Reconstruction of Hyperspectral Images From Spectral Compressed Sensing Based on a Multitype Mixing Model

Zhongliang Wang, Mi He, Zhen Ye, Ke Xu, Yongjian Nian, and Bormin Huang

***Abstract*—Hyperspectral compressed sensing (HCS) based on spectral unmixing technique has shown great reconstruction performance. In particular, the linear mixed model (LMM) has been widely used in HCS reconstruction. However, due to the complexity of environmental conditions, instrumental configurations, and material nonlinear mixing effects, LMM cannot accurately represent the hyperspectral images, which limits the improvement of reconstruction quality. In this article, first, by introducing spectral variability, nonlinear mixing, and residuals, a multitype mixed model (MMM) is proposed to establish a more accurate hyperspectral image model. Then, a novel MMM-based HCS is proposed, which performs spectral compressed sampling at the sampling stage only, and at the reconstruction stage, by using spectral unmixing, an MMM-based HCS super-resolution reconstruction algorithm from spectral compressed sensing data is developed, and the alternating direction multiplier method is employed to estimate each component of the MMM, furthermore, reasonable prior knowledge of each component is introduced to improve the estimation accuracy. Experimental results on hyperspectral datasets demonstrate that the proposed model outperforms those state-of-the-art methods based on the LMM in terms of HCS reconstruction quality.**

***Index Terms*—Compressed sensing, hyperspectral remote sensing, linear mixing model (LMM), spectral unmixing.**

I. Introduction

HYPERSPECTRAL images (HSIs) can provide detailed ground features and are widely used in mineral exploration, agricultural production, environmental monitoring, and military reconnaissance [1]. However, with the continuous development of imaging spectrometers, the spatial resolution and spectral resolution of HSIs are significantly improved, which leads to a large amount of hyperspectral data acquired by imaging spectrometers. Therefore, big data bring tremendous pressure to the power consumption, computing power, and real-time transmission of airborne or satellite-borne imaging systems, making it necessary to effectively compress HSIs by using an advanced compression technique. Traditional compression technology first collects high-resolution hyperspectral data through an optical system and then throws away some information to achieve the effect of data compression. This wasteful model significantly increases the power consumption and computational complexity of the imaging platform. The compressed sensing (CS) technique can collect data at a low sampling rate (much lower than the Nyquist sampling rate), and data acquisition and compression can be completed in the same process [2], which is suitable for data acquisition and real-time transmission with resource-constrained airborne or satellite-borne hyperspectral imaging platforms. Based on the acquired low-resolution sampling data, the CS reconstruction algorithm can achieve the ideal reconstruction of the original data under the premise of sparsity. Therefore, the CS technique also provides an effective way for the super-resolution reconstruction of HSIs.

At present, hyperspectral compressed sensing (HCS) has attracted increasing attention in the field of hyperspectral remote sensing. One of the main concerns is how to reconstruct the original HSIs with high quality from the compressed sampling data. In recent years, many HCS reconstruction algorithms have been proposed according to the structure of HSIs. The most common method is to convert the HSIs into a sparse description form by employing some proper sparsity priors. For example, some effective sparsity terms with ℓ_0, ℓ_1, and ℓ_q $(0 < q < 1)$ norms have been presented to characterize the sparsity [3]–[5]; however, these terms neglect the underlying structure information. Mun and Fowler [6] combined block-based random sampling with projection-driven CS reconstruction, such that the sparseness in the direction transformation domain and the smooth reconstructed image were enhanced at the same time. 3-D CS combines the 3D sparse, smooth, and low-rank priors of HSIs to achieve a certain reconstruction performance [7]. Jia *et al.* [8] proposed the structural similarity prior to HSIs and added it to the objective function, further improving the reconstruction performance. Note that tensor analysis has been applied

Manuscript received February 5, 2020; revised March 23, 2020 and April 23, 2020; accepted May 9, 2020. Date of publication May 15, 2020; date of current version May 29, 2020. This work was supported in part by the Key Logistics Research Project under Grant BLJ18J005, in part by the National Natural Science Foundation of China under Grant 41601344, in part by the Overseas Visiting and Research Project for Excellent Young Key Talents in Higher Education Institutions in Anhui Province under Grant gxgwfx2019056, and in part by the Quality Engineering Project of Universities of Anhui Province under Grant 2016zy126. *(Corresponding author: Yongjian Nian.)*

Zhongliang Wang is with the Department of Electric Engineering, Tongling University, Tongling 244061, China (e-mail: asdwzl@hotmail.com).

Mi He and Yongjian Nian are with the College of Biomedical Engineering and Imaging Medicine, Army Medical University (Third Military Medical University), Chongqing 400038, China (e-mail: hmcherry@126.com; yjnian@126.com).

Zhen Ye is with the School of Electronics and Control Engineering, Chang'an University, Xi'an 710064, China (e-mail: yezhen525@126.com).

Ke Xu is with the College of Electronic Science, National University of Defense and Technology, Changsha 410073, China (e-mail: xuke@nudt.edu.cn).

Bormin Huang is with the School of Information Science and Technology, Southwest Jiaotong University, Chengdu 611756, China (e-mail: bormin.uw@gmail.com).

Digital Object Identifier 10.1109/JSTARS.2020.2994334

to model the spatial–spectral correlation and local smoothness in HSIs, and tensor Tucker decomposition has also been employed to describe the global spatial–spectral correlation between bands [9]. To encode the essential structured sparsity of HSIs and explore the associated advantages for the HCS reconstruction task, a nonlocal tensor sparse and low-rank regularization approach is proposed [10]. A hierarchical reweighted Laplace prior is proposed to model the distribution of sparsity, which relieves the undemocratic penalization of the traditional Laplace prior to nonzero coefficients of a sparse signal [11]. Note that the above algorithms reconstruct the entire hyperspectral data directly; although the convex optimization algorithms, by adding various prior information, have obtained certain reconstruction performance, they usually have high computational complexity due to the big data of HSIs. Furthermore, the design of prior information also has a large impact on reconstruction quality, and unreasonable prior information significantly limits the degree of reconstruction quality.

Compressive-projection principal component analysis (CPPCA) [12] and its extended algorithms [13], [14] explore principal component characteristics of HSIs between the sampling end and reconstruction end, where spectral compressed sampling is the main method for hyperspectral data collection. The reconstruction process consists of eigenvector reconstruction based on convex-set optimization instead of directly reconstructing the original data, which can effectively reduce the reconstruction complexity. Experimental results show that CPPCA and its extended algorithms have the advantages of good reconstruction quality and fast reconstruction speed.

HCS reconstruction based on spectral unmixing reconstructs HSIs from the perspective of matrix decomposition, which has received substantial attention in recent years [15]–[24]. Note that the linear mixed model (LMM), one of the most popular models used in spectral unmixing-based reconstruction, assumes that spectral curves can be considered as the product of the endmembers of various ground objects and their corresponding abundances. Under this hypothesis, the main task of spectral unmixing-based reconstruction is to reconstruct the endmember and abundance of HSIs. Recent works show that spectral unmixing-based reconstruction achieves a perfect reconstruction quality with low reconstruction complexity. Assuming the endmembers are known, the problem of HCS reconstruction can be translated into an estimation of the abundance matrix [15], [16], [18]. In our previous work [22], [23], even if the endmember is unknown, by using a novel sampling matrix, spatially and spectrally separated compressed sampling data of HSIs can be obtained, and the endmember and abundance can be extracted from the above two data points for reconstruction. Spectral compressive acquisition (SpeCA) [19] employs a dual measurement matrix to acquire the compressed sampling data of the original HSIs; its decoder employs vertex component analysis (VCA) [25] to extract the endmember, followed by abundance estimation using the least squares method for reconstruction.

Note that the existing spectral unmixing-based reconstruction algorithms are all based on the LMM assumption. However, the illumination conditions and topography, atmospheric effects, or even the intrinsic variability of the material may cause a variation in a spectral signature for a given material. Moreover, material nonlinear mixing effects caused by multiple scattering and intimate mixing are also of significant influence. The above two main factors, spectral variability and nonlinearity, limit the ability of the LMM to achieve high performance. In view of the above problems, Thouvenin *et al*. [26] added a perturbation term for each endmember to correct spectral variation. The extended LMM (ELMM) [27] introduces scaled versions of reference endmembers to account for the pixelwise spatially coherent local variation in the endmembers. However, it is still difficult to accurately describe the spectral variation using only scalar scaling. Hong *et al*. [28] combined the ELMM and perturbed LMM to correct for spectral variation by employing a spectral variability dictionary, and reasonable prior knowledge for the spectral variability dictionary was also proposed for spectral unmixing. In [29], based on LMM, spectral disturbance is corrected by introducing the spectral correction term, and an LMM for spectral perturbation correction is proposed, which achieve better HCS reconstruction performance than LMM-based HCS algorithm. However, the above algorithm does not consider the nonlinear mixing factor. To my best knowledge, various modified LMM have been successfully applied to spectral unmixing; however, there is still a lack of model research on HCS reconstruction.

Intimate mixture and multilayered scene are two major scenarios inducing nonlinear mixing [1]. The most popular approaches to deal with intimate mixtures are the photometric model of Hapke [30], involving multiple scattering between different materials. In the multilayered configurations, a given material reflects off other materials before reaching the sensor, such as the interaction between vegetation and soil. Based on these concepts, several simplified nonlinear mixing models have been proposed to analytically describe these interactions. The bilinear mixture model (BMM) is an extension of the LMM with a term added to represent the contribution of reflectance between each pair of the observed materials. This reflectance term commonly involves the product of every pair of endmember spectra. However, BMM does not generalize the LMM and results in worse abundance estimation than linear spectral mixture analysis when existing high correlation between the original endmembers and the new introducing endmembers. To alleviate this issue, Halimi *et al*. [31] introduced the generalized bilinear model (GBM) to effectively deal with the underlying assumptions in the BMM. The good results on synthetic and real hyperspectral data show the significant advantage of GBM.

To address the limitations of LMM in HCS reconstruction, the purpose of this article is to reconstruct HSIs from spectral compressed sampling data by establishing a new model by considering the other factors based on LMM. More specifically, the contributions of this article can be summarized as follows.

1) We propose a novel mixing model, called the multitype mixing model (MMM), which simultaneously corrects spectral variation, nonlinear effects, and residual terms. In particular, a correction matrix, which can flexibly correct for spectral variation, is proposed instead of scale factors. Moreover, nonlinear effects and residuals are taken into account to improve the reconstruction accuracy.

2) An HCS framework is proposed according to the proposed MMM, which is succinctly called HCS_MMM. Under the condition that the endmember is known, reasonable prior knowledge is explored to recover the correction terms, and the alternating direction method of multipliers (ADMMs) [32] are utilized to solve the optimization problem to obtain the reconstruction of HSIs.

The rest of this article is organized as follows. In Section II, the proposed MMM is described. In Section III, the proposed HCS algorithm is given. In Section IV, the experimental results and discussion on a real hyperspectral dataset are shown, and finally, Section V gives some concluding remarks.

II. Multitype Mixing Model

The original hyperspectral data are denoted as $\mathbf{X} \in \mathbb{R}^{L\times N}$, where L is the number of bands and N is the total number of pixels in each band. The LMM represents HSIs as the product of an endmember matrix and abundance matrix, that is, $\mathbf{X} = \mathbf{ES}$, where $\mathbf{E} \in \mathbb{R}^{L\times p}$ is the endmember matrix with the ith column vector $\{\mathbf{e}_i\}_{i=1}^{p} \in \mathbb{R}^{L\times 1}$ representing the ith endmember spectrum, p is the number of endmembers, and $\mathbf{S} \in \mathbb{R}^{p\times N}$ denotes the abundant matrix.

Indeed, the LMM provides a simple but effective approximation of the physical process underlying the observations, which is extensively used in the field of spectral unmixing. However, the LMM may be inaccurate in some other specific situations; for instance, multiple scattering and absorption phenomena result in highly nonlinear effects. At this point, nonlinearly mixing spectral signatures should be considered [30], [33]. Another main factor hindering the LMM's ability to yield high performance is spectral variability, referring to a variation in a spectral signature for a given material [34]. Therefore, both nonlinear effects and spectral variability should be taken into account to establish a more accurate model [35]. Note that most of the existing studies are aimed at spectral unmixing, while this article focuses on the model design of HCS reconstruction. Based on the LMM, this article proposes an MMM that can effectively combine nonlinear effects and spectral variability of HSIs. The proposed model can be written as follows:

$$\mathbf{X} = \mathbf{ES} + \Gamma\mathbf{EH} + \mathbf{MG} + \mathbf{R} \tag{1}$$

where $\Gamma \in \mathbb{R}^{L\times L}$ is the correction matrix of the endmember; $\mathbf{M} \in \mathbb{R}^{L\times q}$ is the bilinear endmember matrix; $\mathbf{H} \in \mathbb{R}^{p\times N}$ and $\mathbf{G} \in \mathbb{R}^{q\times N}$ correspond to the abundance matrices of spectral variability and nonlinearity, respectively; and $\mathbf{R} \in \mathbb{R}^{L\times N}$ is a sparse residual term accounting for outliers caused by other factors in addition to nonlinearity and spectral variability. The first term on the right side of (1) expresses the linear mixed effects, the second term describes the impact of spectral variability, the third term illustrates nonlinear effects, and the last term is used to compensate for model mismatch caused by other factors.

