

Journal of
Applied Remote Sensing

RemoteSensing.SPIEDigitalLibrary.org

Exploring error of linear mixed model for hyperspectral image reconstruction from spectral compressive sensing

Zhongliang Wang
Mi He
Zhen Ye
Yongjian Nian
Liang Qiao
Mingsheng Chen

SPIE.

Zhongliang Wang, Mi He, Zhen Ye, Yongjian Nian, Liang Qiao, Mingsheng Chen, "Exploring error of linear mixed model for hyperspectral image reconstruction from spectral compressive sensing," *J. Appl. Remote Sens.* **13**(3), 036514 (2019), doi: 10.1117/1.JRS.13.036514.

Exploring error of linear mixed model for hyperspectral image reconstruction from spectral compressive sensing

Zhongliang Wang,^a Mi He,^b Zhen Ye,^c Yongjian Nian,^{b,*} Liang Qiao,^b and Mingsheng Chen^b

^aTongling University, Department of Electrical Engineering, Tongling, Anhui, China

^bArmy Medical University (Third Military Medical University),

College of Biomedical Engineering and Imaging Medicine, Chongqing, China

^cChang'an University, School of Electronics and Control Engineering, Xi'an, China

Abstract. Linear mixed model (LMM) has been extensively applied for hyperspectral compressive sensing (CS) in recent years. However, the error introduced by LMM that limits the reconstruction performance has not been given full consideration. We propose an algorithm for hyperspectral CS based on LMM under the assumption of known endmembers. At the sampling stage, only spectral compressive sampling is carried out to keep the abundance information as much as possible. At the reconstruction stage, the proposed algorithm estimates abundance by using linear unmixing from the spectral observed data. Moreover, the model error introduced by LMM is explored; a joint convex optimization scheme for estimation of both abundance and model error is established and solved by the alternating iteration approach to achieve the optimal reconstruction. Experimental results on a real hyperspectral dataset demonstrate that the proposed algorithm significantly outperforms the other state-of-the-art hyperspectral CS algorithms. © The Authors. Published by SPIE under a Creative Commons Attribution 4.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JRS.13.036514](https://doi.org/10.1117/1.JRS.13.036514)]

Keywords: hyperspectral images; compressive sensing; spectral unmixing; linear mixture model.

Paper 190362 received May 15, 2019; accepted for publication Sep. 5, 2019; published online Sep. 23, 2019.

1 Introduction

In recent years, hyperspectral images (HSIs) have been widely used in the fields of ground classification, environmental monitoring, military reconnaissance, and so on, since they contain abundant spatial and spectral information of imaging objects. However, with the continuous improvement of both spatial and spectral resolutions of imaging spectrometers, the amount of acquired hyperspectral data increases exponentially, which brings great challenges to the data storage and transmission. Compressive sensing (CS) technique integrates data acquisition and compression, which collect the sparse or compressible signal at a much lower sampling rate than Nyquist sampling theorem and reconstruct the original signal with a higher probability.¹ At present, CS has been successfully applied in HSI compression.²⁻⁶ Hyperspectral CS only captures a small number of sampling data of the original HSI by using incoherent measurement for reconstruction. Compared with traditional compression methods, hyperspectral CS significantly reduces the amount of data that needs to be collected; moreover, it also reduces the resource consumption of imaging. Therefore, hyperspectral CS has attracted remarkable attention in the field of hyperspectral remote sensing.

Hyperspectral CS has two key problems: one is the design of the measurement matrix and the other is the reconstruction algorithm. The core of the former is how to design the measurement matrix so that it can retain the effective information of the original HSI for reconstruction as much as possible; the latter is how to reconstruct the target HSI based on the observed data.

*Address all correspondence to Yongjian Nian, E-mail: [ynjian@126.com](mailto:yjnian@126.com)

In general, the measurement matrix is required to be incoherent with the original HSI, where Gauss random matrix as the measurement matrix is a better choice.⁷ It should be pointed out that the design of the measurement matrix is beyond the scope of this paper; we mainly focus on the reconstruction of hyperspectral CS. At present, many popular reconstruction algorithms have been proposed. Noor and Jacobs⁸ proposed a video imaging measurement scheme for the static scene by exploiting sparsity along the time dimension. After acquiring all measurements required for the first frame, the measurements are acquired only from the areas that change in subsequent frames. In Ref. 9, block compressed sensing-smooth projected Landweber (BCS-SPL) employs a block-based random sampling and a projected Landweber reconstruction based on the sparsity in the directional transform domain with a smooth reconstructed image. Note that convex optimization combined with certain prior knowledge has been widely used in the reconstruction of hyperspectral CS. In Ref. 10, by taking advantage of convex optimization method, three-dimensional compressive sampling (3-D CS) performed the reconstruction by exploiting the sparse, low-rank, and three-dimensional (3-D) smoothing characteristics of HSI and achieved good reconstruction results. In Ref. 11, the structure similarity property was explored to improve the reconstruction accuracy beyond the prior hypothesis of low rank, spatial correlation, and spatial smoothness. Zhang et al.² proposed a reweighted Laplace prior-based hyperspectral CS method to introduce structure information into sparsity prior, which can further improve the reconstruction accuracy. Wang et al.³ adopted the tensor analysis to model spatial correlation and local smoothness, where the tensor tucker decomposition was used to describe the global spatial and spectral correlation and the weighted 3-D total variance (TV) was used to characterize the local smooth image. Another kind of reconstruction algorithm is based on the idea of matrix decomposition using the statistical characteristics. The typical algorithm is the compressive projection principal component analysis (CPPCA)¹² and its derivative algorithms.^{13–15} CPPCA shifts the computational burden of principal component analysis from the resource-constrained encoder to the base-station decoder, which significantly improves the computational efficiency.

Linear mixed model (LMM) is explicitly physically meaningful and is usually a good approximation for HSI, in spite of the existence of nonlinear mixing effects.¹⁶ In recent years, regarded as an effective priori information, LMM has been successfully applied in hyperspectral CS for reconstruction with high accuracy, which decomposes the HSI into some endmembers and an abundance matrix.¹⁷ Based on LMM, the original HSI can be easily reconstructed by the product of endmembers and abundance. With a known spectral library of endmembers, Ramirez et al.¹⁸ proposed a spectral unmixing algorithm directly from the observed data acquired by the coded aperture snapshot spectral imaging system, which shows that estimating abundance from the observed data is completely feasible. In Ref. 19, the theoretical lower bound of the observed values for compressed sensing reconstruction using LMM was given, and the abundance was estimated by using the sparse prior under the assumption that the endmembers are known. Assuming that the endmembers are given, Martín et al.¹⁶ developed a framework for hyperspectral CS called hyperspectral coded aperture, where abundance can be estimated by solving a convex optimization problem including a data term connected to the measurement matrix and a TV regularization term. Instead of using the given endmembers, Zhang et al.⁴ proposed a locally similar sparsity-based hyperspectral CS algorithm to decompose the HSI with an established redundant endmember library. Martín and Bioucas-Dias⁵ proposed a blind hyperspectral reconstruction technique termed spectral compressive acquisition, which supposes that HSI belongs to a low-dimensional subspace that can be learned from the observed data. In our previous work,²⁰ we developed a special kind of measurement matrix that projects the original HSI into spatial and spectral observed data for estimation of endmembers and abundance, respectively, and then spectral unmixing was performed for the reconstruction based on LMM. Based on Ref. 20, a joint optimization model was established to estimate endmembers and abundance, which obtains better reconstruction performance.⁶

Although the above reconstruction algorithms, by using linear unmixing, have achieved good reconstruction, they all neglected the error introduced by the model itself. In fact, nonlinear mixing still occurs in the spectral mixture model due to multiple reflections from a surface. Therefore, the model error of LMM should be given more attention. In our study, we also work under the assumption that endmembers are known or may be estimated from the original HSI. The main contribution of this paper is to propose a hyperspectral CS reconstruction algorithm

based on LMM, in which only spectral sampling is performed on the original HSI to obtain the observed data, and then spectral unmixing is employed for the reconstruction of target HSI based on LMM. In particular, the model error is explored from the observed data by using convex optimization with the TV regularization term, and a joint optimization scheme for estimating abundance and model error is proposed and solved by alternate iteration to obtain the optimal reconstruction quality. Experimental results demonstrate the effectiveness of the proposed algorithm.

The rest of this paper is organized as follows. In Sec. 2, the framework of the proposed algorithm is described. In Sec. 3, the compressive sampling scheme for the proposed algorithm is given. In Sec. 4, the joint optimization scheme for estimating abundance and model error is presented, and in Sec. 5, experimental results on real hyperspectral data are shown. Finally, Sec. 6 gives some concluding remarks.

