIGURE 4: Results on Simulated Data 1 with 30 dB correlated noise: SRE as function of endmember number.

FIGURE 5: Results on Simulated Data 1 with correlated noise when the endmember number is eleven: SRE as function of SNR.

Differing from RSFoBa, BSOMP uses a backtracking process in the whole hyperspectral data to remove some endmembers chosen wrongly from the estimated endmembers set in previous block-processing, which can identify the true endmembers set more accurately than RSFoBa. It indicates that BSOMP behaves better than the other SGA methods.

Figures 7 and 8 show the SRE results obtained on Simulated Data 2 with white noise and correlated noise as a function of SNR, respectively. We can find that BSOMP, SOMP, SMP, RSFoBa, and SUnSAL-TV perform better than SUnSAL and CLSUnSAL. This result indicates that all the SGA methods adopt the block-processing strategy to effectively pick up all the actual endmembers, and SUnSAL-TV combines spatial-contextual coherence regularizer, which can get better performance over SUnSAL and CLSUnSAL. All the SGA methods have comparable performances with SUnSAL-TV; BSOMP and RSFoBa can always obtain the better estimation of the abundances. RSFoBa incorporates the spatial-contextual coherence regularizer to select the actual endmembers more accurately than SOMP and SMP, and BSOMP incorporates a backtracking process in the whole hyperspectral data to identify the true endmembers set more accurately than SOMP and SMP. It can be observed that BSOMP and RSFoBa can select the actual endmembers more accurately than SOMP and SMP.

For visual comparison, Figure 9 shows the true abundance maps and the abundance maps estimated by the different algorithms on Simulated Data 2 with 20 dB white noise. In Figure 9, the abundance maps of endmembers, which were obtained by SUnSAL-TV, contain fewer noise points. However, the abundance maps estimated by SUnSAL-TV may

FIGURE 6: Results obtained by the four SGAs on Simulated Data 1 with 30 dB white noise or correlated noise: number of retained endmembers as function of endmember number.

exhibit an oversmooth visual effect around some pixels; some of the details and edges are not well preserved. Compared with SUnSAL-TV, the estimated abundances by BSOMP may contain more noise points but preserve more edge regions and are closer to the true abundances. Moreover, it is much easier to tell the different materials apart. Compared with SOMP and SMP, the estimated abundances by BSOMP

TABLE 1: Processing times(s) measured after applying the tested methods to the two considered simulated data sets with 30 dB white noise.

Datcube	SUnSAL	CLSunSAL	SUnSAL-TV	SOMP	SMP	RSFoBa	BSOMP
Simulated Data 1 (11 endmembers)	67.67	78.41	363.27	2.083	21.289	3.938	6.418
Simulated Data 2	254.71	279.02	1271.12	9.618	35.444	28.277	106.524

FIGURE 7: Results on Simulated Data 2 with white noise: SRE as function of SNR.

FIGURE 8: Results on Simulated Data 2 with correlated noise: SRE as function of SNR.

TABLE 2: Comparison of the sparse unmixing algorithms with whether the backward step, joint sparsity, and contextual information are considered.

Property	Backward step	Joint sparsity	Contextual information
SMP		.	
SOMP		.	
RSFoBa	.	.	.
BSOMP	.	.	
SUnSAL			
CLSunSAL		.	
SUnSAL-TV			.

contain fewer noise points and preserve more edge regions, which are more accurate and have a better visual effect.

Table 1 shows the processing time measured after applying the tested algorithms to the two simulated data sets with 30 dB white noise. The algorithms were implemented in MATLAB 2009a and executed on a desktop PC equipped with an Intel Core 2 Duo CPU (2.93 GHz) and 2 GB of RAM memory. It can be observed that the SUnSAL-TV algorithm is far more time consuming compared with the other algorithms. We can find that BSOMP incorporating a backtracking technique is more time consuming than SOMP.

In order to make the differences among the BSOMP algorithm and the other sparse unmixing algorithms clearer, we display their comparison concerned with several properties in Table 2.