The LMM is a simple and effective hyperspectral model. However, since multiple scattering and intimate mixing, nonlinearity impedes the effectiveness of LMM [30]. Therefore, in the proposed MMM expressed by (1), we used a GBM [31] to describe the nonlinear effects. The kth column vector of $\mathbf{M}$ is defined as follows:

$$\mathbf{m}_k \equiv \mathbf{e}_i \odot \mathbf{e}_j \quad \text{s.t. } \{(i,j)\,|i<j, i \in \{1,\ldots,p\}, j \in \{1,\ldots,p\}\},\ k=1,\ldots,q \tag{2}$$

where $\odot$ is the Hadamard (elementwise product) operation. The constraint $i < j$ defined in (3) is based on the commutative property of the Hadamard operation, i.e., $\mathbf{e}_i \odot \mathbf{e}_j = \mathbf{e}_j \odot \mathbf{e}_i$. Normally, q is the number of columns of $\mathbf{M}$, and $q = p(p-1)/2$. For spectral unmixing, since p is usually small, making, $q < L$ the estimation of abundance is an over-determined problem. However, for the spectral CS discussed in this article, this condition is difficult to meet. Therefore, when q is larger than the threshold $(\alpha v, 0.5 \le \alpha \le 1)$, $\mathbf{M}$ should be trimmed to avoid the occurrence of underdetermined problems and strong similarities between $\mathbf{m}_k$ when conducting the CS reconstruction process. The pruned version was obtained by simply removing some $\mathbf{m}_k$ values associated with smaller spectral angles.

Another factor that reduces the effectiveness of LMM is spectral variability, caused by geometry and topography of the scene. The scaling factors acting on spectrums are proposed by ELMM to correct the variability [27], [28]. However, the spectral variability of different bands may be inconsistent. A single-scale factor is difficult to accurately correct this variability. Therefore, we employ a correction matrix Γ in (1) to correct different bands with different scales. In other words, the correction method of the spectral curve for different materials may be different. Therefore, even for compressed sampling data, the spectral variability in (1) can suitably compensate for the effects of the spectral variability. Note that this article employs the total variation (TV) constraint for the calculation of Γ, which will be described in Section III in detail.

III. HCS Based on the MMM

A. Spectral Compressed Sampling of HSIs

In recent years, several spectral libraries have been established, such as the United States Geological Survey (USGS) spectral library [36], Jet Propulsion Laboratory spectral library, and Advanced Spaceborne Thermal Emission and Reflection Radiometer spectral library. Therefore, it is feasible to assume that the endmember matrix $\mathbf{E}$ is known [11], [18], [20], [24]. In this case, a reasonable sampling scheme should be determined to achieve good reconstruction quality based on the MMM.

The compressed sampling model of an HSIs usually includes spatial and spectral sampling. In this article, under the condition that the endmember is known, only the spectral compressed sampling manner is employed because the spatial compressed sampling manner may significantly destroy the abundance information of HSIs. This is also the most commonly used compressed sampling method for CS reconstruction algorithms based on spectral unmixing. Actually, we do not need to predict the endmember matrix by employing the spatio–spectral hybrid compressed sampling proposed in our previous work [22], [23], which can extract endmember easily from the compressed data. But this is beyond the scope of this article. The purpose of

the article is to explore the CS reconstruction algorithm using MMM, rather than the compressed sampling method. In addition, spectral compressed sampling is a simple and efficient data acquisition method, and both the pushbroom and the whiskbroom imaging systems are relatively easy to implement in hardware [19], [37], [38]. The spectral sampling based on the MMM can be expressed as

$$\mathbf{Y} = \mathbf{AX} = \mathbf{AES} + \mathbf{A\Gamma EH} + \mathbf{AMG} + \mathbf{AR} \tag{3}$$

where $\mathbf{Y} \in \mathbb{R}^{v \times N}$ denotes the observed data with v bands collected by spectral compressed sampling ($v \ll L$) and $\mathbf{A} \in \mathbb{R}^{v \times L}$ is the Gaussian random 0-1 observation matrix, which is convenient for hardware implementation of optical devices such as digital micromirror devices. Obviously, the sampling rate of spectral compressed sampling can be defined as $SR = v/L$.

As for the spectral sampling data of HSIs, compared with the original HSIs, the number of band is compressed, that is to say, the spectral resolution of spectral sampling data is reduced, where the degree of reduction is determined by the sampling rate. However, the spatial resolution of spectral sampling data is consistent with the original HSIs. Therefore, the following reconstruction process can be regarded as a spectral super-resolution reconstruction of HSIs.

B. HCS Reconstruction Based on the MMM

The goal of the reconstruction is to recover the original HSIs from the reduced spectral sampling data as much as possible. Given the endmember matrix $\mathbf{E}$, observation data $\mathbf{Y}$, and measurement matrix $\mathbf{A}$. Equation (3) shows that the abundance matrices $\mathbf{S}, \mathbf{H}, \mathbf{G}$; correction matrix $\mathbf{\Gamma}$; and sparse residual term $\mathbf{R}$ should be estimated. However, solving these variables directly is a nonconvex optimization problem. The ADMM is employed for the alternate iterative estimation of above variables [32].

We note that the solution of the abundance matrices $\mathbf{S}, \mathbf{H}, \mathbf{G}$ is actually an over-determined problem, given other variables. Although adding prior information (such as the sparsity, TV, sum-to-one, and nonnegative constraints of abundance [39]) helps obtain the optimal solution, compared to the solution of underdetermined problems, over-determined problems more easily yield optimal solutions. Increasing the constraint term of abundance may not help to improve the reconstruction quality and indeed increases the complexity of the solution. Therefore, we use only the least squares method to estimate matrices $\mathbf{S}$, $\mathbf{H}$, and $\mathbf{G}$.

However, the estimations of the correction matrix $\mathbf{\Gamma}$ and the sparse residual term $\mathbf{R}$ are optimization problems for underdetermined equations. These components are difficult to solve directly, and reasonable priors need to be introduced.

The second item on the right of model (1) is to compensate for the LMM mismatch caused by endmember spectral variation. The rows of the matrix $\mathbf{\Gamma}$are used to respond to the interaction between different bands. In general, the interaction between the bands decreases rapidly as the band distance increases. It means that the difference between the adjacent positions of the matrix $\mathbf{\Gamma}$ in the horizontal direction will be very sparse. In other words, the matrix $\mathbf{\Gamma}$ should satisfy a smooth row. Simultaneously, the columns of the matrix $\mathbf{\Gamma}$ represent the weights of endmember spectral of different materials. Since the third item on the right of model (1) can effectively capture the nonlinear effect between the endmember. Then, the second term simply compensates for the variation of each endmember. So, the column vector of the matrix $\mathbf{\Gamma}$ may be a sparse vector. Of course, the difference along the column direction is also sparse. Therefore, in this article, we employs the TV constraint for $\mathbf{\Gamma}$ and the sparse constraint for $\mathbf{R}$. Then, the reconstruction problem can be formulated as the optimization of the following objective function:

$$\min_{\mathbf{S},\mathbf{H},\mathbf{B}} \frac{1}{2} \|\mathbf{Y} - \mathbf{AES} - \mathbf{A\Gamma EH} - \mathbf{AMG} - \mathbf{AR}\|_F^2 + \lambda_1 \mathrm{TV}(\mathbf{\Gamma}) + \lambda_2 \|\mathbf{R}\|_{2,1} \tag{4}$$

where $\|\mathbf{C}\|_F = \sqrt{\mathrm{trace}\{\mathbf{CC}^T\}}$ is the Frobenius norm of $\mathbf{C}$. The superscript T represents the transpose of the matrix. $TV(\mathbf{\Gamma}) \triangleq \|\mathbf{F\Gamma}\|_{1,1} = \|\mathbf{F}_h\mathbf{\Gamma}\|_{1,1} + \|\mathbf{F}_v\mathbf{\Gamma}\|_{1,1}$ is defined as the spatially anisotropic TV norm [40], where $\mathbf{F}_h$ and $\mathbf{F}_v$ represent the discrete gradient operators in the horizontal and vertical directions, respectively. $\|\mathbf{C}\|_{1,1} = \sum_{i=1}^{L} \|\mathbf{C}_i\|_1$, where $\mathbf{C}_i$ is the ith column vector of $\mathbf{C}$. $\|\mathbf{R}\|_{2,1}$ is the so-called $\ell_{2,1}$-norm defined by $\|\mathbf{R}\|_{2,1} = \sum_{i=1}^{N} \|\mathbf{R}_i\|_2$, depicting the row sparsity of $\mathbf{R}$. Finally, $\lambda_1 \geq 0$ and $\lambda_2 \geq 0$ are the regularization parameters, which are used to tradeoff the TV terms and sparse terms.

Note that the ADMM [32] is employed to solve the optimization problem given by (4), which is detailed in Algorithm 1, and the solution to each subproblem is given in Appendix.

IV. Experimental Results and Discussion

In this section, we evaluate the relevance of the MMM proposed in Section II and the corresponding reconstruction accuracy of the HCS_MMM algorithm described in Section III using synthetic and real data. All the experiments are run with MATLAB 2018a on a laptop workstation with a 2.6-GHz CPU and 32-GB RAM.

For the performance assessment of the reconstruction quality, the average signal-to-noise ratio (aSNR) and the average spectral angle mapper (aSAM) were both employed, where the aSNR (measured in dB) is defined as

$$\mathrm{aSNR} = \frac{1}{L} \sum_{i=1}^{L} 10 \log_{10} \frac{\|\mathbf{X}_i\|_2^2}{\left\|\mathbf{X}_i - \hat{\mathbf{X}}_i\right\|_2^2} \tag{5}$$

where $\mathbf{X}_i$ and $\hat{\mathbf{X}}_i$ represent the original image and the reconstruction one of ith band, respectively. The aSAM is defined as

$$\mathrm{aSAM} = \frac{1}{N} \sum_{j=1}^{N} \arccos \frac{\mathbf{X}_j^T \cdot \hat{\mathbf{X}}_j}{\|\mathbf{X}_j\|_2 \cdot \left\|\hat{\mathbf{X}}_j\right\|_2} \tag{6}$$

where $\mathbf{X}_j$ and $\hat{\mathbf{X}}_j$ represent the spectral vectors of the jth pixel of the original and reconstructed images, respectively. Note that aSAM can effectively evaluate the spectral similarity between the original and the reconstruction spectrums, a small aSAM

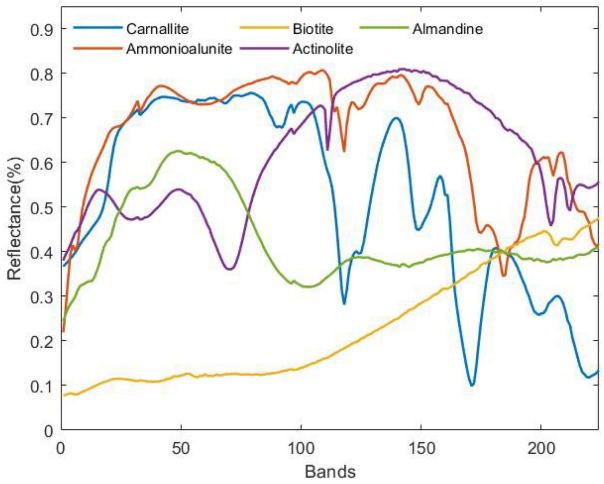


Fig. 1. Reflectance of five endmember spectra.

means a high similarity between the reconstructed imagery and the original one.

A. Experiments With Synthetic Data

In this section, the parameter settings of the proposed HCS_MMM algorithm are tested by synthetic experiments. A 64×64-pixel image composed of five pure spectral components is generated according to the MMM. The endmember spectral signatures are selected from the USGS digital spectral library [36]. Fig. 1 shows the spectral curves of the five materials used in the synthetic data, and the number of bands is 224. The synthetic data are generated according to the following pattern: 60% of the image pixels are generated according to the standard LMM, 20% are generated according to a model that features nonlinear component interactions, and the remaining 20% are added by a spectral variability model described in the second term on the right side of (3). The corresponding abundance fractions for the three models (linear, spectral variability, and nonlinear) are generated by squares, horizontal stripes, and vertical stripes, respectively, and their shapes can be observed from the visual comparison of the synthetic data shown in Fig. 4(a).

The observed data are collected using spectral compressed sampling. The sampling matrix is a binary matrix with 0 and 1 obtained by rounding a uniformly distributed random matrix. The proposed algorithms were tested using different values of the parameters λ_1 and λ_2 on different sampling rata. Fig. 2 reports the aSNR results achieved by different parameter. From Fig. 2, we can see that the results of parameter λ_2 at different sampling rates are basically the same and can maintain high reconstruction accuracy over a wide range. At a higher sampling rate, a noticeable peak in aSNR appears when λ_2 is around 1. At low sampling rates, however, aSNR curves fluctuate little in the $1\text{-}10^4$ range. It means that λ_2 is reasonable in this interval. Therefore, in the following experiments, we set a normal value of 10^3 for parameter λ_2. But aSNR is more sensitive to the parameter λ_1, which can be observed from the experimental results of λ_1. With different sampling rates, the variation trend of aSNR curve varies greatly with λ_1. Fortunately, when $\lambda_1 = 0.1$, the aSNR of all sampling rates can get a higher value. Therefore, we set $\lambda_1 = 0.1$ for the following experiments.