2 Framework of the Proposed Algorithm

In general, HSI can be denoted as $X = [X_1, X_2, \dots, X_L]^T \in \mathbb{R}^{L \times N}$, where N is the total number of spatial pixels in each band, L is the number of spectral bands, and $X_l (1 \leq l \leq L)$ is the vector representation of the l 'th band. As one of the simplest and most effective hypotheses, LMM has been widely used in the spectral unmixing of HSI, which assumes that the abundance represents the relative region associated to certain endmember, thus the spectrum is modeled as a linear combination of several endmembers weighted by the corresponding abundances. Let $E \in \mathbb{R}^{L \times p}$ be the endmember matrix and $S \in \mathbb{R}^{p \times N}$ be the abundance matrix, where p is the number of endmember in HSI. According to LMM, HSI can be written as

$$X = ES + W, \quad (1)$$

where W represents the model error. As seen from Eq. (1), the reconstruction of the target HSI can be converted to the reconstruction of endmembers, abundance, and model error. At present, a large number of advanced endmember extraction techniques have been proposed. On the other hand, many spectral libraries have been established in succession, such as United States Geological Survey and Jet Propulsion Laboratory. Therefore, it is reasonable to suppose E to be known or estimated from the original HSI. On the premise that E is known, the main goal of our study is to estimate the abundance matrix S and error matrix W from the observed data. As for the sampling scheme of the proposed algorithm, it should be pointed out that spatial sampling that only performs measurement within each band may significantly destroy the abundance information, and spectral sampling is recommended because it can keep abundance information more completely. Moreover, the reconstruction by spectral sampling can make full use of the strong spectral correlation, which is important for improving reconstruction performance. According to the above description, the framework of the proposed algorithm is given in Fig. 1. At the sampling stage, the spectral compressive sampling is performed on X to obtain the observed data Y , where A is the measurement matrix. At the reconstruction stage, the initial values of both S and W are first determined, and then the joint convex optimization method is established to perform alternating iterative solution on S and W ; once their optimal values are obtained, the final reconstruction of the target HSI can be achieved according to LMM. In this paper, the proposed algorithm is denoted as spectral compressed reconstruction based on spectral unmixing and error compensation (SCR-SUEC), with more details given in Secs. 3 and 4.

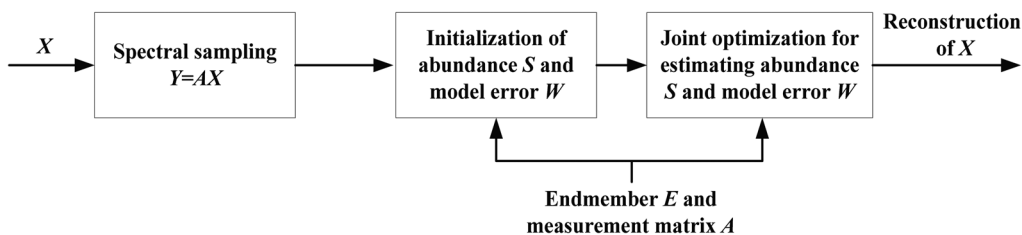


Fig. 1 Framework of the proposed algorithm.

3 Spectral Compressive Sampling

The proposed SCR-SUEC performs compressive sampling on the spectral vector at each pixel to obtain available information for the following reconstruction. Let A be the measurement matrix with size of $J \times L$ ($J \ll L$) and $Y \in \mathbb{R}^{J \times N}$ be the observed data. Combined with Eq. (1), the spectral sampling processing can be expressed as

$$Y = AX = AES + AW. \tag{2}$$

It is obvious that Y has the same size with X in spatial dimension. However, the spectral dimension of Y is reduced by A compared with that of X , where the degree of reduction depends on J value. The sampling rate can be calculated by J/L , where a larger J leads to a high sampling rate and a small J leads to a low sampling rate. In this paper, the measurement matrix A is generated by rounding a Gaussian random sampling matrix to achieve a binary matrix according to the desired sampling rates.

The hardware designs for CS applications have been paid more and more attention. Note that digital micromirror device (DMD) array has been successfully applied for compressive sampling. Based on the works in Refs. 5 and 21, the optical design of the proposed algorithm on a pushbroom platform is given in Fig. 2, which mainly contains four elements: the first one is the grating that splits the reflected light from the object into different wavelengths; the second one is the cylindrical lens that can convert the divergent light at different wavelength into parallel light; the third one is the DMD that generates the measurement matrix controlled by a specific program; and, finally, the cylindrical lens sum the reflected light from the DMD to obtain the observed data by using the linear charge-coupled device array.

4 Proposed Reconstruction Algorithm

4.1 Analysis of Model Error

As mentioned earlier, only spectral sampling is performed on the original HSI, and the abundance information can be well preserved by this sampling manner. According to Eq. (1), if we want to estimate S , it is necessary to obtain the estimation of W first. To deal with this problem, W is temporarily neglected; thus, Eq. (1) becomes $X = ES$. In this case, the spectral sampling processing can be rewritten as

$$Y = AX = AES. \tag{3}$$

Because the observed data Y , measurement matrix A , and endmember matrix E are all known, it is easy to obtain the estimation of S , according to Eq. (3), by using least squares (LS) method as

$$\hat{S} = [(AE)^T AE]^{-1} (AE)^T Y. \tag{4}$$

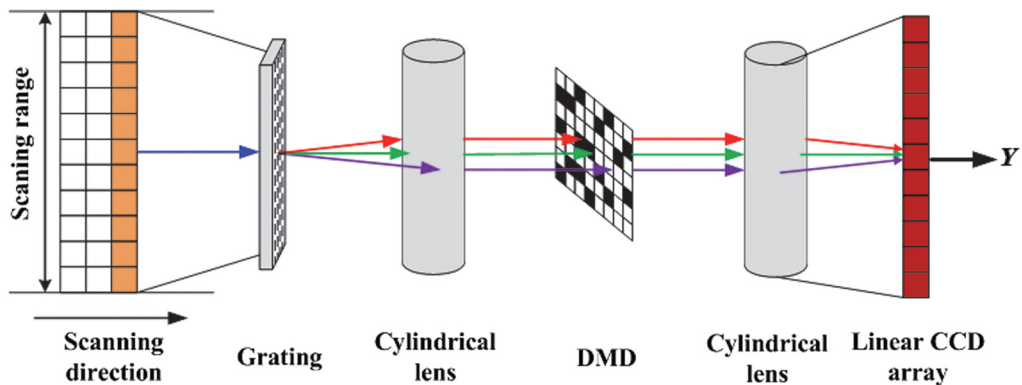


Fig. 2 A simple scheme of optical design on a pushbroom platform.

It must be pointed out that LMM is only an approximation model of HSI, and the use of LMM to describe HSI will not escape from occurrence of error. Note that \mathbf{S} estimated by Eq. (4) is obtained under the condition of neglecting error term of LMM; thus, the final reconstruction can be achieved directly by multiplying \mathbf{E} by \mathbf{S} . However, there must be a certain degree of inevitable model error between this reconstruction and the original one. Unfortunately, the model error of LMM has not been taken into account in the existing algorithms. In this paper, the proposed SCR-SUEC further explores the model error of LMM from spectral observed data, which is of great significance to improve the final reconstruction performance. Note that it is a nondeterministic polynomial-hard problem to calculate \mathbf{S} and \mathbf{W} from \mathbf{Y} simultaneously. Nevertheless, we can utilize alternating iterations to tackle this problem. According to Eq. (2), once \mathbf{S} is estimated, since \mathbf{A} and \mathbf{E} have been known, we can easily obtain the observed data of \mathbf{W} (i.e., \mathbf{AW}). Then, it is possible to estimate \mathbf{W} from \mathbf{AW} by using optimization method with reasonable regularization. To improve the estimation precision, the estimation of \mathbf{W} should be further used to update \mathbf{S} . Therefore, it is necessary to propose a joint optimization scheme to calculate the optimal estimation of both \mathbf{S} and \mathbf{W} .

4.2 Joint Optimization for Abundance and Model Error

In this paper, a joint optimization scheme utilizing an alternating iterative manner to calculate \mathbf{S} and \mathbf{W} is proposed. That is, first assume that \mathbf{W} is fixed and solve \mathbf{S} ; then fix \mathbf{S} and solve \mathbf{W} ; alternate iteration continues until the stopping criterion is satisfied.

4.2.1 Estimation of abundance

Given the measurement matrix \mathbf{A} and the endmember matrix \mathbf{E} , the abundance matrix \mathbf{S} can be estimated from the spectral observed data \mathbf{Y} , subject to LMM. If the sampling rate is high enough, the dimension of the subspace of original HSI is smaller than the spectral dimension of the observed data \mathbf{Y} , that is, $p < J$. In this case, the estimation of \mathbf{S} is actually an over-determined problem. We can simply estimate \mathbf{S} by solving the following constrained optimization problem:

$$\min_{\mathbf{S}} \|\mathbf{AX} - \mathbf{Y}\|_F^2 \quad \text{s.t. } \mathbf{X} = \mathbf{ES} + \mathbf{W}, \quad (5)$$

where $\|\mathbf{C}\|_F \equiv \sqrt{\text{trace}\{\mathbf{CC}^T\}}$ denotes the Frobenius norm of \mathbf{C} . Note that Eq. (5) is actually a constrained LS problem, which can be converted to an unconstrained optimization problem as

$$\mathbf{S}^k = \arg \min_{\mathbf{S}} \|\mathbf{AES} + \mathbf{AW}^{k-1} - \mathbf{Y}\|_F^2 + \lambda_1 \|\mathbf{ES} + \mathbf{W}^{k-1} - \mathbf{X}^{k-1}\|_F^2 = \arg \min_{\mathbf{S}} \|\mathbf{C}^k \mathbf{S} - \mathbf{D}^k\|_F^2, \quad (6)$$

where λ_1 is a regularization parameter, and the superscript k ($k = 1, 2, 3, \dots$) represents the k 'th alternating iteration. Matrix \mathbf{C}^k and \mathbf{D}^k can be expressed as

$$\mathbf{C}^k = \begin{bmatrix} \mathbf{AE} \\ \sqrt{\lambda_1} \mathbf{E} \end{bmatrix}, \quad \mathbf{D}^k = \begin{bmatrix} \mathbf{AW}^{k-1} - \mathbf{Y} \\ \sqrt{\lambda_1} (\mathbf{W}^{k-1} - \mathbf{X}^{k-1}) \end{bmatrix}. \quad (7)$$

Now, \mathbf{S}^k can be estimated by the LS method as

$$\mathbf{S}^k = [(\mathbf{C}^k)^T \mathbf{C}^k]^{-1} (\mathbf{C}^k)^T \mathbf{D}^k, \quad (8)$$

where \mathbf{C}^T is the transposed matrix of \mathbf{C} and \mathbf{C}^{-1} is the inverse matrix of \mathbf{C} .