SUnSAL is a convex relaxation method for sparse unmixing which solves the l_1 norm sparse regression problem. CLSunSAL is convex relaxation method for solving the $l_{2,1}$ optimization problem. The main difference between SUnSAL and CLSunSAL is that the former employs pixelwise independent regressions, while the latter enforces joint sparsity among all the pixels. SUnSAL-TV is the special case of SUnSAL which incorporates the spatial-contextual coherence regularizer into the SUnSAL. As observed in Figures 7 and 8, SUnSAL-TV behaves better than SUnSAL and CLSunSAL, which indicates that the spatial-contextual coherence regularizer can make SUnSAL-TV more effective.

SOMP and SMP are the forward SGAs which select one or more potential endmembers from the spectral library in each iteration. RSFoBa is a forward-backward SGA which incorporates the backward step, spatial-contextual information, and joint sparsity into the greedy algorithm. BSOMP is a forward-backward SGA which incorporates the backward step and joint sparsity to the greedy algorithm. As observed in Figure 6, the numbers of endmembers selected by the SGAs are still larger than the actual number. BSOMP and

FIGURE 9: Continued.

FIGURE 9: Comparison of abundance maps on Synthetic Data 2 with 20 dB white noise. From left column to right column are the abundance maps corresponding to Endmember 1, Endmember 4, and Endmember 9, respectively. From top row to bottom row are true abundance maps or abundance maps obtained by SUnSAL, CLSUnSAL, SUnSAL-TV, SOMP, SMP, RSFoBa, and BSOMP, respectively.

RSFoBa behave better than SMP and SOMP, which indicates that the backward step can make BSOMP and RSFoBa more accurate to select the actual endmembers. Differing from RSFoBa, which adopts a forward step and a backward step to pick some potential endmembers in each block, BSOMP adopts a forward step to pick some potential endmembers in each block and then adopts a backward step in the whole hyperspectral data to remove some endmembers chosen wrongly from the estimated endmembers set. Figure 10 shows the SRE (dB) results obtained on Simulated Data 1 with white noise as a function of block size. We can find that BSOMP performs better than the other SGAs. Figure 11 shows the number of the potential endmembers obtained from the spectral library by all the SGA methods. It can be observed that BSOMP can select the actual endmembers more accurately than RSFoBa. This result indicates that a backward step in the whole hyperspectral data can make BSOMP more effective.

4.2. Evaluation with Real Data. The hyperspectral data set used in our real data experiment is the well-known AVIRIS Cuprite data set (<http://aviris.jpl.nasa.gov/html/aviris.freed-ata.html>). The size of the hyperspectral scene is a 250×191 pixels subset, with 188 spectral bands (water absorption bands and low SNR bands are removed). The spectral library for this data set is the USGS spectral library (containing 498 pure signatures) with the corresponding bands removed. The minerals map (http://speclab.cr.usgs.gov/cuprite95.tgif.2.2um_map.gif) which was produced by a Tricorder 3.3 software (<http://speclab.cr.usgs.gov/PAPERS/tetracorder/>) is shown in Figure 12. Note that the Tricorder map is only suitable for the hyperspectral data collected in 1995; the AVIRIS Cuprite data was collected in 1997. Thus we only use the mineral map as a reference indicator for qualitative analysis of the fractional abundance maps produced by different sparse unmixing methods.

In this section, for the real data experiment, the sparsity of solution and the reconstruction error of the hyperspectral

TABLE 3: Sparsity of the abundance vectors and the reconstruction errors.

Algorithms	SUnSAL	CLSUnSAL	SUnSAL-TV	SOMP	SMP	RSFoBa	BSOMP
Sparsity	17.5629	20.7037	20.472	13.684	15.102	12.702	11.810
RMSE	0.0051	0.0035	0.0028	0.0040	0.0034	0.0032	0.0033

FIGURE 10: Results on Simulated Data 1 with white noise when the endmember number is eleven: SRE as function of block size.