Next, we compare the proposed algorithm with other state-of-the-art spectral unmixing based algorithm, such as SSHCS [22], SpeCA [19], and SSCR_SU [23], on the synthetic data by varying sampling rata from 0.1 to 0.5, with a 0.1 step size. The number of endmembers are set to 5 for all algorithm, and the endmembers of HCS_MMM, SSHCS, and SSCR_SU are extracted by VCA algorithm from the original data. Since the sampling rate of SSHCS and SSCR_SU includes spatial and spectral part. We fix the spatial sampling to 0.01 according to the suggestions by corresponding references. For the SpeCA algorithm, only the number of endmembers and sampling rate need to be set, and the other parameters can be set by default.

Fig. 3 shows the aSNR comparison results of the different algorithms. The proposed HCS_MMM algorithm can achieve far better performance than other spectral unmixing algorithms. At 0.5 sampling rate, the reconstructed aSNR even exceeds SSCR_SU by nearly 10 dB. The performance advantage decreases as the sampling rate decreases. Even so, the aSNR of 5 dB over other algorithms can still be achived, when the sampling rate is 0.1. The reconstruction performance of SSHCS and SSCR_SU is similar for synthetic data, but SpeCA had the worst results.

To visually demonstrate the reconstruction results, we present pseudocolor images (top) and local enlarged views (bottom) with bands (178, 94, 89) of the original synthetic data and the reconstructed data obtained by all algorithms under sampling rates of 0.2 in Fig. 4. However, it is difficult to find the difference between the reconstructed image and the original image. Actually, since the spectral observation data is mainly used in the reconstruction of the several algorithms, there is little disturbance to spatial information. Therefore, the reconstruction preserves the spatial details intact, but the spectrum is distorted. Spectral distortion in pseudocolor images is a slight variation in color, which is difficult to distinguish visually, even in the local enlarged views.

In order to verify that the performance of the proposed HCS_MMM algorithm is significantly better than other algorithms, we provide the 178th bands residual images between original and reconstructed data in Fig. 5. Since all the competing methods can obtain good reconstructed results, the residuals will be small. For the sake of observation, we magnify the residuals of all algorithms by a factor of 20. Note that Fig. 5(a) is the 178th bands original image, the others are residual images. Here, the similarity of the residual image to the original image does not indicate that the reconstruction performance is better. Actually, the higher the brightness of the images, the larger the residuals; conversely, the smaller the residuals. Now, it is obvious that the residual of the proposed HCS_MMM algorithm is much smaller than that of other algorithms. The brightness of residual images achieved by SpeCA and SSCR_SU is similar, but SSHCS is slightly darker.

Since the several algorithms based on spectral unmixing can effectively avoid the distortion of spatial information and are at a disadvantage in terms of the consistency of spectral signatures, it is necessary to show the comparison between the original and the reconstructed spectral curve achieved by

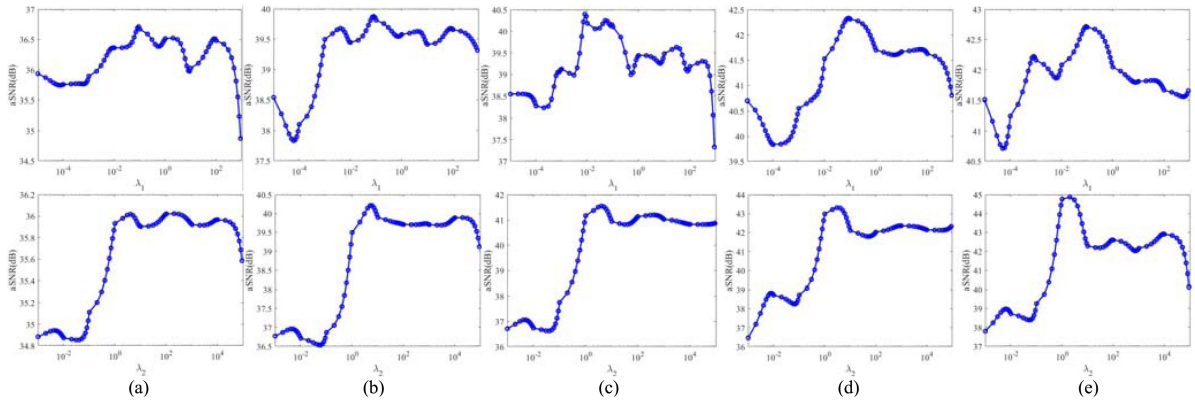


Fig. 2. Regularization parameter sensitivity analysis of the proposed HCS_MMM algorithm with different sampling rate. (a) $SR = 0.1$; (b) $SR = 0.2$; (c) $SR = 0.3$; (d) $SR = 0.4$; (e) $SR = 0.5$.

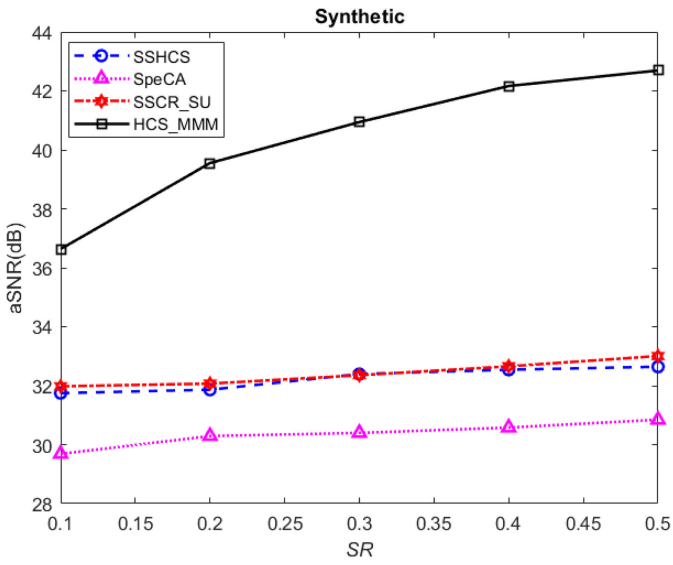


Fig. 3. Comparison of the aSNR values of the reconstructed images achieved by the various algorithms for synthetic data.

various algorithms. Fig. 6 shows the comparison results under 0.2 sampling rate. The reconstructed spectral of HCS_MMM can approach the original curve well. But other algorithms deviate greatly, especially SpeCA. The experiment results of Figs. 5 and 6 confirm the effectiveness of the proposed model and reconstruction algorithm again.

B. Experiments With Real Data

1) Description of the Datasets and Compared Methods: In this section, we choose six real HSIs since their partial ground truth is publicly available. The first image dataset is Samson, which is a simple dataset that is available online. There are 952 × 952 pixels and 156 channels covering the wavelengths from 401 to 889 nm. A region of 95 × 95 pixels is used in our experiments. Three classes, soil, tree, and water, remain.[1] The RGB composite of the Samson data (bands 105, 50, and 18 for red, green, and blue, respectively) is shown in the first line of Fig. 9(a).

Jasper Ridge, a popular hyperspectral dataset used in hyperspectral unmixing analyses, includes 512 × 614 pixels and 224 channels ranging from 380 to 2500 nm. In subsequent experiments, we consider a subimage of 100 × 100 pixels and retain 198 channels after removing some channels (due to the dense water vapour and atmospheric effects). There are four endmembers latent in these data. The RGB composite of the Jasper Ridge data (bands 33, 15, and 4 for red, green, and blue, respectively) is shown in the second line of Fig. 9(a).

The third dataset used in our experiments, University of Pavia (PaviaU), is an urban image acquired by the Reflective Optics System Imaging Spectrometer sensor during a flight campaign over Pavia, northern Italy [45]. After removing the 12 noisiest bands, 103 spectral channels and a region with a spatial coverage of 610 × 340 pixels remain. The ground truth available from the PaviaU data is divided into nine classes. The RGB composite of the PaviaU data (bands 28, 11, and 2 for red, green, and blue, respectively) is shown in the fourth line of Fig. 9(a).

The fourth HSIs group is taken from the Urban dataset, which is one of the most widely used data used in hyperspectral unmixing studies. There are 307 × 307 pixels and 210 wavelengths ranging from 400 to 2500 nm. Due to water absorption and atmospheric effects, we reduced 210 bands to 162 bands by removing bands 1-4, 76, 87, 101-111, 136-153, and 198-210. Six classes representing the various land cover types that occur in this environment were defined for the site. The RGB composite of the Urban data (bands 28, 11, and 2 for red, green, and blue, respectively) is shown in the third line of Fig. 9(a).

For the last two test images, we employed the Salinas and Salinas-A hyperspectral datasets. The Salinas data were collected by the 224-band AVIRIS sensor over Salinas Valley, California. The scene covered comprises 512 lines by 217 samples. A total of 204 spectral channels remained after discarding the 20 water absorption bands. The Salinas scene contains 16 classes, including vegetables, bare soils, and vineyard fields. Salinas-A is a small subscene of the Salinas image, comprising 86 × 83 pixels and including six classes. The RGB composites of the

1.[Online]. Available: http://www.escience.cn/people/feiyunZHU/Dataset_GT.html

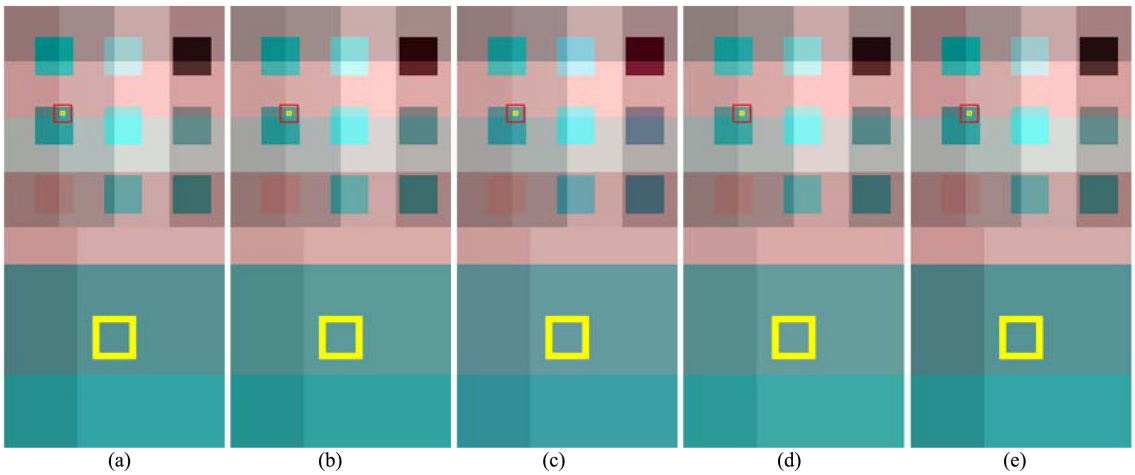


Fig. 4. Original and reconstructed synthetic images achieved by the various algorithms with a 0.2 sampling rate. (a) Original; (b) SSHCS; (c) SpeCA; (d) SSCR_SU; (e) HCS_MMM.

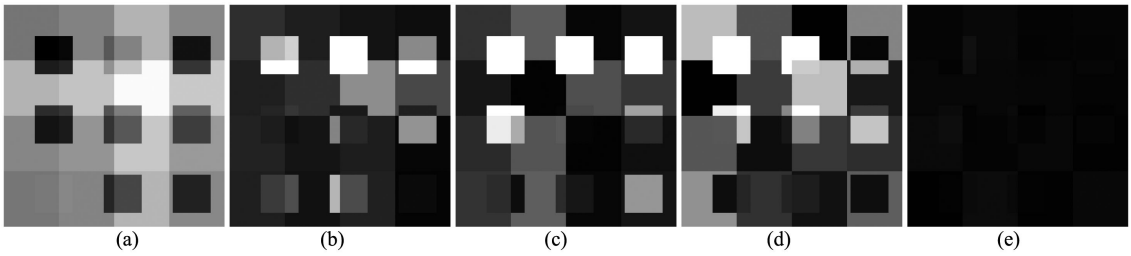


Fig. 5. Visual results of the 178th bands from original synthetic data and 20 times enhancement residuals between original and reconstructed data when sampling rate is 0.2. (a) Original image; (b) SSHCS residual (28.74 dB); (c) SpeCA residual (25.53 dB); (d) SSCR_SU residual (26.57 dB); (e) HCS_MMM residual (48.69 dB).

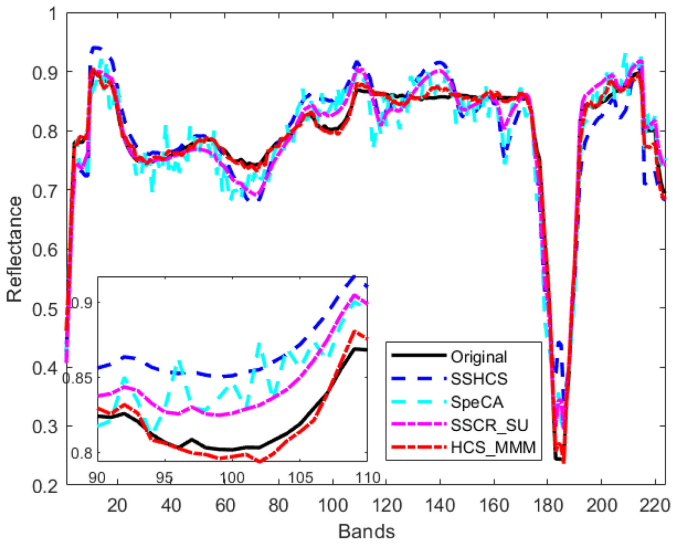


Fig. 6. Original and reconstructed spectral curves for synthetic data.