Before the beginning of the joint optimization process, it is necessary to determine the initial values of both \mathbf{S} and \mathbf{W} . In general, the initial value of \mathbf{W} is set to zero; thus, the initial value of \mathbf{S} can be obtained as

$$\mathbf{S}^0 = [(\mathbf{AE})^T \mathbf{AE}]^{-1} (\mathbf{AE})^T \mathbf{Y}. \quad (9)$$

4.2.2 Estimation of model error

Based on the analysis of the model error in Sec. 4.1, we now estimate \mathbf{W} by solving the following optimization problem with TV regularization:

$$\min_{\mathbf{W}} \frac{1}{2} \|\mathbf{AES} + \mathbf{AW} - \mathbf{Y}\|_F^2 + \lambda_{TV} \text{TV}(\mathbf{W}) \quad \text{s.t. } \mathbf{X} = \mathbf{ES} + \mathbf{W}, \quad (10)$$

where $\lambda_{TV} > 0$ is the regularization parameter of TV regularization term. $\text{TV}(\mathbf{W})$ is defined as the spatial anisotropy TV norm of \mathbf{W} which can be expressed as

$$\text{TV}(\mathbf{W}) \triangleq \|\mathbf{FW}\|_{1,1} = \|\mathbf{F}_v \mathbf{W}^T\|_{1,1} + \|\mathbf{F}_h \mathbf{W}^T\|_{1,1}, \quad (11)$$

where $\|\mathbf{C}\|_{1,1} \equiv \sum_{i=1}^L \|\mathbf{c}_i\|_1$ (\mathbf{c}_i is the i 'th column vector of \mathbf{C}), \mathbf{F}_v and \mathbf{F}_h are the discrete gradient operators in horizontal and vertical directions, respectively. The constrained optimization problem given in Eq. (10) can be translated into an unconstrained optimization problem as

$$\mathbf{W}^k = \arg \min_{\mathbf{W}} \frac{1}{2} \|\mathbf{AES}^k + \mathbf{AW} - \mathbf{Y}\|_F^2 + \lambda_{TV} \|\mathbf{FW}\|_{1,1} + \frac{\lambda_2}{2} \|\mathbf{ES}^k + \mathbf{W} - \mathbf{X}^{k-1}\|_F^2, \quad (12)$$

where $\lambda_2 > 0$ is a regularization parameter. Let $\mathbf{U} = \mathbf{Y} - \mathbf{AES}^k$, $\mathbf{V} = \mathbf{X}^{k-1} - \mathbf{ES}^k$, $\mathbf{H}_1 = \mathbf{AW}$, $\mathbf{H}_2 = \mathbf{W}$, $\mathbf{H}_3 = \mathbf{W}$, $\mathbf{H}_4 = \mathbf{FH}_3$, then Eq. (12) can be rewritten as

$$\begin{aligned} \min_{\mathbf{H}_1, \mathbf{H}_2, \mathbf{H}_3, \mathbf{H}_4} \quad & \frac{1}{2} \|\mathbf{H}_1 - \mathbf{U}\|_F^2 + \frac{\lambda_2}{2} \|\mathbf{H}_2 - \mathbf{V}\|_F^2 + \lambda_{TV} \|\mathbf{H}_4\|_{1,1} \\ \text{s.t. } \quad & \mathbf{AW} = \mathbf{H}_1, \quad \mathbf{W} = \mathbf{H}_2, \quad \mathbf{W} = \mathbf{H}_3, \quad \mathbf{FH}_3 = \mathbf{H}_4. \end{aligned} \quad (13)$$

In this paper, alternative direction multiplier method (ADMM) algorithm is employed to solve this optimization problem.²² The augmented Lagrangian function for \mathbf{W} , \mathbf{H}_1 , \mathbf{H}_2 , and \mathbf{H}_3 is

$$\begin{aligned} \mathcal{L}(\mathbf{W}, \mathbf{H}_1, \mathbf{H}_2, \mathbf{H}_3, \mathbf{H}_4, \mathbf{Q}_1, \mathbf{Q}_2, \mathbf{Q}_3, \mathbf{Q}_4) = & \frac{1}{2} \|\mathbf{H}_1 - \mathbf{U}\|_F^2 + \frac{\lambda_2}{2} \|\mathbf{H}_2 - \mathbf{V}\|_F^2 + \lambda_{TV} \|\mathbf{H}_4\|_{1,1} \\ & + \frac{\mu}{2} \|\mathbf{AW} - \mathbf{H}_1 - \mathbf{Q}_1\|_F^2 + \frac{\mu}{2} \|\mathbf{W} - \mathbf{H}_2 - \mathbf{Q}_2\|_F^2 \\ & + \frac{\mu}{2} \|\mathbf{W} - \mathbf{H}_3 - \mathbf{Q}_3\|_F^2 + \frac{\mu}{2} \|\mathbf{FH}_3 - \mathbf{H}_4 - \mathbf{Q}_4\|_F^2, \end{aligned} \quad (14)$$

where $\mu > 0$ is a positive penalty constant and \mathbf{Q}_1 , \mathbf{Q}_2 , \mathbf{Q}_3 , and \mathbf{Q}_4 denote the Lagrange multipliers.

Note that each iteration of the ADMM only minimizes the augmented Lagrangian function with respect to one variable and fixes the others to the last iteration value. However, the dual variables (Lagrange multipliers) are also updated by the gradient descent method in each iterative loop. By minimizing the augmented Lagrangian function \mathcal{L} with respect to \mathbf{W} , we can obtain

$$\begin{aligned} \mathbf{W}^t = & \arg \min_{\mathbf{W}} \mathcal{L}(\mathbf{W}, \mathbf{H}_1, \mathbf{H}_2, \mathbf{H}_3, \mathbf{H}_4, \mathbf{Q}_1, \mathbf{Q}_2, \mathbf{Q}_3, \mathbf{Q}_4) \\ = & \arg \min_{\mathbf{W}} \frac{\mu}{2} \|\mathbf{AW} - \mathbf{H}_1^{t-1} - \mathbf{Q}_1^{t-1}\|_F^2 \\ & + \frac{\mu}{2} \|\mathbf{W} - \mathbf{H}_2^{t-1} - \mathbf{Q}_2^{t-1}\|_F^2 + \frac{\mu}{2} \|\mathbf{W} - \mathbf{H}_3^{t-1} - \mathbf{Q}_3^{t-1}\|_F^2 \\ = & (\mathbf{A}^T \mathbf{A} + 2\mathbf{I}_L)^{-1} \left[\mathbf{A}^T (\mathbf{H}_1^{t-1} + \mathbf{Q}_1^{t-1}) + \right. \\ & \left. (\mathbf{H}_2^{t-1} + \mathbf{Q}_2^{t-1}) + (\mathbf{H}_3^{t-1} + \mathbf{Q}_3^{t-1}) \right], \end{aligned} \quad (15)$$

where $\mathbf{I}_L \in \mathbb{R}^{L \times L}$ is an identity matrix. The superscript t represents the t 'th iteration for calculating \mathbf{W} . The optimization problem to calculate \mathbf{H}_1 at t 'th iteration can be written as

$$\begin{aligned}
 \mathbf{H}_1^t &= \arg \min_{\mathbf{H}_1} \mathcal{L}(\mathbf{W}, \mathbf{H}_1, \mathbf{H}_2, \mathbf{H}_3, \mathbf{H}_4, \mathbf{Q}_1, \mathbf{Q}_2, \mathbf{Q}_3, \mathbf{Q}_4) \\
 &= \arg \min_{\mathbf{H}_1} \frac{1}{2} \|\mathbf{H}_1 - \mathbf{U}\|_F^2 + \frac{\mu}{2} \|\mathbf{A}\mathbf{W}^t - \mathbf{H}_1 - \mathbf{Q}_1^{t-1}\|_F^2 \\
 &= \frac{1}{1 + \mu} [\mathbf{U} + \mu(\mathbf{A}\mathbf{W}^t - \mathbf{Q}_1^{t-1})].
 \end{aligned} \tag{16}$$

Similarly, \mathbf{H}_2 at the t 'th iteration can be obtained by solving the following optimization problem

$$\begin{aligned}
 \mathbf{H}_2^t &= \arg \min_{\mathbf{H}_2} \mathcal{L}(\mathbf{W}, \mathbf{H}_1, \mathbf{H}_2, \mathbf{H}_3, \mathbf{H}_4, \mathbf{Q}_1, \mathbf{Q}_2, \mathbf{Q}_3, \mathbf{Q}_4) \\
 &= \arg \min_{\mathbf{H}_2} \frac{\lambda_2}{2} \|\mathbf{H}_2 - \mathbf{V}\|_F^2 + \frac{\mu}{2} \|\mathbf{W}^t - \mathbf{H}_2 - \mathbf{Q}_2^{t-1}\|_F^2 \\
 &= \frac{1}{\lambda_2 + \mu} [\mathbf{V} + \mu(\mathbf{W}^t - \mathbf{Q}_2^{t-1})].
 \end{aligned} \tag{17}$$