FIGURE 11: Results on Simulated Data 1 with 30 dB white noise when the endmember number is eleven: number of retained endmembers as function of block size.

image are used to evaluate the performances of the sparse unmixing algorithms, since the true abundance maps of the real data are not available. The sparsity is obtained by counting the average number of nonzero abundances estimated by each algorithm. The same as in [22], we define the fractional abundances larger than 0.001 as nonzero abundances to avoid counting negligible values. The reconstruction error is calculated by the original hyperspectral image and the one reconstructed by the actual endmembers and their corresponding abundances [32]. The root mean square error (RMSE) is used to evaluate the reconstruction error between the original hyperspectral image and the reconstructed one:

$$\text{RMSE} = \sqrt{\frac{1}{K} \sum_{j=1}^K (Y_j - \hat{Y}_j)^2}, \quad (14)$$

where Y denotes the original hyperspectral image and $\hat{Y}$ denotes the reconstruction hyperspectral image. If the RMSE of a sparse unmixing algorithm is smaller, it also reflects the unmixing performance of this algorithm.

The regularization parameter used for CLSUnSAL was empirically set to 0.01, while the one corresponding to SUnSAL was set to 0.001. The parameters λ and λ_{TV} of SUnSAL-TV were set to 0.001 and 0.001. The block sizes of the SGA methods were empirically set to 10. The stopping threshold parameter σ of SOMP and SMP was set to 0.01. The stopping threshold parameters σ and δ of BSOMP were set to 0.01 and 0.1. The parameters λ and σ of RSFoBa were set to 1 and 0.01. Table 2 shows the sparsity of the abundance vectors and the reconstruction error obtained by the different algorithm. In Table 3, we can find that BSOMP, SOMP, SMP, and SUnSAL perform better than SUnSAL-TV in the sparsest solutions, although SUnSAL-TV has the lowest reconstruction errors. BSOMP can achieve nearly as few reconstruction errors as SUnSAL-TV but obtain much sparser solutions. In general, BSOMP can obtain the sparsest solutions and achieve the lower reconstruction errors, which indicates the effectiveness of the proposed algorithm.

For visual comparison, Figure 13 shows the estimated fractional abundance maps by applying the proposed BSOMP, SMP, RSFoBa, SOMP, SUnSAL, and SUnSAL-TV algorithms to the AVIRIS Cuprite data for three different minerals (Alunite, Buddingtonite, and Chalcedony). For comparison, the classification maps of three materials extracted from the Tricorder 3.3 software are also presented in Figure 13. It can be observed that the abundances estimated by BSOMP are comparable or higher in the regions assigned to the respective minerals in comparison to the other considered algorithms, with regard to the classification maps produced by the Tricorder 3.3 software. We also analyzed the sparsity of the abundances estimated by BSOMP; the average

FIGURE 12: USGS map showing the distribution of different minerals in the Cuprite mining district in Nevada.

number of endmembers with abundances higher than 0.001 estimated by BSOMP is 11.810 (per pixel). This means that the algorithm uses a minimum number of members to explain the data. Overall, the qualitative results and visual results reported in this section indicate that the BSOMP algorithm is more effective for sparse unmixing than the other algorithms.

5. Conclusions

In this paper, we present a new SGA termed BSOMP for sparse unmixing of hyperspectral data. The main idea of the proposed algorithm is that BSOMP incorporates a backtrack step to detect the previous chosen endmembers' reliability and then deletes the unreliable endmembers. It is proved that BSOMP can recover the optimal endmembers from the spectral library under certain conditions. Experiments on both synthetic data and real data also demonstrate that the BSOMP algorithm is more effective than the other SGAs for sparse unmixing.

The topic we want to discuss is why BSOMP is effective for sparse unmixing. Endmember selection and abundance estimation are the two main parts of the SGAs. The number of the potential endmembers selected by the SGAs

is far smaller than the number of spectral signatures in the spectral library, and then the abundance estimation method makes the abundance vectors sparse further with the relatively small endmembers subset. The main difference among BSOMP, SOMP, SMP, and RSFoBa is the endmember selection. The SGAs adopt a block-processing strategy to efficiently select endmembers, but the numbers of endmembers selected are still larger than the actual number. However the nonexisting endmembers will have a negative effect on the estimation of the abundances corresponding to the actual endmembers.