Salinas and Salinas-A data (bands 40, 20, and 10 for red, green, and blue, respectively) are shown in the fifth and sixth lines of Fig. 9(a), respectively.

In our real-data experiments, we compare the proposed algorithm with several state-of-the-art reconstruction algorithms of HCS using spectral unmixing based on an LMM, SSHCS [22], SpeCA [19], and SSCR_SU [23]. SSHCS and SSCR_SU can estimate the number of endmembers automatically. However, SpeCA does not specify the endmember number estimation algorithm. Therefore, for the sake of fairness, the number of endmembers for all algorithm experiments is the same as the number of classes included in the image ground truth.

In this section, the compressed data are collected by a random Gaussian sampling matrix $\mathbf{A}$ with a sampling rate ranging from 0.1 to 0.5 for our HCS_MMM algorithm. The other three comparison algorithms adopt the sampling patterns given by their references. In this study, the proposed method assumes that the necessary endmembers are obtained from the spectrum library. Although the off-the-shelf spectrum library contains a large number of standard spectra of different objects, the standard spectra do not necessarily match the spectra in the real scene on account of the various weather and imaging conditions. Therefore, we implement the VCA method on each dataset to extract certain endmembers.

2) Performance Comparison: Fig. 7 shows the results of the four algorithms in terms of the aSNR with different sampling rates. The results show that the reconstruction performance

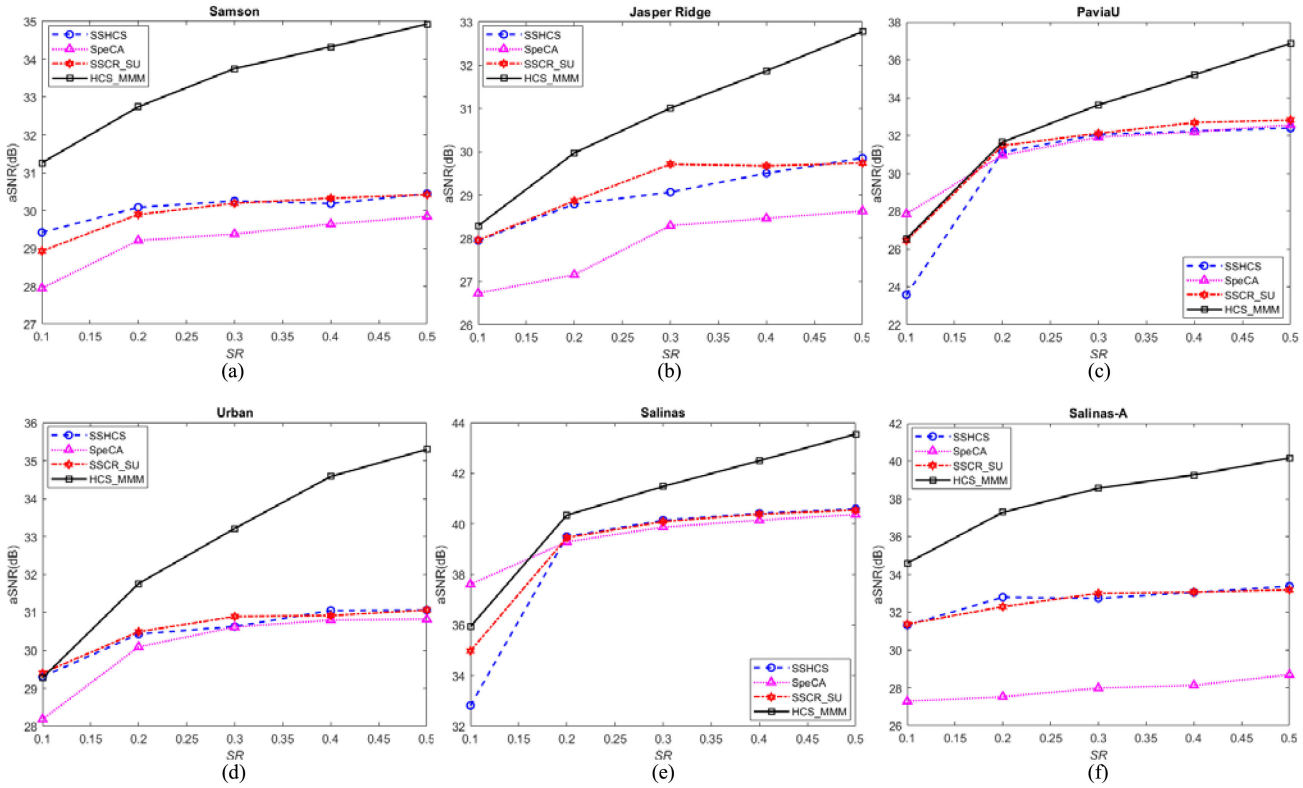


Fig. 7. Comparison of the aSNR values of the reconstructed images achieved by the various algorithms. (a) Samson data; (b) Jasper ridge data; (c) PaviaU data; (d) Urban data; (e) Salinas data; and (f) Salinas-A data.

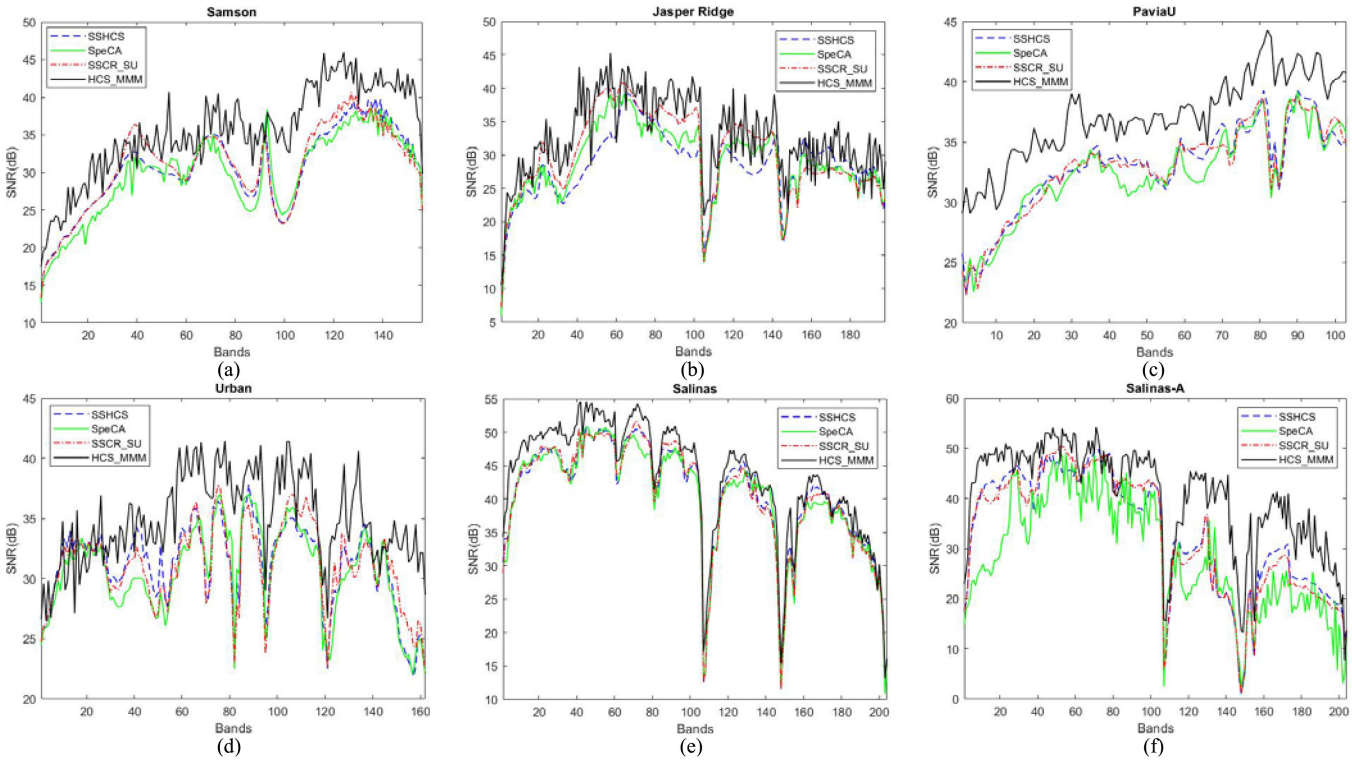


Fig. 8. SNR per band for different datasets with a 0.5 sampling rate. (a) Samson data; (b) Jasper ridge data; (c) PaviaU data; (d) Urban data; (e) Salinas data; and (f) Salinas-A data.

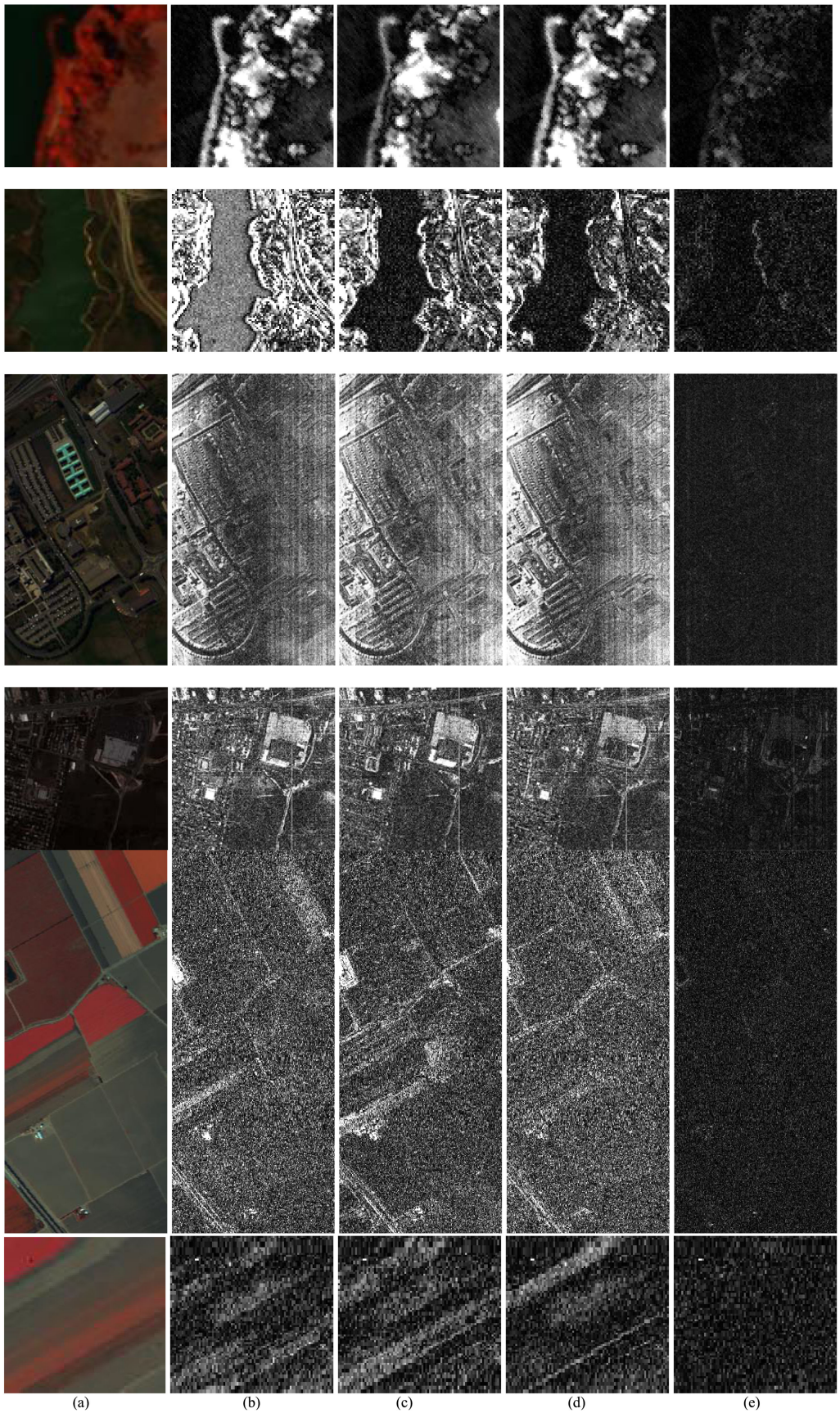


Fig. 9. Original pseudocolor and residual images for the different datasets with a 0.5 sampling rate; from top to bottom, Samson data, Jasper Ridge data, PaviaU data, Urban data, Salinas data, and Salinas-A data. (a) Original; (b) SSHCS; (c) SpeCA; (d) SSCR_SU; and (e) HCS_MMM.