\mathbf{H}_3 at the t 'th iteration can be obtained by solving the following optimization problem

$$\begin{aligned}
 \mathbf{H}_3^t &= \arg \min_{\mathbf{H}_3} \mathcal{L}(\mathbf{W}, \mathbf{H}_1, \mathbf{H}_2, \mathbf{H}_3, \mathbf{H}_4, \mathbf{Q}_1, \mathbf{Q}_2, \mathbf{Q}_3, \mathbf{Q}_4) \\
 &= \arg \min_{\mathbf{H}_3} \frac{\mu}{2} \|\mathbf{W}^t - \mathbf{H}_3 - \mathbf{Q}_3^{t-1}\|_F^2 + \frac{\mu}{2} \|\mathbf{F}\mathbf{H}_3 - \mathbf{H}_4^{t-1} - \mathbf{Q}_4^{t-1}\|_F^2 \\
 &= (\mathbf{F}^T \mathbf{F} + \mathbf{I})^{-1} [\mathbf{W}^t - \mathbf{Q}_3^{t-1} + \mathbf{F}^T(\mathbf{H}_4^{t-1} + \mathbf{Q}_4^{t-1})],
 \end{aligned} \tag{18}$$

and \mathbf{H}_4 at the t 'th iteration can be obtained by solving the following optimization problem

$$\begin{aligned}
 \mathbf{H}_4^t &= \arg \min_{\mathbf{H}_4} \mathcal{L}(\mathbf{W}, \mathbf{H}_1, \mathbf{H}_2, \mathbf{H}_3, \mathbf{H}_4, \mathbf{Q}_1, \mathbf{Q}_2, \mathbf{Q}_3, \mathbf{Q}_4) \\
 &= \arg \min_{\mathbf{H}_4} \lambda_{\text{TV}} \|\mathbf{H}_4\|_{1,1} + \frac{\mu}{2} \|\mathbf{F}\mathbf{H}_3^t - \mathbf{H}_4 - \mathbf{Q}_4^{t-1}\|_F^2 \\
 &= \text{soft}\left(\mathbf{F}\mathbf{H}_3^t - \mathbf{Q}_4^{t-1}, \frac{\lambda_{\text{TV}}}{\mu}\right),
 \end{aligned} \tag{19}$$

where $\text{soft}(\ast)$ denotes the component-wise application of the soft-threshold function. At the end of each iteration, for calculating \mathbf{W} , the Lagrange multipliers \mathbf{Q}_1 , \mathbf{Q}_2 , \mathbf{Q}_3 , and \mathbf{Q}_4 are updated by the gradient descent method as follows:

$$\begin{aligned}
 \mathbf{Q}_1^t &= \mathbf{Q}_1^{t-1} - \mathbf{A}\mathbf{W}^t + \mathbf{H}_1^t, \\
 \mathbf{Q}_2^t &= \mathbf{Q}_2^{t-1} - \mathbf{W}^t + \mathbf{H}_2^t, \\
 \mathbf{Q}_3^t &= \mathbf{Q}_3^{t-1} - \mathbf{W}^t + \mathbf{H}_3^t, \\
 \mathbf{Q}_4^t &= \mathbf{Q}_4^{t-1} - \mathbf{F}\mathbf{H}_3^t + \mathbf{H}_4^t.
 \end{aligned} \tag{20}$$

Note that all functions that form the objective function \mathcal{L} in Eq. (14) are closed, proper, and convex. If Eq. (13) exists as an optimal solution, the following constraints will be met.

$$\varepsilon^t = \|\mathbf{A}\mathbf{W} - \mathbf{H}_1\|_F + \|\mathbf{W} - \mathbf{H}_2\|_F + \|\mathbf{W} - \mathbf{H}_3\|_F + \|\mathbf{F}\mathbf{H}_3 - \mathbf{H}_4\|_F \rightarrow 0 \quad \text{as } t \rightarrow \infty. \tag{21}$$

If Eq. (13) does not have a solution, then ε will diverge. As for the stopping criterion in Algorithm 1, we use $\varepsilon^t \leq \sqrt{N}\varepsilon_{\text{ref}}$, where the reference value $\varepsilon_{\text{ref}} > 0$. In the later experiments, we empirically set ε_{ref} equal to 10^{-5} . In summary, the estimation of \mathbf{W} by ADMM is described in Algorithm 1.

Algorithm 1 ADMM for estimating W^k .**Inputs:** A, E, Y, S_k **Set parameters:** λ_2, λ_{TV} , and μ **Initialize:** $W^0, H_1^0, \dots, H_4^0, Q_1^0, \dots, Q_4^0$, and set $t = 1$ **Repeat:**

1. compute W^t by minimizing \mathcal{L} with respect to W by Eq. (15);
2. compute H_1^t by minimizing \mathcal{L} with respect to H_1 by Eq. (16);
3. compute H_2^t by minimizing \mathcal{L} with respect to H_2 by Eq. (17);
4. compute H_3^t by minimizing \mathcal{L} with respect to H_3 by Eq. (18);
5. compute H_4^t using soft-threshold operation by Eq. (19);
6. update Lagrange multipliers Q_1, Q_2, Q_3 , and Q_4 by Eq. (20);
7. update iteration: $t = t + 1$;

until stopping criterion is satisfied;**Output:** model error matrix $W^k = W^t$.**4.3 Reconstruction of HSI**

The reconstruction of the target HSI in the k 'th alternate iteration can be expressed as

$$X^k = ES^k + W^k. \quad (22)$$

The termination parameter is defined as the relative change of the adjacent alternate iteration results

$$\zeta = \frac{\|X^k - X^{k-1}\|_F}{\|X^k\|_F}. \quad (23)$$

Note that the alternate iteration process will be stopped once ζ is smaller than a preset value. In summary, the proposed reconstruction is described in Algorithm 2.

Algorithm 2 The proposed SCR-SUEC algorithm.**Inputs:** A, E, Y **Set parameters:** $\lambda_1, \lambda_2, \lambda_{TV}$, and μ **Initialize:** W^0, S^0 , and set $k = 1$ **Repeat:**

1. compute S^k using constrained LS method by Eq. (8);
2. compute W^k by Algorithm 1;
3. compute X^k by Eq. (22);
4. update iteration $k = k + 1$;

until stopping criterion is satisfied;**Output:** The reconstructed HSI $\hat{X} = X^k$.

5 Experimental Results and Discussion

Experimental results are reported in terms of sampling rates, signal-to-noise ratio (SNR) structural similarity (SSIM) and computational complexity, where SNR can effectively evaluate the reconstruction quality of each algorithm under certain sampling rate. The band-based SNR measured in decibels is defined as

$$\text{SNR}(X_l, \hat{X}_l) = 10 \log_{10} \frac{\|X_l\|_2^2}{\|X_l - \hat{X}_l\|_2^2}. \quad (24)$$

Note that there have been several popular algorithms that were used for hyperspectral CS, such as BCS-SPL,⁹ 3-D CS,¹⁰ and CPPCA,¹² and several improved algorithms based on CPPCA. To illustrate the efficiency of the proposed algorithm, the reconstruction quality of the proposed SCR-SUEC was compared with that of the above algorithms and the SCR-SUEC without error compensation (SCR-SU). In this paper, we use the raw data of HSIs as the tested data. As we know, raw data are the original data that are acquired by airborne or spaceborne platforms without any preprocessing, which are helpful in evaluating the performance of the algorithm objectively. The raw HSIs acquired by airborne visible infrared imaging spectrometer (AVIRIS) are employed for performance evaluation. The six tested data are referred to *aviris_sc0.raw*, *aviris_sc3.raw*, *aviris_sc11.raw*, *aviris_sc18.raw*, *hawaii_sc01.raw*, and *maine_sc10.raw*.²³ Owing to the limitation of 3-D CS algorithms, the spatial size of all tested images were intercepted to 180×180 , with 224 spectral bands and 2 bytes for each pixel, where the 90th band of each image is shown in Fig. 3.

5.1 Parameters Selection for the Proposed SCR-SUEC

Since the performance of the proposed SCR-SUEC is sensitive to the setting of parameters λ_1 , λ_2 , λ_{TV} , and μ , we first investigate the reasonable range for the above parameters setting with a