Compared with SOMP and SMP, BSOMP and RSFoBa adopt a backward step to detect the previous chosen endmembers' reliability and delete the nonexisting endmembers. So BSOMP and RSFoBa can select the actual endmembers more accurately than SOMP and SMP. The main difference between BSOMP and RSFoBa is that the former adopts a backward step in the whole hyperspectral data, while the latter adopts a backward step in each block. As shown in Figures 10 and 11, we can observe that the block size will significantly affect the number of the endmembers obtained by RSFoBa, while BSOMP is not affected by the block size. In conclusion, BSOMP can select the actual endmembers more accurately than the other SGAs and be beneficial to improve estimation effectiveness.

FIGURE 13: Continued.

FIGURE 13: Abundance maps estimated by SUNSAL, CLSUnSAL, SUNSAL-TV, SOMP, SMP, RSFoBa, and BSOMP and the distribution maps produced by Tricorder 3.3 software for the AVIRIS Cuprite scene. From top row to bottom row are the maps produced or estimated by Tricorder software, SUNSAL, CLSUnSAL, SUNSAL-TV, SOMP, SMP, RSFoBa, and BSOMP, respectively. From left column to right column are the maps corresponding to Alunite, Buddingtonite, and Chalcedony, respectively.

Appendix

Proof of Theorem 7

First we should introduce some definitions which will be used during the proof process.

Definition A.1. Given hyperspectral image data matrix $Y \in R^{L \times K}$ and available spectral library matrix $A \in R^{L \times m}$,

$\tilde{A}_t$ is a submatrix of the spectral library A , where A_t contains the k columns indexed in S_t , $\tilde{A}_t$ contains the remaining columns, $A = [A_t, \tilde{A}_t]$;

$A_{\sim \text{opt}}$ is a submatrix of A_t , which contains the remaining $(k - T)$ columns indexed in $S_t \setminus S_{\text{opt}}$;

$X_{\sim \text{opt}}$ is a submatrix of X_t and contains the rows listed in $S_t \setminus S_{\text{opt}}$.

Now we will prove the theorem on the performance of the algorithm.

Proof. Let P_{opt} denote the orthogonal projector onto the range of A_{opt} :

$$P_{\text{opt}} = A_{\text{opt}} (A_{\text{opt}}^* A_{\text{opt}})^{-1} A_{\text{opt}}^*, \quad (\text{A.1})$$

where A_{opt}^* denotes the conjugate transpose of A_{opt} . Since the columns of A_t are orthogonal to the columns of R_t ($R_t = Y - Y_t$), we have

$$\|A^* R_t\|_{\infty, \infty} = \|\tilde{A}_t^* R_t\|_{\infty, \infty}. \quad (\text{A.2})$$

The maximum correlation between $\tilde{A}_t$ (a submatrix of the spectral library A) and the residual $R_{\text{opt}} = Y - Y_{\text{opt}}$ is smaller

than the maximum correlation between the spectral library A and the residual R_{opt} , so we have

$$\|\tilde{A}_t^* (Y - Y_{\text{opt}})\|_{\infty, \infty} \leq \|A^* (Y - Y_{\text{opt}})\|_{\infty, \infty}. \quad (\text{A.3})$$

Invoke (A.2) and (A.3) and apply the triangle inequality to get

$$\begin{aligned} \|A^* R_t\|_{\infty, \infty} &= \|\tilde{A}_t^* (Y - Y_{\text{opt}} + Y_{\text{opt}} - Y_t)\|_{\infty, \infty} \\ &\leq \|A^* (Y - Y_{\text{opt}})\|_{\infty, \infty} \\ &\quad + \|\tilde{A}_t^* (Y_t - Y_{\text{opt}})\|_{\infty, \infty}. \end{aligned} \quad (\text{A.4})$$