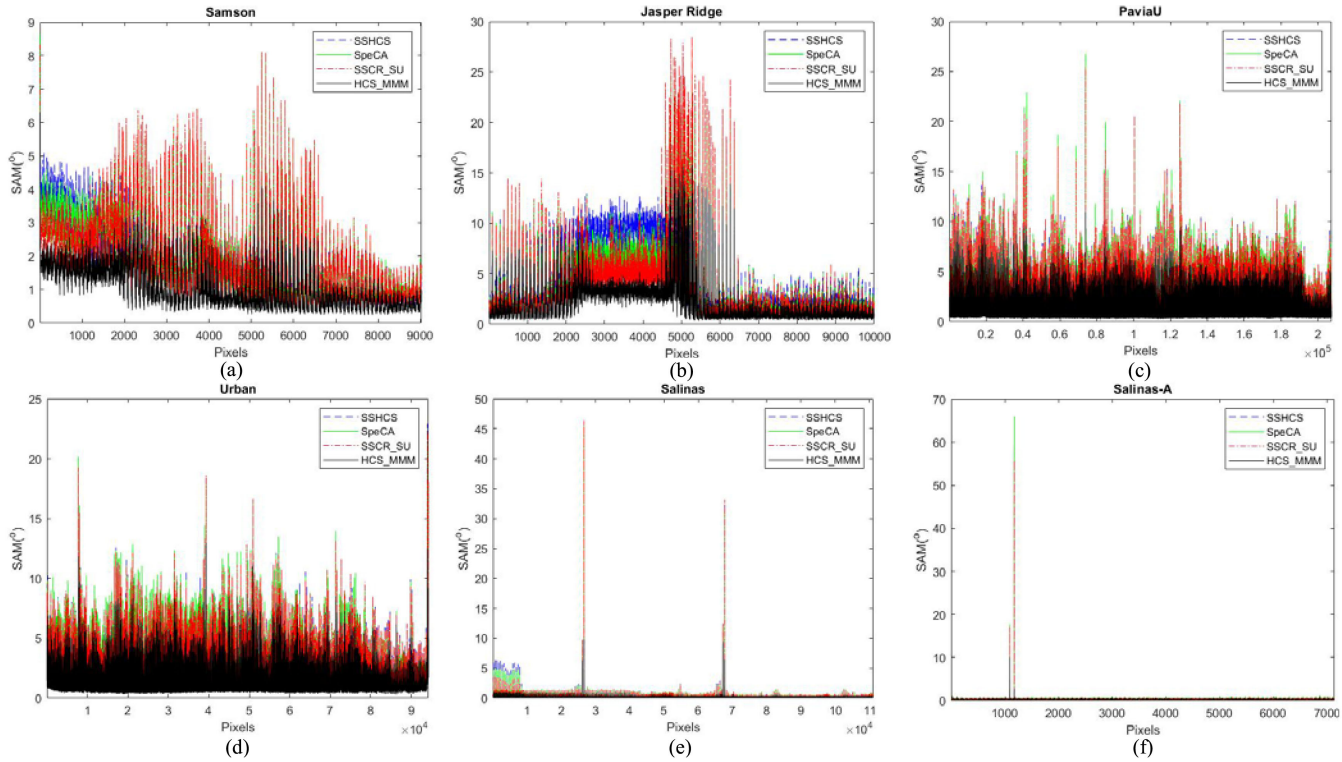


Fig. 10. SAM per pixel for different datasets with a 0.5 sampling rate. (a) Samson data; (b) Jasper Ridge data; (c) PaviaU data; (d) Urban data; (e) Salinas data; and (f) Salinas-A data.

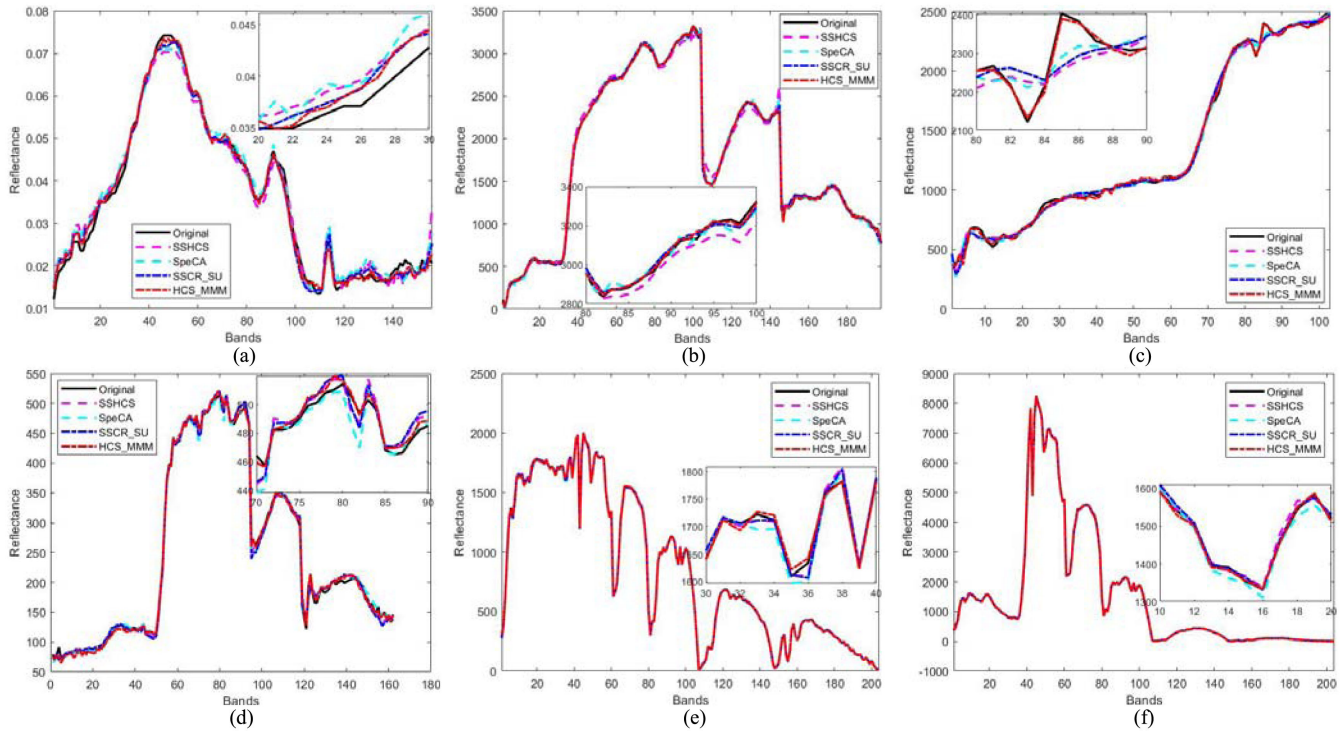


Fig. 11. Original and reconstructed spectral curves for different pixels with a 0.5 sampling rate. (a) Samson data (10,10); (b) Jasper Ridge data (10,10); (c) PaviaU data (10,10); (d) Urban data (100,100); (e) Salinas data (10,10); and (f) Salinas-A data (10,10).

TABLE I
COMPARISON OF aSAM (°) ACHIEVED BY THE VARIOUS ALGORITHMS (THE BEST RESULTS ARE IN BOLD)

SR	0.1	0.2	0.3	0.4	0.5
Results on the Samson dataset					
SSHCS	2.1036	2.3969	1.9597	2.0213	2.0223
SpeCA	2.7965	2.1644	2.1546	2.0875	2.0372
SSCR_SU	2.2828	2.0603	2.0306	2.0415	2.0288
HCS_MMM	**1.7258**	**1.4241**	**1.2963**	**1.1832**	**1.0759**
Results on the Jasper Ridge dataset					
SSHCS	4.1412	3.7857	3.7346	3.7063	3.351
SpeCA	4.7481	4.4319	4.2536	3.8976	3.803
SSCR_SU	4.1432	3.6532	3.5906	3.59	3.5772
HCS_MMM	**3.52**	**2.943**	**2.5138**	**2.2383**	**1.9149**
Results on the PaviaU dataset					
SSHCS	4.7104	1.714	1.5518	1.4928	1.4635
SpeCA	**2.3019**	1.7051	1.5262	1.4777	1.4497
SSCR_SU	3.4129	1.6725	1.5277	1.4201	1.4046
HCS_MMM	3.1151	**1.4937**	**1.2058**	**1.0375**	**0.8971**
Results on the Urban dataset					
SSHCS	2.2701	2.0498	2.0039	1.9533	1.9512
SpeCA	2.4454	2.0726	2.0175	1.9728	1.9829
SSCR_SU	**2.2292**	2.0415	1.9864	1.9601	1.9436
HCS_MMM	2.3275	**1.7199**	**1.4364**	**1.2305**	**0.1053**
Results on the Salinas dataset					
SSHCS	0.8719	0.3439	0.3079	0.2968	0.2904
SpeCA	**0.4187**	0.3414	0.3159	0.3038	0.2967
SSCR_SU	0.6717	0.3428	0.3096	0.297	0.2889
HCS_MMM	0.5892	**0.3262**	**0.2784**	**0.243**	**0.209**
Results on the Salinas-A dataset					
SSHCS	**0.4436**	0.4011	0.3631	0.3898	0.3882
SpeCA	0.5307	0.4489	0.4403	0.4237	0.4143
SSCR_SU	0.4506	0.4051	0.3978	0.3915	0.3897
HCS_MMM	0.5553	**0.3999**	**0.3307**	**0.2917**	**0.2532**

of all algorithms increases with increasing sampling rate. The values for SSHCS, SpeCA, and SSCR_SU increases slowly, while that of HCS_MMM increases rapidly. This illustrates that the proposed algorithm has a more obvious advantage over the others at higher sampling rates. For example, when the SNR is 0.5, HCS_MMM exceeds other competing methods on aSNR by at least 4.5 dB on the Samson dataset, 2.9 dB on the Jasper Ridge dataset, 4 dB on the PaviaU dataset, 4.2 dB on the Urban dataset, 3 dB on the Salinas dataset, and 6.8 dB on the Salinas-A dataset. When the sampling rate exceeds 0.2, the reconstruction quality of the HCS_MMM algorithm is significantly better than that of the other three algorithms. However, when the sampling rate is 0.1, the reconstruction accuracy of HCS_MMM is close to that of SSCR_SU of the Jasper Ridge, PaviaU, Urban, and Salinas datasets. This is because when the sampling rate is too low, the last three terms on the right side of (3) are difficult to estimate accurately. As a result, the multitype mixing model degrades to the LMM. Thus, HCS_MMM performs similarly to SSCR_SU for some datasets. However, for the Samson and Salinas-A datasets, the aSNR of HCS_MMM still exceeds that of SSCR_SU by 2.3 and 3.2 dB, respectively.

It is worth noting that although the Salinas-A scene is part of the Salinas dataset, their reconstruction results are quite different. The Salinas-A region is smaller, the ground truth contains fewer classes, and the reconstruction quality is lower. The performance of SpeCA is particularly evident. This may also mean that CS reconstruction based on spectral unmixing is more suitable for large-area, multiclass HSIs than for other images types.

SNR per band are also provided in Fig. 8. We can see that the reconstructed SNR of the HCS_MMM algorithm fluctuates greatly with the band, which may be caused by the estimation of the correction matrix $\mathbf{\Gamma}$ and sparse residual term $\mathbf{R}$. The

TABLE II
RUNTIME OF THE VARIOUS ALGORITHMS

SR	0.1	0.2	0.3	0.4	0.5
Results on the Samson dataset					
SSHCS	0.1327	0.0361	0.0369	0.0124	0.0199
SpeCA	0.424	0.4282	0.4133	0.4178	0.428
SSCR_SU	1.6556	1.5648	1.6250	1.5814	1.6010
HCS_MMM	7.7162	8.8022	10.2434	11.33	12.1502
Results on the Jasper Ridge dataset					
SSHCS	0.1415	0.0398	0.0406	0.0181	0.0225
SpeCA	1.2183	1.1908	1.1752	1.1759	1.1894
SSCR_SU	2.6319	2.5226	2.7644	2.6899	2.7933
HCS_MMM	9.4844	11.0458	12.9761	14.5479	15.9423
Results on the PaviaU dataset					
SSHCS	0.2170	0.1006	0.1092	0.0860	0.0941
SpeCA	35.6766	35.3031	35.3430	35.4372	37.5306
SSCR_SU	47.8656	47.7762	48.2199	49.3058	49.1932
HCS_MMM	86.0260	100.8798	117.1089	130.8514	142.9889
Results on the Urban dataset					
SSHCS	0.2133	0.0853	0.0897	0.0626	0.0709
SpeCA	19.2170	19.5150	19.4149	18.8062	19.1980
SSCR_SU	22.3427	21.7655	21.9956	21.1543	21.4681
HCS_MMM	73.5405	81.4823	92.8507	99.5101	114.3168
Results on the Salinas dataset					
SSHCS	0.1934	0.1144	0.1171	0.1024	0.1088
SpeCA	205.7025	204.9305	204.8881	204.8763	205.3877
SSCR_SU	33.416	33.9727	33.3061	33.6829	34.0547
HCS_MMM	98.8468	116.9637	132.2422	150.7448	166.1454
Results on the Salinas-A dataset					
SSHCS	0.1361	0.0354	0.0367	0.0143	0.0204
SpeCA	1.9514	2.0551	2.0488	2.1930	2.0554
SSCR_SU	2.0812	2.0263	2.0171	2.0677	2.0756
HCS_MMM	8.4254	9.4492	10.7745	11.7650	13.6022

estimation of $\mathbf{\Gamma}$ and $\mathbf{R}$ is an underdetermined problem. Even if the appropriate prior constraint is used, it is still difficult to find the accurate solution. Therefore, the reconstruction performance may be slightly different between different bands. However, the other three algorithms mainly solve positive definite or overdetermined problems and can obtain consistent analytical solutions. Even so, the lowest point of the SNR curve achieved by the proposed HCS_MMM algorithm is still higher than that of other algorithms on multiple datasets.