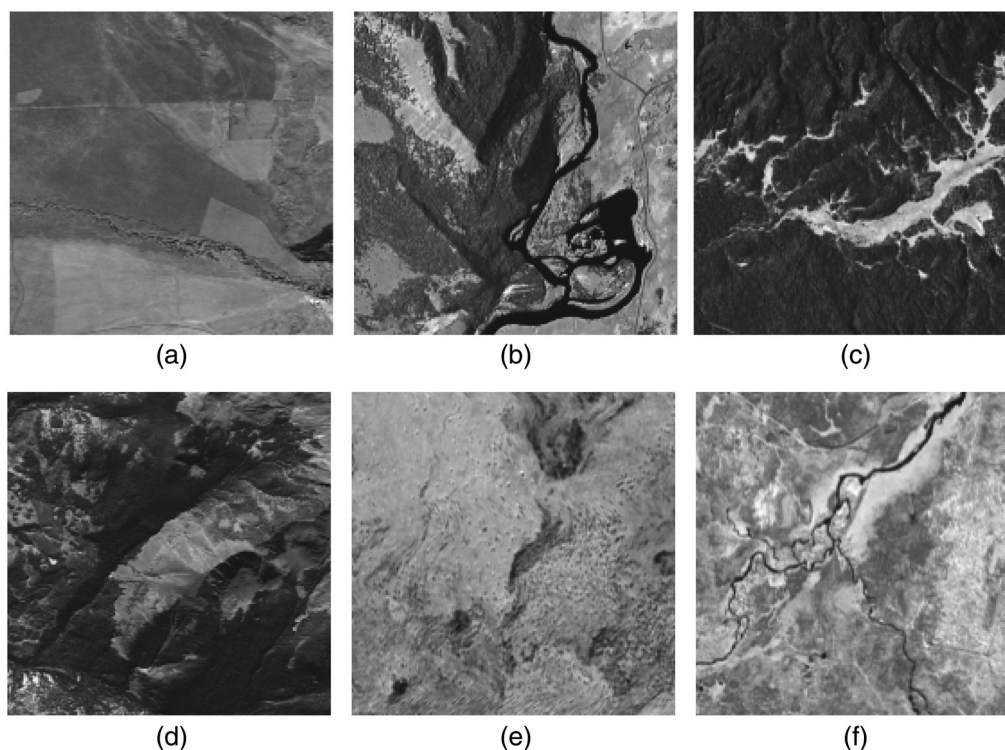


Fig. 3 The 90th band of original images of six HSIs: (a) *aviris_sc0.raw*, (b) *aviris_sc3.raw*, (c) *aviris_sc11.raw*, (d) *aviris_sc18.raw*, (e) *hawaii_sc01.raw*, and (f) *maine_sc10.raw*.

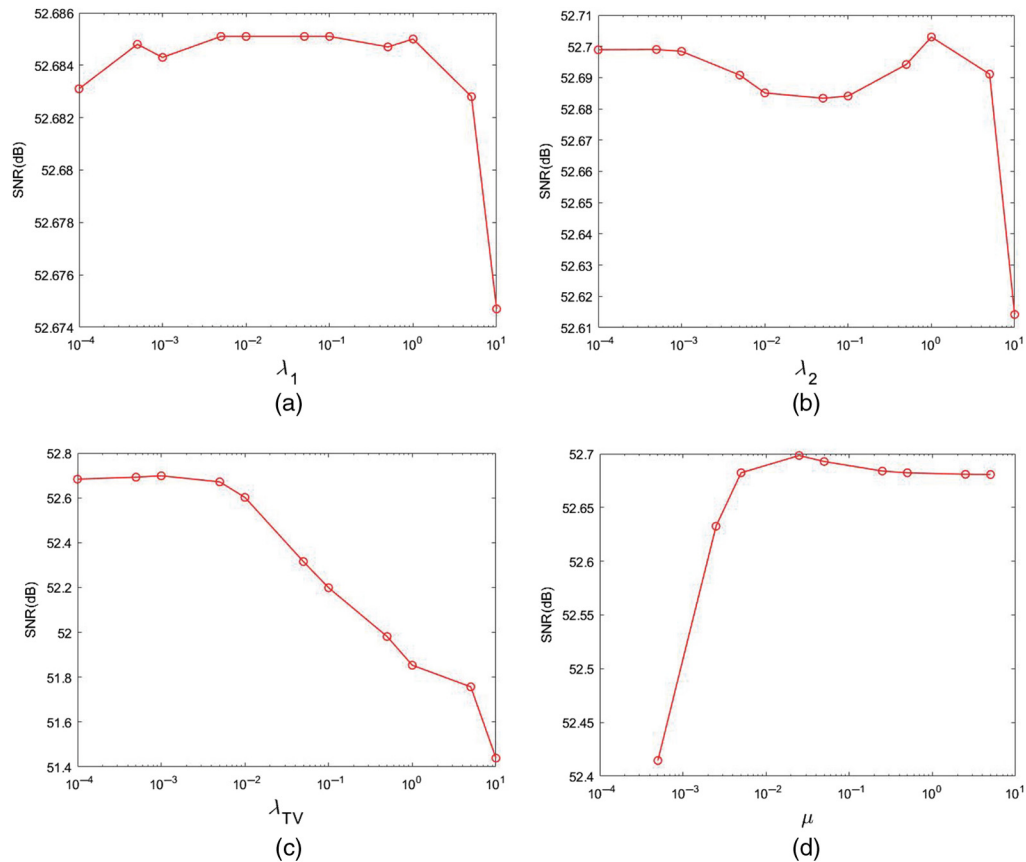


Fig. 4 Value analysis of three regularization parameters λ_1 , λ_2 , and λ_{TV} as well as the penalty constant μ for the proposed SCR-SUEC algorithm: (a) λ_1 , (b) λ_2 , (c) λ_{TV} , and (d) μ

sampling rate of 0.5. Figure 4 shows the reconstructed quality by using different values of the above four parameters for the proposed SCR-SUEC. We can empirically observe that the parameter λ_{TV} plays a dominant role. Note that the three parameters (λ_1 , λ_2 , and λ_{TV}) can obtain good reconstruction results within a certain range of values. In general, the recommended ranges of λ_1 , λ_2 , and λ_{TV} are [0.001, 1], [0.001, 1] and [0.0001, 0.01], respectively. Note that the parameter μ controls the convergence speed of the ADMM algorithm. Adaptive adjustment of parameter μ will perform well in most cases. Similar with the above parameters, the value of parameter μ in a large range can guarantee high reconstruction quality, where the recommended range of μ is [0.005, 1]. As for the proposed SCR-SUEC, according to the above recommended range, λ_1 and λ_2 are both set to 0.1, λ_{TV} is set to 0.003, and μ is set to 0.05.

5.2 Comparison of Reconstruction Performance

In this study, we assume that the endmembers are known, which can be selected from the spectrum library. Although there have been several off-the-shelf spectrum libraries that provide a large amount of standard spectrum of different objects, those standard ones are obviously different with the spectrum in the actual HSI due to the weather influence, imaging conditions, and so on.⁴ Therefore, we employed the popular vertex component analysis²⁴ algorithm on each considered HSI to extract a certain amount of endmembers to establish a spectrum library for spectral unmixing, which is an international common practice for such research.

The performance of the proposed SCR-SUEC was first compared with BCS-SPL, 3-D CS, CPPCA, and SCR-SU, with the sampling rates ranging from 0.1 to 0.5 with a step size of 0.1. The relationship between SNR and sampling rates is reported in Fig. 5. Note that the block size employed by BCS-SPL is set to 50×50 . As we know, BCS-SPL has only spatial sampling process, while the others have only spectral sampling process. For the sake of fairness, the total

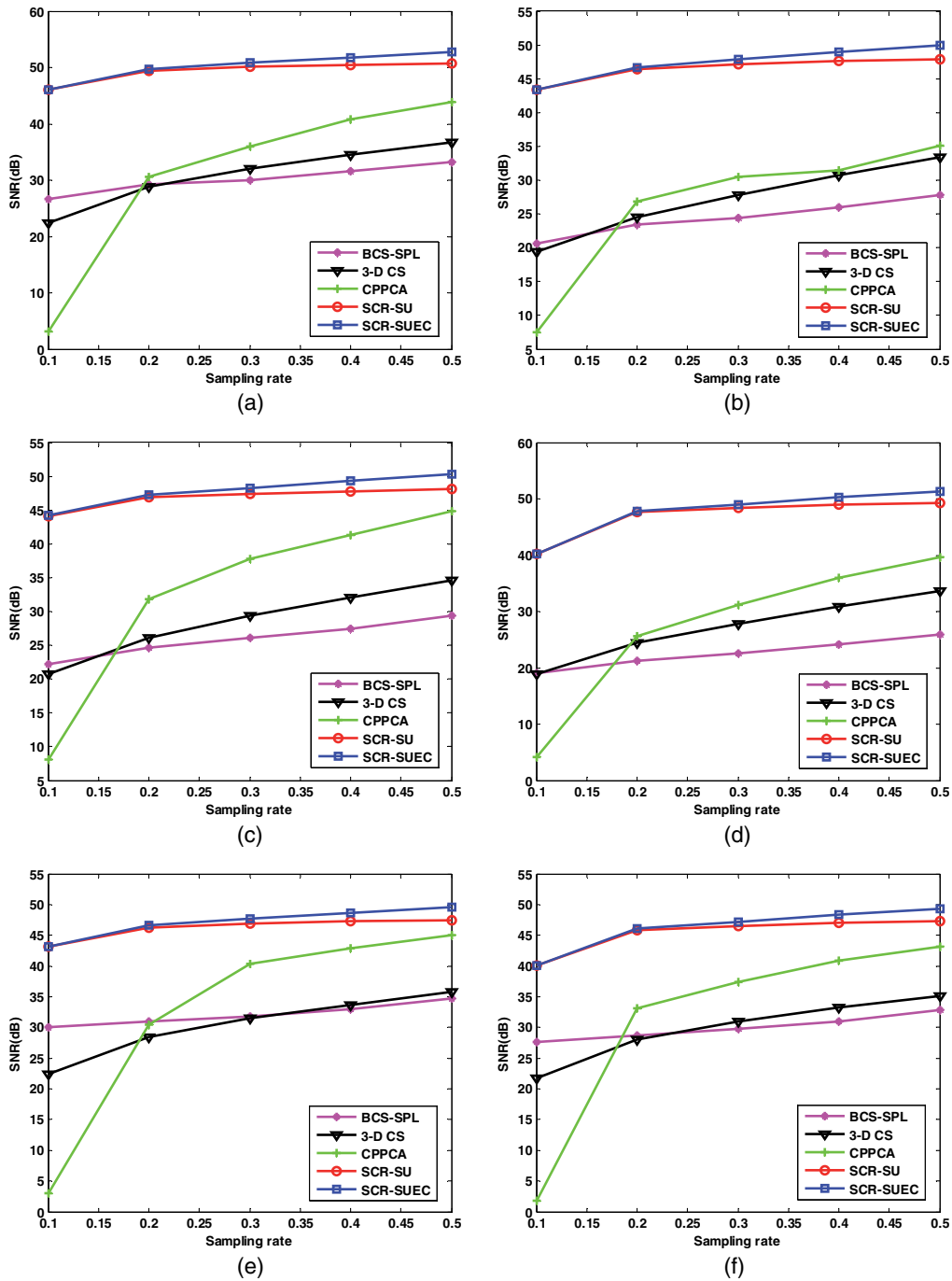


Fig. 5 Comparison of reconstruction performance using SNR: (a) aviris_sc0.raw, (b) aviris_sc3.raw, (c) aviris_sc11.raw, (d) aviris_sc18.raw, (e) hawaii_sc01.raw, and (f) maine_sc10.raw.