Then the matrix $Y_t - Y_{\text{opt}}$ can be expressed in the following manner:

$$\begin{aligned} Y_t - Y_{\text{opt}} &= (I - P_{\text{opt}}) Y_t = (I - P_{\text{opt}}) A_t X_t \\ &= (I - P_{\text{opt}}) A_{\sim \text{opt}} X_{\sim \text{opt}}. \end{aligned} \quad (\text{A.5})$$

The last equality holds because $(I - P_{\text{opt}})$ is orthogonal to the range of A_{opt} . Applying the triangle inequality, the second term of (A.4) can be rewritten as follows:

$$\begin{aligned} \|\tilde{A}_t^* (Y_t - Y_{\text{opt}})\|_{\infty, \infty} &= \|\tilde{A}_t^* (I - P_{\text{opt}}) A_{\sim \text{opt}} X_{\sim \text{opt}}\|_{\infty, \infty} \\ &\leq \left[\|\tilde{A}_t^* A_{\sim \text{opt}}\|_{\infty, \infty} + \|\tilde{A}_t^* P_{\text{opt}} A_{\sim \text{opt}}\|_{\infty, \infty} \right] \\ &\quad \times \|X_{\sim \text{opt}}\|_{\infty, \infty}. \end{aligned} \quad (\text{A.6})$$

We will apply the cumulative coherence function to deduce the estimates of (A.6); the deduction has five steps.

(1) Each row of the matrix $\tilde{A}_t^* A_{\sim \text{opt}}$ lists the inner products between a spectral endmember and $(k - T)$ distinct endmembers; we have

$$\|\tilde{A}_t^* A_{\sim \text{opt}}\|_{\infty, \infty} \leq \mu(k - T). \quad (\text{A.7})$$

(2) Use the fact that $\|\cdot\|_{\infty, \infty}$ is submultiplicative and invoke (A.1) to get

$$\begin{aligned} \|\tilde{A}_t^* P_{\text{opt}} A_{\sim \text{opt}}\|_{\infty, \infty} &\leq \|\tilde{A}_t^* A_{\text{opt}}\|_{\infty, \infty} \\ &\quad \times \|(A_{\text{opt}}^* A_{\text{opt}})^{-1}\|_{\infty, \infty} \\ &\quad \times \|A_{\text{opt}}^* A_{\sim \text{opt}}\|_{\infty, \infty}. \end{aligned} \quad (\text{A.8})$$

(3) Making use of the same deduction as in step (1), we have

$$\begin{aligned} \|\tilde{A}_t^* A_{\text{opt}}\|_{\infty, \infty} &\leq \mu(T), \\ \|A_{\text{opt}}^* A_{\sim \text{opt}}\|_{\infty, \infty} &\leq \mu(k - T). \end{aligned} \quad (\text{A.9})$$

(4) Since the columns of A are normalized, the Gram matrix $A_{\text{opt}}^* A_{\text{opt}}$ lists the inner products between T different

endmembers, so the matrix $A_{\text{opt}}^* A_{\text{opt}}$ has a unit diagonal, and then we can get

$$A_{\text{opt}}^* A_{\text{opt}} = I - X_{\text{opt}}, \quad (\text{A.10})$$

where $\|X_{\text{opt}}\|_{\infty, \infty} \leq \mu(T)$; then we can expand the inverse in a Neumann series to get

$$\begin{aligned} \|(A_{\text{opt}}^* A_{\text{opt}})^{-1}\|_{\infty, \infty} &= \|(I - X_{\text{opt}})^{-1}\|_{\infty, \infty} \\ &\leq \frac{1}{1 - \|X_{\text{opt}}\|_{\infty, \infty}} \leq \frac{1}{1 - \mu(T)}. \end{aligned} \quad (\text{A.11})$$

(5) Invoke the bounds from steps (1) to (4) into (A.6); we can obtain that

$$\begin{aligned} \|\tilde{A}_t^* (Y_t - Y_{\text{opt}})\|_{\infty, \infty} &\leq \left[\mu(k - T) + \frac{\mu(T) \mu(k - T)}{1 - \mu(T)} \right] \times \|X_{\sim \text{opt}}\|_{\infty, \infty}. \end{aligned} \quad (\text{A.12})$$