Similar to synthetic data, all algorithms can well preserve the spatial information of the image on real datasets, and the slightly visual differences are difficult to observe. So we also adopt the residual images to demonstrate the significant performance of the proposed algorithm for real datasets. Fig. 9 shows the original RGB composite and residual images between original data and reconstructed data achieved by different algorithms with a 0.5 sampling rate. Now, we can easily judge the advantages of each algorithm. The residual image of HCS_MMM is significantly darker than other algorithms, indicating that the residual of the proposed algorithm is smaller and the reconstruction accuracy is higher. When the sampling rate is 0.5, the reconstruction performance of the other three algorithms is similar according to Fig. 7. Therefore, the brightness of residual images for the three algorithms is also visually similar. Only on the Jasper Ridge data, the reconstruction of SSHCS algorithm in this band is unsatisfactory.

To further illustrate the superiority of the proposed HCS_MMM on spectrum reconstruction, we compare the aSAM values of the different algorithms. In Table I, we provide the performance of all methods using the aSAM results over all the above datasets.

The results in Table I are essentially consistent with the trends in Fig. 1. At higher sampling rates, the aSAM value reconstructed by the proposed algorithm is lower than those

of the other three spectral unmixing-based reconstruction algorithms, which verifies the effectiveness of HCS_MMM. When the sampling rate is 0.1, the realization of each algorithm on different datasets has its own advantages. For example, the HCS_MMM algorithm leads the pack for the Samson and Jasper Ridge datasets. For the PaviaU and Salinas datasets, SpeCA is the best. The SSCR_SU algorithm has the lowest aSAM for the Urban dataset, and the SSHCS algorithm has the best performance for the Salinas-A datasets with a 0.1 sampling rate.

Fig. 10 shows the reconstructed SAM per pixel for different data sets. The results of the proposed HCS_MMM algorithm are marked with a solid black line. From the first four data sets, we can easily distinguish that the SAM of the proposed algorithm is lower than other algorithms. For the latter two data sets, however, the observation of discrimination is not obvious due to the inappropriate spikes of SAM reconstructed by SpeCA and SSCR_SU. Only on the first 10 000 pixels of Salinas data, it can still be clearly observed that HCS_MMM is superior to other algorithms.

To further demonstrate the superiority of the proposed algorithm over the other algorithms on spectrum reconstruction, we illustrate the original and reconstructed spectral curves with a 0.5 sampling rate for the different pixels in the above datasets in Fig. 11. The spectral unmixing-based CS reconstruction algorithm can completely recover the original spectral curves, especially for the Salinas and Salinas-A datasets. However, the curves recovered by the proposed HCS_MMM algorithm are closer to the original curves than with several other algorithms, which can be seen from the enlarged subgraph in Fig. 11. The experimental results in Figs. 4 and 5 further confirm the superiority of the proposed algorithm, both in the reconstruction of spatial images and the recovery of spectral curves.

The runtime of the several algorithms is listed in Table II. Obviously, the reconstruction speed of SSHCS algorithm is the fastest, which is more than two orders of magnitude higher than other algorithms. Because the computational cost of SSHCS is dominated by VCA algorithm, which has the order of complexity O(p^2N). What's more, SSHCS algorithm does not require iteration. The runtime of SSCR_SU and SpeCA is basically in the same magnitude, but the proposed algorithm is slightly longer than the two algorithms. For example, the runtime of HCS_MMM is about five times that of SSCR_SU. This is because HCS_MMM continuously adjusts the step size μ during iteration to balance the primal and dual residuals, which makes the algorithm more difficult to converge than other methods. So 100 iterations are necessary for HCS_MMM, whereas SSCR_SU typically takes only 20 to achieve good performance. The complexity of the SpeCA algorithm is dominated by the term p^3L^3, which indicates that the runtime of the algorithm will be sensitive to the number of endmembers. When p is large, the computational cost is disastrous. For example, SpeCA runs even longer than HCS_MMM on Salinas data. It should be noted that, the computational complexity of proposed algorithm is O(vLN). Therefore, the runtime is lengthened by the increase of v, when the sampling rate is increased. But the other algorithms are almost constant.

V. Conclusion

In this article, due to the limitation of LMM in HCS reconstruction, we proposed a more accurate hyperspectral inages model, called the MMM, which considers not only the linear mixing properties but also some other spectral conditions, such as spectral variability, nonlinear effects, and sparse residuals. Based on the proposed MMM, the corresponding HCS algorithm is also presented, which only performs spectral compressed sampling on HSIs at the sampling stage and HCS reconstruction based on the proposed MMM at the reconstruction stage. As for the reconstruction process, ADMM is utilized to estimate each component of the MMM, furthermore, reasonable prior knowledge is introduced to improve the estimation accuracy. With the proposed HCS_MMM, we successfully realized the reconstruction of the original HSIs with high quality from the spectral compressed sampling data. Experimental results show that the proposed HCS_MMM can obtain better reconstruction performance than the other state-of-the-art algorithms, which means that by taking multiple factors beyond linear effects into consideration is generally superior to the LMM. Note that how to improve the reconstruction performance at low sampling rate for the proposed HCS_MMM is the future work that needs to be solved.

Appendix
Solution to ALMM-Based CS Reconstrction

The object function in (4) is not convex with respect to all variables simultaneously, but it is a convex problem regarding the separate variable when other variables are fixed. As a result, we successively minimize the object function with respect to $\mathbf{S}$, $\mathbf{H}$, $\mathbf{G}$, $\mathbf{\Gamma}$, $\mathbf{R}$ as follows:

1) *Subproblem for* $\mathbf{S}, \mathbf{H}, \mathbf{G}$: After fixing other variables, since $p \ll N$, the estimation of abundance $\mathbf{S}$ in (4) is found to be a standard least squares problem

$$\min_{\mathbf{S}^k} \frac{1}{2} \left\| \mathbf{Y} - \mathbf{A}\mathbf{E}\mathbf{S}^k - \mathbf{A}\mathbf{\Gamma}^k\mathbf{E}\mathbf{H}^k - \mathbf{A}\mathbf{M}\mathbf{G}^k - \mathbf{A}\mathbf{R}^k \right\|_F^2 \quad (7)$$

where the superscript k indicates the $k-\mathrm{th}$ iteration.
The least squares problem has analytical solutions with high accuracy and high reliability

$$\mathbf{S}^{k+1} \leftarrow \left[(\mathbf{A}\mathbf{E})^T (\mathbf{A}\mathbf{E}) \right]^{-1} (\mathbf{A}\mathbf{E})^T \times \left(\mathbf{Y} - \mathbf{A}\mathbf{\Gamma}^k\mathbf{E}\mathbf{H}^k - \mathbf{A}\mathbf{M}\mathbf{G}^k - \mathbf{A}\mathbf{R}^k \right) \quad (8)$$

where $\mathbf{C}^{-1}$ denotes the inverse matrix of $\mathbf{C}$.
Similarly, the minimization of $\mathbf{H}^k$ and $\mathbf{G}^k$ is the same as that of $\mathbf{S}^k$ and is also a least squares problem. Therefore, an analytical solution is given directly

$$\mathbf{H}^{k+1} \leftarrow \left[\left(\mathbf{A}\mathbf{\Gamma}^k\mathbf{E}\right)^T \mathbf{A}\mathbf{\Gamma}^k\mathbf{E} \right]^{-1} \left(\mathbf{A}\mathbf{\Gamma}^k\mathbf{E}\right)^T \times \left(\mathbf{Y} - \mathbf{A}\mathbf{E}\mathbf{S}^{k+1} - \mathbf{A}\mathbf{M}\mathbf{G}^k - \mathbf{A}\mathbf{R}^k \right) \quad (9)$$

$$\mathbf{G}^{k+1} \leftarrow \left[(\mathbf{A}\mathbf{M})^T \mathbf{A}\mathbf{M} \right]^{-1} (\mathbf{A}\mathbf{M})^T \times \left(\mathbf{Y} - \mathbf{A}\mathbf{E}\mathbf{S}^{k+1} - \mathbf{A}\mathbf{\Gamma}^k\mathbf{E}\mathbf{H}^{k+1} - \mathbf{A}\mathbf{R}^k \right). \quad (10)$$

2) *Subproblem for* $\mathbf{\Gamma}$: Fixing other variables, the subproblem of minimizing $\mathbf{\Gamma}$ can be written as

$$\min_{\mathbf{\Gamma}^k} \frac{1}{2}\left\|\mathbf{Y}-\mathbf{AES}^{k+1}-\mathbf{A\Gamma}^k\mathbf{EH}^{k+1}-\mathbf{AMG}^{k+1}-\mathbf{AR}^k\right\|_F^2 + \lambda_1\left\|\mathbf{F\Gamma}^k\right\|_{1,1}. \tag{11}$$

Introducing auxiliary variables $\mathbf{T}_1^k=\mathbf{A\Gamma}^k$, $\mathbf{T}_2^k=\mathbf{\Gamma}^k$, and $\mathbf{T}_3^k=\mathbf{FT}_2^k$, the unconstrained optimization problem of (11) can be converted into the following constrained optimization problem:

$$\min_{\mathbf{\Gamma}^k,\mathbf{T}_1^k,\mathbf{T}_2^k,\mathbf{T}_3^k} \frac{1}{2}\left\|\mathbf{Y}-\mathbf{AES}^{k+1}-\mathbf{T}_1^k\mathbf{EH}^{k+1}-\mathbf{AMG}^{k+1} - \mathbf{AR}^k\right\|_F^2 + \lambda_1\left\|\mathbf{T}_3^k\right\|_{1,1}$$
$$\text{s.t. } \mathbf{A\Gamma}^k-\mathbf{T}_1^k=\mathbf{0},\mathbf{\Gamma}^k-\mathbf{T}_2^k=\mathbf{0},\mathbf{FT}_2^k-\mathbf{T}_3^k=\mathbf{0}. \tag{12}$$

The augmented Lagrange function for, $\mathbf{\Gamma}^k$,$\mathbf{T}_1^k$,$\mathbf{T}_2^k$,$\mathbf{T}_3^k$ is

$$\begin{aligned}&\mathcal{L}\left(\mathbf{\Gamma}^k,\mathbf{T}_1^k,\mathbf{T}_2^k,\mathbf{T}_3^k,\mathbf{D}_1^k,\mathbf{D}_2^k,\mathbf{D}_3^k\right)\\ &=\frac{1}{2}\left\|\mathbf{Y}-\mathbf{AES}^{k+1}-\mathbf{T}_1^k\mathbf{EH}^{k+1}-\mathbf{AMG}^{k+1}-\mathbf{AR}^k\right\|_F^2\\ &+\lambda_1\left\|\mathbf{T}_3^k\right\|_{1,1}+\frac{\mu}{2}\left\|\mathbf{A\Gamma}^k-\mathbf{T}_1^k-\mathbf{D}_1^k\right\|_F^2\\ &+\frac{\mu}{2}\left\|\mathbf{\Gamma}^k-\mathbf{T}_2^k-\mathbf{D}_2^k\right\|_F^2+\frac{\mu}{2}\left\|\mathbf{FT}_2^k-\mathbf{T}_3^k-\mathbf{D}_3^k\right\|_F^2\end{aligned} \tag{13}$$

where $\mathbf{D}_1^k$, $\mathbf{D}_2^k$, and $\mathbf{D}_3^k$ denotes the Lagrange multipliers and μ is a positive penalty constant used to control the iterative convergence speed. How to choose a suitable value for the parameter μ is an active research topic. Adaptive selection of parameters based on the primal and dual ADMM variables, where the updated objective of μ is to keep the ratio between the primal and dual residual norms within a given positive interval, usually performs very well in many cases [39], [41], [42]. In this article, we adopt the parameter selection rules suggested by [42]. The augmented Lagrange function of (13) is also a nonconvex optimization problem for minimizing all variables at the same time. Therefore, the ADMM is applied again to update each variable alternately.

Minimization with respect to $\mathbf{\Gamma}^k$**:** The optimization problem of $\mathbf{\Gamma}^k$ can be formulated as follows:

$$\min_{\mathbf{\Gamma}^k}\frac{\mu}{2}\left\|\mathbf{A\Gamma}^k-\mathbf{T}_1^k-\mathbf{D}_1^k\right\|_F^2+\frac{\mu}{2}\left\|\mathbf{\Gamma}^k-\mathbf{T}_2^k-\mathbf{D}_2^k\right\|_F^2. \tag{14}$$

Therefore, the analytical solution for $\mathbf{\Gamma}^{k+1}$ can be simply obtained by the following formula:

$$\mathbf{\Gamma}^{k+1}\leftarrow\left(\mathbf{A}^T\mathbf{A}+\mathbf{I}_L\right)^{-1}\left[\mathbf{A}^T\left(\mathbf{T}_1^k+\mathbf{D}_1^k\right)+\left(\mathbf{T}_2^k+\mathbf{D}_2^k\right)\right] \tag{15}$$

where $\mathbf{I}_L$ denotes the identity matrix with $L\times L$.