sampling rates in all of the above algorithms are identical. It is easy to see that, with the increase of sampling rate, the reconstruction performance of all algorithms improves gradually. However, the growth speed of both SCR-SU and SCR-SUEC is relatively slow compared with the other algorithms. The performance of the above two algorithms has relatively low sensitivity on sampling rates, which indicates that the proposed algorithm can also achieve a perfect performance at a low sampling rate. Although BCS-SPL employs the block-based CS scheme to perform hyperspectral CS, its average performance is the worst, since BCS-SPL handles each band separately without considering the spectral correlation. Another problem for BCS-SPL is the difficulty to select the optimal block size. Note that a large block size provides better performance than a

small block size. However, a large block size also leads to a high computational complexity. By utilizing both the spatial and spectral sparsity, 3-D CS can improve the reconstruction performance, compared with BCS-SPL at high sampling rates. However, the performance of 3-D CS is worse than that of BCS-SPL at low sampling rates. In general, CPPCA may perform better than 3-D CS with low computational efficiency. In our study, the performance achieved by CPPCA is better than that achieved by 3-D CS for all tested HSIs except that at a low sampling rate of 0.1 because CPPCA cannot work normally at very low sampling rates, which is a significant drawback for CPPCA.

It is obvious that SCR-SU performs much better than BCS-SPL, CPPCA, and 3-D CS at all sampling rates, which demonstrates the superiority of reconstruction quality by using spectral unmixing. Compared with 3-D CS that uses convex optimization to reconstruct the target HSI, the proposed SCR-SUEC uses convex optimization to reconstruct the error image instead of the target HSI. In general, SCR-SU can be regarded as a special case of SCR-SUEC without error compensation. Since there is model error between ES and X , despite SCR-SU yielding satisfied reconstruction, by exploiting model error, SCR-SUEC can further improve the reconstruction performance. As can be seen, with the lowest sampling rate (i.e., 0.1), since it is really difficult to explore the error image due to the lack of necessary information, SCR-SUEC may degenerate into SCR-SU. Even so, it still provides much better reconstructed quality than other algorithms. With the increase of sampling rate, SCR-SUEC achieves a prominent performance gain compared with SCR-SU since the model error is well explored. Note that when the sampling rate reaches 0.5, the performance gap between SCR-SUEC and SCR-SU can be more than 2 dB. In fact, the onboard hyperspectral data usually have raw format without any correction and other processing; therefore, the experimental results on raw dataset can be more convincing. On the other hand, in terms of onboard compression, low compression ratio is usually employed in order to preserve the information as much as possible. Because low sampling rates can easily destroy significant amounts of information in HSI, higher sampling rates are preferred. In this case, SCR-SUEC can significantly outperform all the other algorithms.

In addition to SNR, SSIM is also used to evaluate the reconstructed quality. The comparison of reconstructed quality using SSIM is reported in Fig. 6. As can be seen, 3-D CS outperforms BCS-SPL significantly. When the sampling rate is lower, such as 0.1, since CPPCA fails at low sampling rates, the SSIM achieved by CPPCA is too small. As for the other sampling rates, the performance of CPPCA is better than the two algorithms mentioned above. Note that SCR-SU performs better than the other algorithms. Moreover, SCR-SUEC can significantly improve SSIM, compared with SCR-SU. Therefore, the results in Fig. 6 further demonstrate the effectiveness of the proposed SCR-SUEC in CS reconstruction for HSIs.

Compared with objective quality evaluation by using SNR, subjective quality evaluation should also be considered. The reconstructed images of the 90th band of *aviris_sc18.raw* achieved by various algorithms with a sampling rate of 0.5 are shown in Fig. 7. Because each pixel of the tested images has 2 bytes, the reconstructed images are displayed by using their normalized values. As can be seen from this figure, the visual quality of the reconstructed images by BCS-SPL is not satisfying, which is too rough to distinguish the detailed features for the following processing. Compared with BCS-SPL, 3-D CS only slightly improves the performance, despite considering the spectral sparsity, while CPPCA significantly improves the performance, which effectively retains the detailed features with a higher SNR. Note that SCR-SU outperforms the above three algorithms not only in visual effect but also in SNR, where the detailed textures are more close to the original image. By exploring the model error, SCR-SUEC can further improve the reconstruction quality compared with SCR-SU, where the maximum performance gain can be more than 5 dB for some bands. Although this improvement is not evident in the visual effect, it may be very important for some applications, such as classification, segmentation, and target detection.

The comparison of reconstructed images achieved by various algorithms at a low sampling rate of 0.2 is also given in Fig. 8. As can be seen, the quality of the reconstructed images obtained by both BCS-SPL and 3-D CS is poor, and the former has obvious block effect. CPPCA provides general reconstruction quality, which is better than that provided by both BCS-SPL and 3-D CS. Note that the reconstructed images by SCR-SU and SCR-SUEC are very comfortable for eyes, where the detailed features are more distinct than the other reconstructed ones. This indicates that

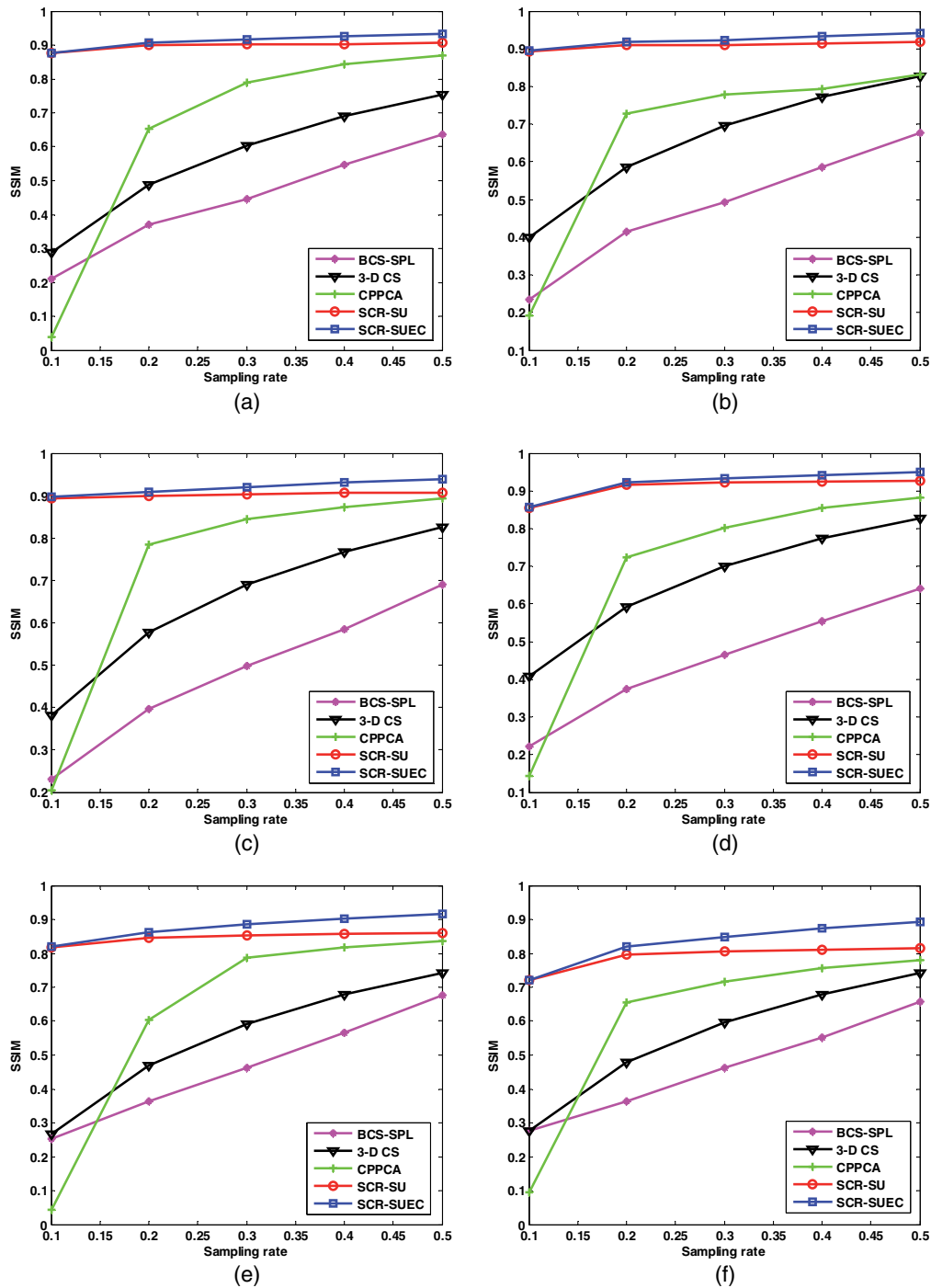


Fig. 6 Comparison of reconstruction performance using SSIM: (a) aviris_sc0.raw, (b) aviris_sc3.raw, (c) aviris_sc11.raw, (d) aviris_sc18.raw, (e) hawaii_sc01.raw, and (f) maine_sc10.raw.

at a lower sampling rate, SCR-SU and SCR-SUEC provide great advantages over the other algorithms. Note that the proposed SCR-SUEC can still achieve a performance gain compared with SCR-SU even at a low sampling rate. To further compare the performance of above algorithms, Fig. 9 shows the comparison of reconstructed images at a low sampling rate of 0.1. It is clear that the quality of reconstructed images obtained by BCS-SPL, 3-D CS, and CPPCA is too bad, where there is serious block effect in the reconstructed images obtained by BCS-SPL and regular horizontal noise strips in the reconstructed images obtained by CPPCA. Although the error exploration can hardly work at very low sampling rate, we can still obtain the reconstruction

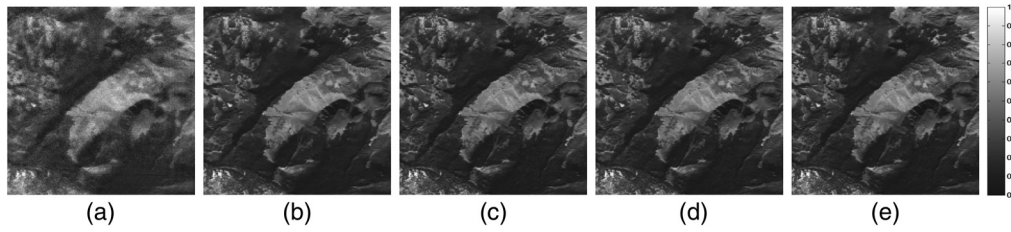


Fig. 7 Comparison of reconstructed images of the 90th band of aviris_sc18.raw with a sampling rate of 0.5: (a) BCS-SPL, SNR = 19.86 dB, (b) 3-D CS, SNR = 31.60 dB, (c) CPPCA, SNR = 42.03 dB, (d) SCR-SU, SNR = 49.97 dB, and (e) SCR-SUEC, SNR = 55.07 dB.

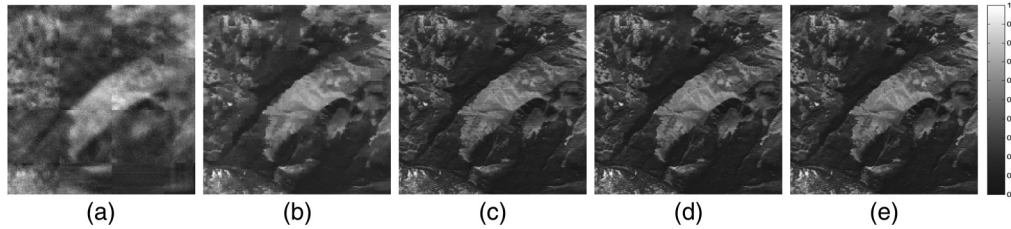


Fig. 8 Comparison of reconstructed images of the 90th band of aviris_sc18.raw with a sampling rate of 0.2: (a) BCS-SPL, SNR = 16.22 dB, (b) 3-D CS, SNR = 21.74 dB, (c) CPPCA, SNR = 30.44 dB, (d) SCR-SU, SNR = 48.16 dB, and (e) SCR-SUEC, SNR = 48.78 dB.

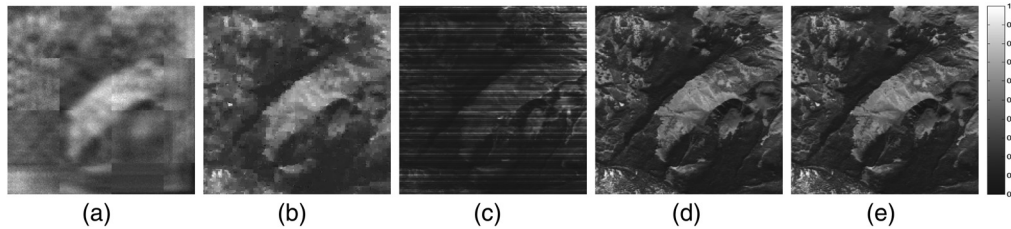


Fig. 9 Comparison of reconstructed images of the 90th band of aviris_sc18.raw with a sampling rate of 0.1: (a) BCS-SPL, SNR = 14.42 dB, (b) 3-D CS, SNR = 16.59 dB, (c) CPPCA, SNR = 2.97 dB, (d) SCR-SU, SNR = 39.31 dB, and (e) SCR-SUEC, SNR = 39.33 dB.

with much higher quality, which shows the great advantage over BCS-SPL, 3-D CS, and CPPCA.

To further illustrate the performance from the perspective of spectral feature preservation, the average spectral angle distance (SAD) is introduced to measure the quality of the reconstructed images, which can be calculated as

$$SAD(\mathbf{r}, \mathbf{z}) = \arccos \left[\frac{\sum_{k=1}^L r_k z_k}{\sqrt{\sum_{k=1}^L r_k^2} \sqrt{\sum_{k=1}^L z_k^2}} \right], \quad (25)$$

where \mathbf{r} and \mathbf{z} are the two spectral curves with L pixels. Note that SAD can effectively evaluate the distortion of spectral curves, where the larger the SAD value, the greater the distortion. Table 1 gives the comparison of average SAD achieved by various algorithms at a sampling rate of 0.5. As can be seen, SCR-SU provides lower SAD for all tested images than BCS-SPL, 3-D CS, and CPPCA, and the proposed SCR-SUEC further reduces SAD values based on SCR-SU, demonstrating its excellent preservation ability of spectral characteristic.

The comparison of spectral curves of aviris_sc0 is given in Fig. 10, where the spatial location is (50, 50) and the sampling rate is 0.2. It is obvious that the spectral curve obtained by the proposed SCR-SUEC has the highest similarity with the original one even at the low sampling

Table 1 Comparison of SAD achieved by various algorithms (sampling rate = 0.5). The smallest SAD for each data is specified in boldface.

	BCS-SPL	3-D CS	CPPCA	SCR-SU	SCR-SUEC
aviris_sc0.raw	0.0268	0.0242	0.0053	0.0026	0.0020
aviris_sc3.raw	0.0606	0.0372	0.0157	0.0038	0.0028
aviris_sc11.raw	0.0491	0.0350	0.0047	0.0036	0.0027
aviris_sc18.raw	0.0657	0.0388	0.0087	0.0031	0.0024
hawaii_sc01.raw	0.0230	0.0490	0.0049	0.0038	0.0029
maine_sc10.raw	0.0369	0.0697	0.0042	0.0029	0.0022

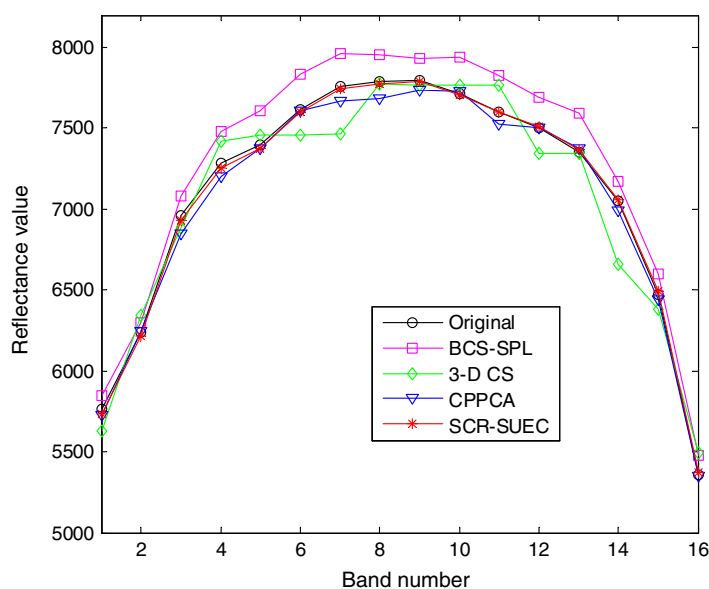


Fig. 10 Comparison of spectral curves at the same spatial location obtained by various algorithms.

Table 2 Comparison of reconstruction performance achieved by various algorithms (unit: decibels). The maximum SNR at each sampling rates is specified in boldface.

Sampling rates	0.2	0.3	0.4	0.5
CPPCA	13.95	16.80	20.11	22.65
C-CPPCA	16.55	19.65	21.55	23.62
MH(NU)-CPPCA	17.59	22.03	25.11	27.55
MH(U)-CPPCA	17.26	21.52	24.53	26.92
MH(NU)-C-CPPCA	19.16	23.13	25.44	27.95
MH(U)-C-CPPCA	18.99	22.48	24.94	27.34
SCR-SU	31.17	32.29	32.76	33.16
SCR-SUEC	31.47	33.06	34.06	35.14

Table 3 Comparison of complexity achieved by different algorithms (unit: seconds).

	Running time/band
BCS-SPL	5.612
3-D CS	3.179
CPPCA	0.005
SCR-SU	0.001
SCR-SUEC	3.317

rate, which is very important for the postprocessing applications of reconstructed images to achieve better results.

The performance of the proposed algorithm was also compared with some improved CPPCA algorithms, such as class-dependent compressive projection principal component analysis (C-CPPCA),²⁵ MH(U)-CPPCA, MH(NU)-CPPCA, MH(U)-C-CPPCA, and MH(NU)-C-CPPCA,¹⁵ where MH denotes the multiple hypotheses, U denotes the uniform grouping, and NU denotes nonuniform grouping. Table 2 gives the comparison of reconstruction performance achieved by various algorithms tested in University of Pavia. As can be seen, SCR-SU performs much better than the other state-of-art algorithms; moreover, the proposed SCR-SUEC can further improve the reconstruction performance, compared with SCR-SU, which demonstrates the superiority of error estimation and compensation.

5.3 Comparison of Complexity

As for the computational complexity of the SCR-SUEC algorithm, the most time-consuming step is to calculate W and H_3 in each iteration, where the orders of complexity are $O(NJ^2)$ and $O(JN \log N)$, respectively. The overall order of complexity in per iteration is given as $O(NJ^2 + JN \log N)$. Furthermore, we use the average running time per band to evaluate the complexity of each algorithm, which is reported in Table 3. Note that the algorithms were implemented using Matlab2014 on a desktop PC equipped with an Intel Core CPU (3.30 GHz) and 8 GB of RAM memory. As can be seen, BCS-SPL has the highest complexity; meanwhile, its reconstructed performance is the worst. The complexity of 3-D CS is much lower than that of BCS-SPL. CPPCA also has relatively low complexity. SCR-SU provides high reconstructed performance with the lowest complexity. Owing to the iterative operation, the complexity of the proposed SCR-SUEC is much higher than that of SCR-SU. However, SCR-SUEC can outperform all the other algorithms in terms of reconstructed performance. Because the reconstruction process is carried out on the ground, where the computing resources are often sufficient, therefore, a certain degree of increase in complexity will not have a significant impact on practical application.

6 Conclusion

In this paper, a reconstruction algorithm from spectral compressive sampling by exploring model error of LMM is proposed. Supposing the endmembers are known, compressive sampling only performed on the spectral direction of the original HSI to retain the information of abundance. At the reconstruction stage, considering the limitation of LMM, the error introduced by LMM is further explored by solving an optimization problem with the TV regularization. To further improve the reconstruction accuracy, a joint optimization scheme for estimating abundance and model error is proposed and an alternate iterative updating manner for their calculation is carried out. Once we obtain the optimal results of abundance and model error, the final reconstruction of the target HSI can be achieved according to LMM. Experimental results on several real hyperspectral datasets demonstrate that the proposed SCR-SUEC outperforms the other classical algorithms in terms of both SNR and visual effect. In future research, more powerful sparse

representation should be explored for the estimation of model error, which may further improve reconstruction performance, especially under low sampling rates.

Acknowledgments

This work was supported by the Overseas Visiting and Research Project for Excellent Young Key Talents in Higher Education Institutions in Anhui Province (No. gxgwfx2019056), the Quality Engineering Project of Universities of Anhui Province (No. 2016zy126), and Chongqing Research Program of Basic Research and Frontier Technology (No. cstc2016jcyjA0539). The authors would like to thank the anonymous reviewers for their outstanding comments and suggestions, which greatly improved the technical quality of this paper.

References

1. D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory* **52**(4), 1289–1306 (2006).
2. L. Zhang et al., "Exploring structured sparsity by a reweighted Laplace prior for hyperspectral compressive sensing," *IEEE Trans. Image Process.* **25**(10), 4974–4988 (2016).
3. Y. Wang et al., "Compressive sensing of hyperspectral images via joint tensor Tucker decomposition and weighted total variation regularization," *IEEE Geosci. Remote Sens. Lett.* **14**(12), 2457–2461 (2017).
4. L. Zhang et al., "Locally similar sparsity-based hyperspectral compressive sensing using unmixing," *IEEE Trans. Comput. Imaging* **2**(2), 86–100 (2016).
5. G. Martín and J. M. Bioucas-Dias, "Hyperspectral blind reconstruction from random spectral projections," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **9**(6), 2390–2399 (2016).
6. L. Wang et al., "Compressed sensing reconstruction of hyperspectral images based on spectral unmixing," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **11**(4), 1266–1284 (2018).
7. J. Hahn et al., "Compressive sensing and adaptive direct sampling in hyperspectral imaging," *Digital Signal Process.* **26**, 113–126 (2014).
8. I. Noor and E. L. Jacobs, "Adaptive compressive sensing algorithm for video acquisition using a single-pixel camera," *J. Electron. Imaging* **22**(2), 021013 (2013).
9. S. Mun and J. E. Fowler, "Block compressed sensing of images using directional transforms," in *Proc. IEEE Int. Conf. Image Process.*, Cairo, Egypt, pp. 3021–3024 (2009).
10. X. B. Shu and N. Ahuja, "Imaging via three-dimensional compressive sampling (3DCS)," in *Proc. IEEE Int. Conf. Comput. Vision*, Barcelona, Spain, pp. 439–446 (2011).
11. Y. B. Jia, Y. Feng, and Z. L. Wang, "Reconstructing hyperspectral images from compressive sensors via exploiting multiple priors," *Spectrosc. Lett.* **48**(1), 22–26 (2014).
12. J. E. Fowler, "Compressive-projection principal component analysis," *IEEE Trans. Image Process.* **18**(10), 2230–2242 (2009).
13. W. Li, S. Prasad, and J. E. Fowler, "Integration of spectral-spatial information for hyperspectral image reconstruction from compressive random projections," *IEEE Geosci. Remote Sens. Lett.* **10**(6), 1379–1383 (2013).
14. N. H. Ly, Q. Du, and J. E. Fowler, "Reconstruction from random projections of hyperspectral imagery with spectral and spatial partitioning," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **6**(2), 466–472 (2013).
15. C. Chen et al., "Reconstruction of hyperspectral imagery from random projections using multihypothesis prediction," *IEEE Trans. Geosci. Remote Sens.* **52**(1), 365–374 (2014).
16. G. Martín, J. M. Bioucas-Dias, and A. Plaza, "HYCA: a new technique for hyperspectral compressive sensing," *IEEE Trans. Geosci. Remote Sens.* **53**(5), 2819–2831 (2015).
17. J. M. Bioucas-Dias et al., "Hyperspectral unmixing overview: geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **5**(2), 354–379 (2012).
18. A. Ramirez, G. R. Arce, and B. M. Sadler, "Spectral image unmixing from optimal coded-aperture compressive measurements," *IEEE Trans. Geosci. Remote Sens.* **53**(1), 405–415 (2015).

19. M. Golbabaee, S. Arberet, and P. Vandergheynst, "Compressive source separation: theory and methods for hyperspectral imaging," *IEEE Trans. Image Process.* **22**(12), 5096–5110 (2013).
20. Z. L. Wang, Y. Feng, and Y. B. Jia, "Spatio-spectral hybrid compressive sensing of hyperspectral imagery," *Remote Sens. Lett.* **6**(3), 199–208 (2015).
21. J. E. Fowler, "Compressive pushbroom and whiskbroom sensing for hyperspectral remote-sensing imaging," in *Proc. IEEE Int. Conf. Image Process.*, Paris, France, pp. 684–688 (2014).
22. S. Boyd et al., "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.* **3**(1), 1–122 (2010).
23. Jet Propulsion Lab., "AVIRIS free data," California Institute of Technology, Pasadena, <http://aviris.jpl.nasa.gov/html/aviris.freedata.html> (2008).
24. J. M. P. Nascimento and J. M. B. Dias, "Vertex component analysis: a fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.* **43**(4), 898–910 (2005).
25. W. Li, S. Prasad, and J. E. Fowler, "Classification and reconstruction from random projections for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.* **51**(2), 833–843 (2013).

Zhongliang Wang received his PhD in signal and information processing in 2015 from the School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China. Currently, he is an associate professor at the Department of Electric Engineering, Tongling University, Tongling, China. His research interests include image enhancement and hyperspectral remote sensing.

Mi He received her PhD in information and communication engineering in 2012 from the National University of Defense and Technology, Changsha, China. Since 2015, she has been an associate professor at the College of Biomedical Engineering and Imaging Medicine, Army Medical University. Her research involves remote sensing and biomedical signal processing.

Zhen Ye received her PhD in information and communication engineering from the Northwestern Polytechnical University, Xi'an, China, in 2015. She spent one year as an exchange student at Mississippi State University. Currently, she is an associate professor at the School of Electronics and Control Engineering, Chang'an University. Her research interests include hyperspectral image (HSI) analysis, pattern recognition, and machine learning.

Yongjian Nian received his PhD in information and communication engineering from the National University of Defense and Technology, Changsha, China. He is an associate professor at the College of Biomedical Engineering and Imaging Medicine, Army Medical University. His research fields include HSI processing, pattern recognition, and machine learning.

Liang Qiao received his PhD in biomedical engineering in 2017 from Army Medical University, Chongqing, China. He is an associate professor at College of Biomedical Engineering and Imaging Medicine, Army Medical University. His research interests include pattern recognition and machine learning.

Mingsheng Chen received his PhD from the College of Electronic Science and Engineering, National University of Defense Technology, Changsha, in 2012. From 2014 to 2018, he studied as postdoctoral fellow at the College of Biomedical Engineering and Imaging Medicine at Army Medical University, Chongqing. His research interest fields include pattern recognition and machine learning.