Simplify expression (A.12) to conclude that

$$\|\tilde{A}_t^* (Y_t - Y_{\text{opt}})\|_{\infty, \infty} \leq \frac{\mu(k - T)}{1 - \mu(T)} \|X_{\sim \text{opt}}\|_{\infty, \infty}. \quad (\text{A.13})$$

In order to produce the an upper bound of $\|X_{\sim \text{opt}}\|_{\infty, \infty}$, we deduce the following formula by using the triangle inequality:

$$\begin{aligned} \|A_t^* (Y - Y_{\text{opt}})\|_{\infty, \infty} &= \|A_t^* (Y - Y_t + Y_{\text{opt}} - Y_t)\|_{\infty, \infty} \\ &\leq \|A_t^* (Y - Y_t)\|_{\infty, \infty} + \|A_t^* (Y_t - Y_{\text{opt}})\|_{\infty, \infty} \\ &\leq \|A_t^* (Y_t - Y_{\text{opt}})\|_{\infty, \infty}. \end{aligned} \quad (\text{A.14})$$

Since $(Y - Y_t)$ is orthogonal to the range of A_t . Introduce (A.5) into (A.14); we can get

$$\begin{aligned} \|A_t^* (Y_t - Y_{\text{opt}})\|_{\infty, \infty} &= \|A_{\sim \text{opt}}^* (I - P_{\text{opt}}) A_{\sim \text{opt}} X_{\sim \text{opt}}\|_{\infty, \infty}. \end{aligned} \quad (\text{A.15})$$

The equality holds because $(I - P_{\text{opt}})$ is orthogonal to the range of A_{opt} ; we can replace A_t by $A_{\sim \text{opt}}$. Let $D = A_{\sim \text{opt}}^* (I - P_{\text{opt}}) A_{\sim \text{opt}}$ and rewrite $D = I - (I - D)$; we can expand its inverse in a Neumann series to obtain

$$\|DX_{\sim \text{opt}}\|_{\infty, \infty} \geq (1 - \|I - D\|_{\infty, \infty}) \|X_{\sim \text{opt}}\|_{\infty, \infty}. \quad (\text{A.16})$$

Recall the definition of D , and apply the triangle inequality to obtain

$$\begin{aligned} \|I - D\|_{\infty, \infty} &= \|I - A_{\sim \text{opt}}^* A_{\sim \text{opt}} + A_{\sim \text{opt}}^* P_{\text{opt}} A_{\sim \text{opt}}\|_{\infty, \infty} \\ &\leq \|I - A_{\sim \text{opt}}^* A_{\sim \text{opt}}\|_{\infty, \infty} \\ &\quad + \|A_{\sim \text{opt}}^* P_{\text{opt}} A_{\sim \text{opt}}\|_{\infty, \infty}. \end{aligned} \quad (\text{A.17})$$

The right-hand side of inequality can be deduced completely analogous with steps (1)–(5), we can obtain

$$\begin{aligned} \|I - D\|_{\infty, \infty} &\leq \mu(k - T) + \frac{\mu(T)\mu(k - T)}{1 - \mu(T)} \\ &= \frac{\mu(k - T)}{1 - \mu(T)}. \end{aligned} \quad (\text{A.18})$$

Introduce (A.18) and (A.16) into (A.14); we have

$$\begin{aligned} \|A_t^*(Y - Y_{\text{opt}})\|_{\infty, \infty} \\ \geq \left[1 - \frac{\mu(k - T)}{1 - \mu(T)}\right] \|X_{\sim \text{opt}}\|_{\infty, \infty}. \end{aligned} \quad (\text{A.19})$$

Using the same deduction as (A.3), it follows that $\|A_t^*(Y - Y_{\text{opt}})\|_{\infty, \infty} \leq \|A^*(Y - Y_{\text{opt}})\|_{\infty, \infty}$. So (A.19) can be rewritten as follows:

$$\begin{aligned} \|A^*(Y - Y_{\text{opt}})\|_{\infty, \infty} \\ \geq \left[1 - \frac{\mu(k - T)}{1 - \mu(T)}\right] \|X_{\sim \text{opt}}\|_{\infty, \infty}. \end{aligned} \quad (\text{A.20})$$

Introduce (A.20) into (A.13); we can get

$$\begin{aligned} \|\tilde{A}_t^*(Y_t - Y_{\text{opt}})\|_{\infty, \infty} \\ \leq \frac{\mu(k - T)}{1 - \mu(T) - \mu(k - T)} \|A^*(Y - Y_{\text{opt}})\|_{\infty, \infty}. \end{aligned} \quad (\text{A.21})$$

Introduce (A.21) into (A.4) to get

$$\begin{aligned} \|A^*R_t\|_{\infty, \infty} \\ \leq \frac{1 - \mu(T)}{1 - \mu(T) - \mu(k - T)} \|A^*(Y - Y_{\text{opt}})\|_{\infty, \infty}. \end{aligned} \quad (\text{A.22})$$

Finally, assume we have a bound $\|A^*(Y - Y_{\text{opt}})\|_{\infty, \infty} \leq \tau$, so (A.22) can be rewritten as follows:

$$\|A^*R_t\|_{\infty, \infty} \leq \frac{1 - \mu(T)}{1 - \mu(T) - \mu(k - T)} \tau. \quad (\text{A.23})$$

This completes the argument. $\square$

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work has been supported by a grant from the NUAA Fundamental Research Funds (NS2013085). The authors would like to thank M. D. Iordache, J. Bioucas-Dias, and A. Plaza for sharing Simulated Data 2 and their codes for the algorithms of CLSUnSAL, SUnSAL, and SUnSAL-TV. The authors would also like to thank Shi Z, Tang W, Duren Z, and Jiang Z for sharing their codes for the algorithms of SMP and RSFoBa. Moreover, the authors would also like to thank those anonymous reviewers for their helpful comments to improve this paper.

References

- [1] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 44–57, 2002.
- [2] Y. H. Hu, H. B. Lee, and F. L. Scarpace, "Optimal linear spectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 1, pp. 639–644, 1999.
- [3] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon et al., "Hyperspectral unmixing overview: geometrical, statistical, and sparse regression-based approaches," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 354–379, 2012.
- [4] M. Graña, I. Villaverde, J. O. Maldonado, and C. Hernandez, "Two lattice computing approaches for the unsupervised segmentation of hyperspectral images," *Neurocomputing*, vol. 72, no. 10–12, pp. 2111–2120, 2009.
- [5] J. Li and J. M. Bioucas-Dias, "Minimum volume simplex analysis: a fast algorithm to unmix hyperspectral data," in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS '08)*, pp. III-250–III-253, Boston, Mass, USA, July 2008.
- [6] H. Pu, B. Wang, and L. Zhang, "Simplex geometry-based abundance estimation algorithm for hyperspectral unmixing," *Scientia Sinica Informationis*, vol. 42, no. 8, pp. 1019–1033, 2012.
- [7] G. Zhou, S. Xie, Z. Yang, J.-M. Yang, and Z. He, "Minimum-volume-constrained nonnegative matrix factorization: enhanced ability of learning parts," *IEEE Transactions on Neural Networks*, vol. 22, no. 10, pp. 1626–1637, 2011.
- [8] T.-H. Chan, C.-Y. Chi, Y.-M. Huang, and W.-K. Ma, "A convex analysis-based minimum-volume enclosing simplex algorithm for hyperspectral unmixing," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4418–4432, 2009.
- [9] M. Arngren, M. N. Schmidt, and J. Larsen, "Bayesian nonnegative matrix factorization with volume prior for unmixing of hyperspectral images," in *Proceedings of the IEEE International Workshop on Machine Learning for Signal Processing (MLSP '09)*, pp. 1–6, IEEE, Grenoble, France, September 2009.
- [10] V. P. Pauca, J. Piper, and R. J. Plemmons, "Nonnegative matrix factorization for spectral data analysis," *Linear Algebra and Its Applications*, vol. 416, no. 1, pp. 29–47, 2006.
- [11] N. Wang, B. Du, and L. Zhang, "An endmember dissimilarity constrained non-negative matrix factorization method for hyperspectral unmixing," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 2, pp. 554–569, 2013.
- [12] Z. Wu, S. Ye, J. Liu, L. Sun, and Z. Wei, "Sparse non-negative matrix factorization on GPUs for hyperspectral unmixing," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 8, pp. 3640–3649, 2014.
- [13] J. B. Greer, "Sparse demixing of hyperspectral images," *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 219–228, 2012.
- [14] J. A. Tropp and S. J. Wright, "Computational methods for sparse solution of linear inverse problems," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 948–958, 2010.
- [15] M. Elad, *Sparse and Redundant Representations*, Springer, New York, NY, USA, 2010.
- [16] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.

- [17] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Transactions on Information Theory*, vol. 55, no. 5, pp. 2230–2249, 2009.
- [18] I. Marques and M. Graña, "Hybrid sparse linear and lattice method for hyperspectral image unmixing," in *Hybrid Artificial Intelligence Systems*, vol. 8480 of *Lecture Notes in Computer Science*, pp. 266–273, Springer, 2014.
- [19] I. Marques and M. Graña, "Greedy sparsification WM algorithm for endmember induction in hyperspectral images," in *Natural and Artificial Computation in Engineering and Medical Applications*, vol. 7931 of *Lecture Notes in Computer Science*, pp. 336–344, Springer, Berlin, Germany, 2013.
- [20] M. V. Afonso, J. M. Bioucas-Dias, and M. A. Figueiredo, "An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Transactions on Image Processing*, vol. 20, no. 3, pp. 681–695, 2011.
- [21] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Total variation spatial regularization for sparse hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 11, pp. 4484–4502, 2012.
- [22] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Collaborative sparse unmixing of hyperspectral data," in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS '12)*, pp. 7488–7491, IEEE, Munich, Germany, July 2012.
- [23] J. M. Bioucas-Dias and M. A. T. Figueiredo, "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 '10)*, pp. 1–4, Reykjavik, Iceland, June 2010.
- [24] Z. Shi, W. Tang, Z. Duren, and Z. Jiang, "Subspace matching pursuit for sparse unmixing of hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 6, pp. 3256–3274, 2014.
- [25] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Algorithms for simultaneous sparse approximation. Part I: greedy pursuit," *Signal Processing*, vol. 86, no. 3, pp. 572–588, 2006.
- [26] W. Tang, Z. Shi, and Y. Wu, "Regularized simultaneous forward-backward greedy algorithm for sparse unmixing of hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 9, pp. 5271–5288, 2014.
- [27] J. Chen and X. Huo, "Theoretical results on sparse representations of multiple-measurement vectors," *IEEE Transactions on Signal Processing*, vol. 54, no. 12, pp. 4634–4643, 2006.
- [28] G. Davis, S. Mallat, and M. Avellaneda, "Greedy adaptive approximation," *Journal of Constructive Approximation*, vol. 13, no. 1, pp. 57–98, 1997.
- [29] J. A. Tropp, "Greed is good: algorithmic results for sparse approximation," *IEEE Transactions on Information Theory*, vol. 50, no. 10, pp. 2231–2242, 2004.
- [30] D. L. Donoho and M. Elad, "Maximal sparsity representation via l_1 minimization," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, pp. 2197–2202, 2003.
- [31] R. N. Clark, G. A. Swayze, R. Wise et al., *USGS Digital Spectral Library splib06a*, U.S. Geological Survey, Denver, Colo, USA, 2007.
- [32] J. Plaza, A. Plaza, P. Martínez, and R. Pérez, "H-COMP: a tool for quantitative and comparative analysis of endmember identification algorithms," in *Proceedings of the Geoscience and Remote Sensing Symposium (IGARSS '03)*, vol. 1, pp. 291–293, 2003.