Minimization with respect to $\mathbf{T}_1^k$: The subproblem of $\mathbf{T}_1^k$ can be written as

$$\min_{\mathbf{T}_1^k}\frac{1}{2}\left\|\mathbf{Y}-\mathbf{AES}^{k+1}-\mathbf{T}_1^k\mathbf{EH}^{k+1}-\mathbf{AMG}^{k+1}-\mathbf{AR}^k\right\|_F^2 + \frac{\mu}{2}\left\|\mathbf{A\Gamma}^{k+1}-\mathbf{T}_1^k-\mathbf{D}_1^k\right\|_F^2 \tag{16}$$

which is readily solved by

$$\begin{aligned}\mathbf{T}_1^{k+1}\leftarrow&\left[\left(\mathbf{Y}-\mathbf{AES}^{k+1}-\mathbf{AMG}^{k+1}-\mathbf{AR}^k\right)\right.\\ &\times\left(\mathbf{EH}^{k+1}\right)^T+\mu\left(\mathbf{A\Gamma}^{k+1}-\mathbf{D}_1^k\right)\Big]\\ &\times\left[\left(\mathbf{EH}^{k+1}\right)\left(\mathbf{EH}^{k+1}\right)^T+\mu\mathbf{I}_L\right]^{-1}.\end{aligned} \tag{17}$$

Minimization with respect to $\mathbf{T}_2^k$: To compute $\mathbf{T}_2^{k+1}$, the optimization problem to be solved is

$$\min_{\mathbf{T}_2^k}\frac{\mu}{2}\left\|\mathbf{\Gamma}^{k+1}-\mathbf{T}_2^k-\mathbf{D}_2^k\right\|_F^2+\frac{\mu}{2}\left\|\mathbf{FT}_2^k-\mathbf{T}_3^k-\mathbf{D}_3^k\right\|_F^2. \tag{18}$$

Similar to $\mathbf{\Gamma}^{k+1}$, $\mathbf{T}_2^{k+1}$ is updated as follows:

$$\mathbf{T}_2^{k+1}\leftarrow\left(\mathbf{F}^T\mathbf{F}+\mathbf{I}_L\right)^{-1}\left[\mathbf{F}^T\left(\mathbf{T}_3^k+\mathbf{D}_3^k\right)+\left(\mathbf{\Gamma}^{k+1}-\mathbf{D}_2^k\right)\right]. \tag{19}$$

Minimization with respect to $\mathbf{T}_3^k$: To compute $\mathbf{T}_3^{k+1}$, we solve the optimization problem

$$\min_{\mathbf{T}_3^k}\lambda_1\left\|\mathbf{T}_3^k\right\|_{1,1}+\frac{\mu}{2}\left\|\mathbf{FT}_2^{k+1}-\mathbf{T}_3^k-\mathbf{D}_3^k\right\|_F^2 \tag{20}$$

whose solution is the well-known *soft threshold*

$$\mathbf{T}_3^{k+1}\leftarrow\text{soft}\left(\mathbf{FT}_2^{k+1}-\mathbf{D}_3^k,\frac{\lambda_1}{\mu}\right) \tag{21}$$

where $\text{soft}(*)$ denotes the soft-threshold function.

Lagrange multipliers update $\mathbf{D}_1^{k+1}$, $\mathbf{D}_2^{k+1}$, $\mathbf{D}_3^{k+1}$: Lagrange multipliers can be updated by means of gradient descent

$$\begin{aligned}\mathbf{D}_1^{k+1}&\leftarrow\mathbf{D}_1^k-\left(\mathbf{A\Gamma}^{k+1}-\mathbf{T}_1^{k+1}\right)\\ \mathbf{D}_2^{k+1}&\leftarrow\mathbf{D}_2^k-\left(\mathbf{\Gamma}^{k+1}-\mathbf{T}_2^{k+1}\right)\\ \mathbf{D}_3^{k+1}&\leftarrow\mathbf{D}_3^k-\left(\mathbf{FT}_2^{k+1}-\mathbf{T}_3^{k+1}\right).\end{aligned} \tag{22}$$

3) *Subproblem for* $\mathbf{R}$: With the other variables fixed, the minimum subproblem with respect to $\mathbf{R}^k$ can be expressed as follows:

$$\min_{\mathbf{R}^k}\frac{1}{2}\left\|\mathbf{Y}-\mathbf{AES}^{k+1}-\mathbf{A\Gamma}^{k+1}\mathbf{EH}^{k+1}-\mathbf{AMG}^{k+1} -\mathbf{AR}^k\right\|_F^2+\lambda_2\left\|\mathbf{R}^k\right\|_{2,1}. \tag{23}$$

After introducing the auxiliary variables $\mathbf{U}^k=\mathbf{R}^k$, the unconstrained optimization problem of (23) can be converted into the following constrained optimization problem:

$$\min_{\mathbf{R}^k,\mathbf{U}^k}\frac{1}{2}\left\|\mathbf{Y}-\mathbf{AES}^{k+1}-\mathbf{A\Gamma}^{k+1}\mathbf{EH}^{k+1}-\mathbf{AMG}^{k+1} -\mathbf{AR}^k\right\|_F^2+\lambda_2\left\|\mathbf{U}^k\right\|_{2,1}\text{s.t.}\mathbf{R}^k-\mathbf{U}^k=\mathbf{0}. \tag{24}$$

The augmented Lagrange function is

$$\begin{aligned}\mathcal{L}\left(\mathbf{R}^k,\mathbf{U}^k\right)=&\frac{1}{2}\left\|\mathbf{Y}-\mathbf{AES}^{k+1}-\mathbf{A\Gamma}^{k+1}\mathbf{EH}^{k+1}\right.\\ &\left.-\mathbf{AMG}^{k+1}-\mathbf{AR}^k\right\|_F^2\\ &+\lambda_2\left\|\mathbf{U}^k\right\|_{2,1}+\frac{\mu}{2}\left\|\mathbf{R}^k-\mathbf{U}^k-\mathbf{V}^k\right\|_F^2\end{aligned} \tag{25}$$

where $\mathbf{V}^k$ is a Lagrange multiplier.

Algorithm 1: Pseudocode of the HCS_MMM Algorithm.

Input: $\mathbf{Y}$, $\mathbf{A}$, and $\mathbf{E}$
Set parameters: λ_1, λ_2, μ, maxiters $= 100$
Step 1. **Initialization**:
$\mathbf{H}^0 = \mathbf{0}, \mathbf{G}^0 = \mathbf{0}, \mathbf{\Gamma}^0 = \mathbf{I}_L, \mathbf{T}_1^0 = \mathbf{A}\mathbf{\Gamma}^0,$
$\mathbf{T}_2^0 = \mathbf{\Gamma}^0, \mathbf{T}_3^0 = \mathbf{F}\mathbf{T}_2^0, \mathbf{R}^0 = \mathbf{0}, \mathbf{U}^0 = \mathbf{R}^0$,
$\mathbf{D}_1^0 = \mathbf{0}, \mathbf{D}_2^0 = \mathbf{0}, \mathbf{D}_3^0 = \mathbf{0}, \mathbf{V}^0 = \mathbf{0}, k = 1$, thr $= 10^{-5}$, res $= \infty$
Step 2. While $k <$ maxiters and (resp $>$ throrresd $>$ thr)
Step 3. Update $\mathbf{S}^{k+1}$ according to (8);
Step 4. Update $\mathbf{H}^{k+1}$ according to (9);
Step 5. Update $\mathbf{G}^{k+1}$ according to (10);
Step 6. Update $\mathbf{\Gamma}^{k+1}$ according to (15);
Step 7. Update $\mathbf{T}_1^{k+1}$ according to (17);
Step 8. Update $\mathbf{T}_2^{k+1}$ according to (19);
Step 9. Update $\mathbf{T}_3^{k+1}$ according to (21);
Step 10. Update Lagrange multipliers $\mathbf{D}_1^{k+1}$,$\mathbf{D}_2^{k+1}$,$\mathbf{D}_3^{k+1}$ by (22);
Step 11. Update $\mathbf{R}^{k+1}$ according to (27);
Step 12. Update $\mathbf{U}^{k+1}$ according to (29);
Step 13. Update the Lagrange multipliers $\mathbf{V}^{k+1}$ by (30);
Step 14. Compute the convergence conditions:

$$resp = \|\mathbf{A}\mathbf{\Gamma}^{k+1} - \mathbf{T}_1^{k+1}\|_F + \|\mathbf{\Gamma}^{k+1} - \mathbf{T}_2^{k+1}\|_F + \|\mathbf{F}\mathbf{T}_2^{k+1} - \mathbf{T}_3^{k+1}\|_F + \|\mathbf{R}^{k+1} - \mathbf{U}^{k+1}\|_F$$

$$resd = \|\mu\mathbf{A}^T\mathbf{T}_1^{k+1} - \mathbf{T}_1^{k}\|_F + \|\mu\mathbf{T}_2^{k+1} - \mathbf{T}_2^{k}\|_F + \|\mu\mathbf{F}\mathbf{T}_3^{k+1} - \mathbf{T}_3^{k}\|_F + \|\mu\mathbf{U}^{k+1} - \mathbf{U}^{k}\|_F$$

Step 15. Update μ:

$$\mu^{k+1} = \begin{cases} 2\mu^k & \text{if resp} > \text{resd} \\ \mu^k/2 & \text{if resd} > \text{resp} \\ \mu^k & \text{otherwise} \end{cases}$$

End While
Step 16. Estimate $\hat{\mathbf{X}}$ by (31):
Output: $\hat{\mathbf{X}}$

The augmented Lagrange function $\mathcal{L}(\mathbf{R}^k, \mathbf{U}^k)$ minimizes $\mathbf{R}^k$ to obtain the following optimization subproblems and $\mathbf{R}^k$ can be computed by solving the following optimization problem:

$$\min_{\mathbf{R}^k} \frac{1}{2}\left\|\mathbf{Y} - \mathbf{A}\mathbf{E}\mathbf{S}^{k+1} - \mathbf{A}\mathbf{\Gamma}^{k+1}\mathbf{E}\mathbf{H}^{k+1} - \mathbf{A}\mathbf{M}\mathbf{G}^{k+1} - \mathbf{A}\mathbf{R}^k\right\|_F^2 + \frac{\mu}{2}\left\|\mathbf{R}^k - \mathbf{U}^k - \mathbf{V}^k\right\|_F^2. \tag{26}$$

$\mathbf{R}^{k+1}$ can be updated as follows:

$$\mathbf{R}^{k+1} \leftarrow \left(\mathbf{A}^T\mathbf{A} + \mathbf{I}_L\right)^{-1} \times \left[\mathbf{A}^T\left(\mathbf{Y} - \mathbf{A}\mathbf{E}\mathbf{S}^{k+1} - \mathbf{A}\mathbf{\Gamma}^{k+1}\mathbf{E}\mathbf{H}^{k+1} - \mathbf{A}\mathbf{M}\mathbf{G}^{k+1}\right) + \mu\left(\mathbf{U}^k + \mathbf{V}^k\right)\right]. \tag{27}$$

We minimize $\mathcal{L}(\mathbf{R}^k, \mathbf{U}^k)$ with respect to $\mathbf{U}^k$:

$$\min_{\mathbf{U}^k} \lambda_2\left\|\mathbf{U}^k\right\|_{2,1} + \frac{\mu}{2}\left\|\mathbf{R}^{k+1} - \mathbf{U}^k - \mathbf{V}^k\right\|_F^2 \tag{28}$$

whose solution is the well-known *vect-soft threshold* [43], [44]

$$\mathbf{U}^{k+1} \leftarrow \text{vect} - \text{soft}\left(\mathbf{\Phi}, \frac{\lambda_2}{\mu}\right) \tag{29}$$

where $\text{vect} - \text{soft}(*)$ denotes the vect-soft-threshold function, $\mathbf{\Phi} = \mathbf{R}^{k+1} - \mathbf{V}^k$, and $\text{vect} - \text{soft}(\mathbf{\Phi}, \frac{\lambda_2}{\mu}) = \frac{\max\{\|\mathbf{\Phi}\|_2 - \frac{\lambda_2}{\mu}, 0\}}{\max\{\|\mathbf{\Phi}\|_2 - \frac{\lambda_2}{\mu}, 0\} + \frac{\lambda_2}{\mu}} \times \mathbf{\Phi}$.

Lagrange multiplier update $\mathbf{V}^{k+1}$

$$\mathbf{V}^{k+1} \leftarrow \mathbf{V}^k - \left(\mathbf{R}^{k+1} - \mathbf{U}^{k+1}\right). \tag{30}$$

After k iterations, the original HSIs can be estimated as follows:

$$\hat{\mathbf{X}} = \mathbf{E}\mathbf{S}^{k+1} + \mathbf{\Gamma}^{k+1}\mathbf{E}\mathbf{H}^{k+1} + \mathbf{M}\mathbf{G}^{k+1} + \mathbf{R}^{k+1}. \tag{31}$$

In this article, the HCS reconstruction algorithm based on the MMM is marked as HCS_MMM. The specific calculation process is as follows.

In Algorithm 1, the computational cost of the proposed HCS_MMM is dominated by matrix products. For each iteration, the most costly steps are the calculus of $\mathbf{A}\mathbf{R}^k$ and $\mathbf{A}^T\mathbf{Y}$, which have the order of complexity O(vLN). The complexities of the step 3, 4, 5, 7, and 11 are dominated by the term vLNassociated with the solution calculus of $\mathbf{A}\mathbf{R}^k$ or $\mathbf{A}^T\mathbf{Y}$. Therefore, the overall order of complexity per iteration is given then by O(vLN).

ACKNOWLEDGMENT

The authors would like to thank those anonymous reviewers for their insightful comments in improving the quality of this article.

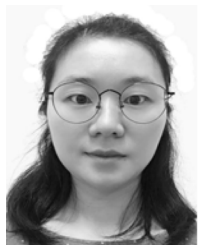
REFERENCES

[1] 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.*, vol. 5, no. 2, pp. 354–379, May 2012.
[2] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
[3] Q. Sun, "Recovery of sparsest signals via q-minimization," *Appl. Comput. Harmonic Anal.*, vol. 32, no. 3, pp. 329–341, May 2012, doi: 10.1016/j.acha.2011.07.001.
[4] J. Yang and Y. Zhang, "Alternating direction algorithms for L1-problems in compressive sensing," *SIAM J. Sci. Comput.*, vol. 33, no. 1, pp. 250–278, Jun. 2011, doi: 10.1137/090777761.
[5] Z. Xu, H. Zhang, Y. Wand, X. Chang, and Y. Liang, "$L_{1/2}$ regularization," *Sci. China Info. Sci.*, vol. 53, no. 06, pp. 1159–1169, May 2010, doi: 10.1007/s11432-010-0090-0.
[6] S. Mun and J. E. Fowler, "Block compressed sensing of images using directional transforms," in *Proc. 16th IEEE Int. Conf. Image Process.*, Nov. 7–10, 2009, pp. 3021–3024.
[7] S. Xianbiao and N. Ahuja, "Imaging via three-dimensional compressive sampling (3DCS)," in *Proc. IEEE Int. Conf. Comput. Vis.*, Nov. 6–13, 2011, pp. 439–446.
[8] Y. B. Jia, Y. Feng, and Z. L. Wang, "Reconstructing hyperspectral images from compressive sensors via exploiting multiple priors," *Spectrosc. Lett.*, vol. 48, no. 1, pp. 22–26, Aug. 2015, doi: 10.1080/00387010.2013.850727.
[9] Y. Wang, L. Lin, Q. Zhao, T. Yue, D. Meng, and Y. Leung, "Compressive sensing of hyperspectral images via joint tensor tucker decomposition and weighted total variation regularization," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 12, pp. 2457–2461, Nov. 2017, doi: 10.1109/lgrs.2017.2771212.
[10] J. Xue, Y. Zhao, W. Liao, and J. Chan, "Nonlocal tensor sparse representation and low-rank regularization for hyperspectral image compressive sensing reconstruction," *Remote Sens.*, vol. 11, no. 2, pp. 1–24, Aug. 2019, doi: 10.3390/rs11020193.

[11] L. Zhang, W. Wei, C. Tian, F. Li, and Y. Zhang, "Exploring structured sparsity by a reweighted Laplace prior for hyperspectral compressive sensing," *IEEE Trans. Image Process.*, vol. 25, no. 10, pp. 4974–4988, Aug. 2016.
[12] J. E. Fowler, "Compressive-projection principal component analysis," *IEEE Trans. Image Process.*, vol. 18, no. 10, pp. 2230–2242, Jun. 2009.
[13] N. H. Ly, Q. Du, and J. E. Fowler, "Reconstruction from random projections of hyperspectral imagery with spectral and spatial partitioning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 2, pp. 466–472, Oct. 2013.
[14] C. Chen, L. Wei, E. W. Tramel, and J. E. Fowler, "Reconstruction of hyperspectral imagery from random projections using multihypothesis prediction," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 365–374, Feb. 2014.
[15] G. Martin, J. M. B. Dias, and A. J. Plaza, "A new technique for hyperspectral compressive sensing using spectral unmixing," *Proc. SPI*, vol. 8514, Oct. 19, 2012, Art. no. 85140N.
[16] G. Martin, J. M. Bioucas-Dias, and A. Plaza, "Hyperspectral coded aperture (HYCA): A new technique for hyperspectral compressive sensing," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 22–27, 2012, pp. 1381–1384.
[17] M. Golbabaee, S. Arberet, and P. Vandergheynst, "Compressive source separation: Theory and methods for hyperspectral imaging," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 5096–5110, Sep. 2013.
[18] G. Martín, J. M. Bioucas-Dias, and A. Plaza, "HYCA: A new technique for hyperspectral compressive sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2819–2831, Nov. 2015.
[19] G. Martín and J. M. Bioucas-Dias, "Hyperspectral blind reconstruction from random spectral projections," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 6, pp. 2390–2399, Apr. 2016.
[20] L. Zhang, W. Wei, Y. Zhang, H. Yan, F. Li, and C. Tian, "Locally similar sparsity-based hyperspectral compressive sensing using unmixing," *IEEE Trans. Comput. Imag.*, vol. 2, no. 2, pp. 86–100, Mar. 2016.
[21] G. Martín and J. M. Bioucas-Dias, "Spatial-spectral hyperspectral image compressive sensing," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 23–28, 2017, pp. 3988–3991.
[22] Z. Wang, Y. Feng, and Y. Jia, "Spatio-spectral hybrid compressive sensing of hyperspectral imagery," *Remote Sens. Lett.*, vol. 6, no. 3, pp. 199–208, Mar. 2015, doi: 10.1080/2150704x.2015.1024892.
[23] L. Wang, Y. Feng, Y. Gao, Z. Wang, and M. He, "Compressed sensing reconstruction of hyperspectral images based on spectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 4, pp. 1266–1284, Jan. 2018.
[24] Z. Wang, M. He, Z. Ye, Y. Nian, L. Qiao, and M. Chen, "Exploring error of linear mixed model for hyperspectral image reconstruction from spectral compressive sensing," *J. Appl. Remote Sens.*, vol. 13, no. 3, pp. 1–18, Sep. 2019, doi: 10.1117/1.jrs.13.036514.
[25] J. M. P. Nascimento and J. M. B. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, Apr. 2005.
[26] P. Thouvenin, N. Dobigeon, and J. Tourneret, "Hyperspectral unmixing with spectral variability using a perturbed linear mixing model," *IEEE Trans. Signal Process.*, vol. 64, no. 2, pp. 525–538, Oct. 2016.
[27] L. Drumetz, M. Veganzones, S. Henrot, R. Phlypo, J. Chanussot, and C. Jutten, "Blind hyperspectral unmixing using an extended linear mixing model to address spectral variability," *IEEE Trans. Image Process.*, vol. 25, no. 8, pp. 3890–3905, Jun. 2016.
[28] D. Hong, N. Yokoya, J. Chanussot, and X. X. Zhu, "An augmented linear mixing model to address spectral variability for hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 1923–1938, Nov. 2019.
[29] Z. Wang, M. He, Z. Ye, and Y. Nian, "For hyperspectral compressed sensing method on linear mixed spectrum model," *J. Remote Sens.*, vol. 24, no. 3, pp. 63–75, 2020, doi: 10.11834/jrs.20209211.
[30] N. Dobigeon, J.-Y. Tourneret, C. Richard, J. Bermudez, S. McLaughlin, and A. Hero, "Nonlinear unmixing of hyperspectral images: Models and algorithms," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 82–94, 2014.
[31] A. Halimi, Y. Altmann, N. Dobigeon, and J. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4153–4162, Nov. 2011.
[32] J. Eckstein and D. P. Bertsekas, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Math. Program.*, vol. 55, no. 1, pp. 293–318, Apr. 1992, doi: 10.1007/bf01581204.
[33] B. Somers, G. P. Asner, L. Tits, and P. Coppin, "Endmember variability in spectral mixture analysis: A review," *Remote Sens. Environ.*, vol. 115, no. 7, pp. 1603–1616, 2011, doi: 10.1016/j.rse.2011.03.003.
[34] A. Zare and K. Ho, "Endmember variability in hyperspectral analysis: Addressing spectral variability during spectral unmixing," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 95–104, Jan. 2014.
[35] A. Halimi, P. Honeine, and J. M. Bioucas-Dias. "Hyperspectral unmixing in presence of endmember variability, nonlinearity or mismodeling effects," *IEEE Trans. Image Process.*, vol. 25, no. 10, pp. 4565–4579, Oct. 2016.
[36] R. N. Clark *et al.*, "*USGS Digital Spectral Library splib06a*." [Online]. Available: http://speclab.cr.usgs.gov/spectral.lib06
[37] J. E. Fowler, "Compressive pushbroom and whiskbroom sensing for hyperspectral remote-sensing imaging," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 27–30, 2014, pp. 684–688.
[38] J. Bacca, C. V. Correa, and H. Arguello, "Noniterative hyperspectral image reconstruction from compressive fused measurements," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 4, pp. 1231–1239, Apr. 2019.
[39] M. D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Total variation spatial regularization for sparse hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4484–4502, May 2012.
[40] L. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. D, Nonlinear Phenom.*, vol. 60, no. 1, pp. 259–268, Nov. 1992, doi: 10.1016/0167-2789(92)90242-F.
[41] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jul. 2010, doi: 10.1561/2200000016.
[42] Y. Lin, B. Wohlberg, and V. Vesselinov, "ADMM Penalty Parameter Selection with Krylov Subspace Recycling Technique for Sparse Coding," in *Proceedings of IEEE Int. Conf. Image Process.*, Sep. 17–20, 2017, pp. 1945–1949.
[43] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2479–2493, Mar. 2009.
[44] M. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Collaborative sparse regression for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 341–354, Feb. 2014.
[45] P. Gamba, "A collection of data for urban area characterization," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Sep. 20–24. 2004, pp. 1–72.

Zhongliang Wang received the Ph.D. degree in signal and information processing from the School of Electronics and Information, Northwestern Polytechnical University, Xi'an, China, in 2015.

He is currently an Associate Professor with the Department of Electric Engineering, Tongling University, Tongling, China. His research interests include image enhancement and hyperspectral remote sensing.



Mi He was born in Sichuan province, China, in 1982. She received the B.S. degree in applied mathematics, the M.S. degree in computational mathematics, and the Ph.D. degree in information and communication engineering from the National University of Defense and Technology, Changsha, China, in 2005, 2007, and 2012, respectively.

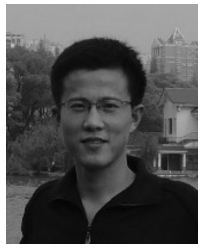
From 2009 to 2010, she was a Visiting Ph.D. Student with Delft University of Technology, Delft, The Netherlands. Since 2015, she has been an Associate Professor with the School of Biomedical Engineering and Imaging Medicine, Army Medical University, Chongping, China. Her research interests include remote sensing and biomedical signal processing.

Zhen Ye received the Ph.D. degree in information and communication engineering from Northwestern Polytechnical University, Xi'an, China, in 2015.

She spent one year as an Exchange Student of Mississippi State University. She is currently an Associate Professor with the School of Electronics and Control Engineering, Chang'an University, Xi'an, China. Her research interests include hyperspectral image analysis, pattern recognition, and machine learning.

Ke Xu was born in Sichuan province, China, in 1982. He received the M.S. and the Ph.D. degrees in information and communication engineering from the National University of Defense and Technology, Changsha, China, in 2006 and 2011, respectively.

From 2008 to 2009, he was a Visiting Ph.D. Student with the University of Windsor, Windsor, ON, Canada. He is currently an Associate Professor with the College of Electronic Science, National University of Defense and Technology. His current research interests include intelligent information processing, underwater acoustic signal processing, and array signal processing.



Yongjian Nian was born in Shandong province, China, in 1982. He received the Ph.D. degree in information and communication engineering from the National University of Defense and Technology, Changsha, China, in 2010.

He is currently an Associate Professor with the School of Biomedical Engineering and Imaging Medicine, Army Medical University, Chongping, China. His research interests include hyperspectral images processing, pattern recognition, and machine learning.

Bormin Huang received the M.S.E. degree in aerospace engineering from the University of Michigan, Ann Arbor, MI, USA, in 1992, and the Ph.D. degree in satellite remote sensing from the University of Wisconsin-Madison, Madison, WI, USA, in 1998.

He is currently a Professor with the School of Information Science and Technology, Southwest Jiaotong University, Chengdu, China. He served as the Director of the Intel Parallel Computing Center with the University of Wisconsin-Madison, a Research Professor with the University of Las Palmas de Gran Canaria, Las Palmas, Spain, and a Guest Professor with several universities in China. His research interests include remote sensing, satellite data compression, artificial intelligence, and high-performance parallel computing.

Dr. Huang is a NVIDIA CUDA Fellow and a SPIE Fellow. He serves as an Associate Editor for the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING.