

# Spatial Group Sparsity Regularized Nonnegative Matrix Factorization for Hyperspectral Unmixing

Xinyu Wang, Yanfei Zhong, *Senior Member, IEEE*, Liangpei Zhang, *Senior Member, IEEE*, and Yanyan Xu

**Abstract**—In recent years, blind source separation (BSS) has received much attention in the hyperspectral unmixing field due to the fact that it allows the simultaneous estimation of both endmembers and fractional abundances. Although great performances can be obtained by the BSS-based unmixing methods, the decomposition results are still unstable and sensitive to noise. Motivated by the first law of geography, some recent studies have revealed that spatial information can lead to an improvement in the decomposition stability. In this paper, the group-structured prior information of hyperspectral images is incorporated into the nonnegative matrix factorization optimization, where the data are organized into spatial groups. Pixels within a local spatial group are expected to share the same sparse structure in the low-rank matrix (abundance). To fully exploit the group structure, image segmentation is introduced to generate the spatial groups. Instead of a predefined group with a regular shape (e.g., a cross or a square window), the spatial groups are adaptively represented by superpixels. Moreover, the spatial group structure and sparsity of the abundance are integrated as a modified mixed-norm regularization to exploit the shared sparse pattern, and to avoid the loss of spatial details within a spatial group. The experimental results obtained with both simulated and real hyperspectral data confirm the high efficiency and precision of the proposed algorithm.

**Index Terms**—Hyperspectral unmixing (HU), nonnegative matrix factorization (NMF), spatial group sparsity.

## I. INTRODUCTION

HYPERSPECTRAL sensors collect the reflection and radiation information of a scene, creating tens to hundreds of narrow bands that correspond to different narrow wavelengths. Recent advances in remote sensing and spectral imaging have led the way for the development of hyperspectral sensors. Hyperspectral remote sensing has become a rapidly developing technology, and plays an important role in fields such as agriculture, mineral exploration, and environmental monitoring [1]. Hyperspectral images (HSIs) provide abundant spectral information, and these data are often used to precisely identify the subtle differences between materials that are

Manuscript received December 7, 2016; revised May 21, 2017; accepted July 3, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 41622107 and Grant 41371344, and in part by the Natural Science Foundation of Hubei Province under Grant 2016CFA029. (*Corresponding author: Yanfei Zhong*)

The authors are with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China, and also with the Collaborative Innovation Center of Geospatial Technology, Wuhan University, Wuhan 430079, China (e-mail: wangxinyu@whu.edu.cn; zhongyanfei@whu.edu.cn; zlp62@whu.edu.cn; xuyy@whu.edu.cn).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TGRS.2017.2724944

present in a scene. Nevertheless, due to the spatial resolution limitation of hyperspectral imaging systems, each pixel in an HSI may be a mixture of several different materials [2].

The wide existence of “mixed” pixels is a major challenge for the hyperspectral data analysis. The precise analysis of mixed pixels at the subpixel level can be achieved by the spectral unmixing technique, through decomposing each mixed pixel into a collection of distinct or pure spectral signatures, called “endmembers,” and their corresponding percentages, called “fractional abundances.” The majority of the existing hyperspectral unmixing (HU) methods are based on the linear mixing model (LMM), which is generally recognized to be a straightforward and efficient way to simplify the mixing process [3]. In addition, two physical abundance constraints—the abundance nonnegative constraint (ANC) and the abundance sum-to-one constraint (ASC)—also need to be considered in linear unmixing.

In previous studies, the LMM-based HU problem has been divided into two subproblems—endmember extraction (EE) and abundance inversion—which are then solved separately. Convex geometry-based methods can efficiently address the EE problem, based on the fact that hyperspectral data distribute in a simplex set after affine transformation, with the end-members located in the vertices. Therefore, EE is equivalent to searching for a set of pure pixels or generating a set of endmembers that can describe a simplex, corresponding to the existence and nonexistence of pure pixels, respectively. Many geometrical-based HU methods have been proposed in the last two decades, including the pixel purity index [4], N-FINDR [5], vertex component analysis (VCA) [6], and the simplex growing algorithm [7] for pure pixel cases, as well as the minimum-volume enclosing simplex [8] for cases without pure pixels. The estimation of the fractional abundances can then be regarded as an inversion problem and solved by least-squares-based approaches such as fully constrained least squares (FCLS) [9]. However, the inversion result depends to a large extent on the precision of the EE.

Statistical-based HU is a powerful alternative for reducing the accumulative errors in the two-step-based HU methods, where the HU is regarded as a blind source separation (BSS) problem, and the endmembers and abundances are estimated simultaneously. BSS deals with the separation of a set of source signals from a set of mixed signals, with very little prior information about the mixing process and the source signals. Independent component analysis (ICA) [10], [11] and nonnegative matrix factorization (NMF) [12] are two typical approaches for solving BSS problems, which have

recently been introduced into the hyperspectral data analysis to decompose mixed pixels. By assuming the mutual independence of the source signals (abundances), ICA can be used to address the unmixing issue. However, such an assumption is not appropriate for linear unmixing, and goes against the ASC, which assumes that the source signals are statistically dependent. The other approach NMF is one of the recently emerged BSS methods for multivariate data analysis, where the nonnegative data can be factorized into a product of two low-dimensional nonnegative matrices, in which one matrix is termed the “mixing matrix” and the other is termed the “source matrix.” In view of the nonnegative properties, NMF is ideal for linear unmixing. However, it should be noted that the standard NMF methods may not guarantee solution uniqueness, and the decomposition result usually depends on the initialization. To address such problems, a number of constrained NMF (CNMF) methods have been proposed for blind HU, through adding some specific constraints on the endmembers or abundances. In detail, the geometric constraint of endmembers has been introduced into the NMF-based unmixing. Typical examples include minimum-volume CNMF (MVCNMF) [13], where the simplex volume is taken into account, and iterated constrained endmembers (ICE) [14], where the difference between endmembers is controlled in the iterative process. Due to the sparse distribution of ground materials, sparse priors have also been used in various sparsity promoting NMF methods, such as sparsity promoting ICE (SPICE) [15], dictionary learning (DL) [16], sparse component analysis (BSUSCA) [17], and  $L_{1/2}$  sparsity CNMF ( $L_{1/2}$ NMF) [18].  $L_{1/2}$ NMF is a case in point, where the  $L_{1/2}$ -norm is used instead of the  $L_1$ -norm to obtain a sparser solution. The  $L_{1/2}$  regularization has been shown to provide a better unmixing performance, and can be found in other NMF-based HU methods, such as graph regularized  $L_{1/2}$ -NMF (GLNMF) [19] and structure constrained sparse NMF (CSNMF) [20].

It should be noted that, in practice, the endmember signatures may be known in advance in HU. A typical example is multiple endmember spectral mixture analysis [21], which estimates the abundances by searching for the endmembers from a spectral library, and allows each pixel to be represented by a unique collection of the endmembers from the spectral library. However, the HU accuracy depends on an appropriate spectral library which is designed to address the spectral variability [22], [23]. In contrast, the advantage of NMF-based HU lies in the automatic estimation of the endmembers and abundances, with little prior knowledge about the scene.

Due to the spatial autocorrelation of the ground surface, some recent studies have found that incorporating spatial information in spectral unmixing can lead to improvements in both EE and abundance estimation [24]. The automated morphological EE algorithm [25], which was designed as an iterative process with different sliding windows, is one of the famous EE approaches incorporating spatial information. The purest pixels and the most mixed pixels can be identified by dilation and erosion operations, respectively, and then a morphological eccentricity index is used to determine the final endmember candidates. However, this is a time-consuming

process when the maximum window size is large. In addition, some other interesting spatial and spectral EE approaches have been proposed, including spatial purity-based EE [26], which aims at searching for the purest spatial neighborhoods, instead of a single pixel, and spatial preprocessing [27], where the pure pixels are assumed to be more likely to be located in spatially homogenous areas. Incorporating a spatial prior in the abundance estimation is another way to promote the unmixing results. Usually, the abundance of a pixel is estimated independently of its neighborhood by minimizing the reconstruction error. To enforce spatial homogeneity of the abundances, Iordache *et al.* [28] incorporated a spatial constraint in sparse unmixing by means of a total variation regularizer, which promotes piecewise smooth transitions in the abundance of the same endmember among neighboring pixels. Liu *et al.* [29] proposed abundance separation and smoothness CNMF (ASSNMF) by introducing the abundance separation and abundance smoothness constraints into NMF to act on the spectral and spatial domains, respectively. It should be noted that the spatial neighborhood in all of the above methods is described as a set of pixels with a predefined regular shape (e.g., a cross or square window).

In this paper, a new blind unmixing method called the spatial group sparsity regularized NMF (SGSNMF) is proposed, by incorporating a constraint, namely, a spatial group sparsity regularizer, into the NMF-based unmixing process. The major objective of this paper is to explore the use of spatial information and the sparse structure of HSIs. Motivated by the first law of geography, the pixels within a local spatial group are expected to share the same sparse pattern in the low-rank matrix (abundance). However, to date, such spatial group structures of HSIs have not been fully explored in HU. In order to exploit the group information, instead of the predefined pixel grids, simple linear iterative clustering (SLIC) image segmentation [30] is introduced to the hyperspectral data to generate local spatial groups with irregular shapes (sets of pixels, also called superpixels). Superpixels are natural and adaptive representations of a scene, where each of the contiguous pixels is similar with respect to their spectral features. The spatial prior and the sparsity of the abundance are then integrated as a modified mixed-norm regularization into NMF to exploit the shared sparse pattern. Compared to the well-known sparsity-inducing regularizers, such as the  $L_1$ -norm and the  $L_{1/2}$ -norm, the proposed regularizer has been shown to be able to induce a sparser representation at the group level. This regularization is reasonable and effective, according to the fact that only a few different substances play a dominant role in a continuous region. In addition, since there is only one regularizer in SGSNMF, it is more effective than the other NMF-based approaches with two constraints, such as ASSNMF, GLNMF, and CSNMF. As a result of the integration of the spatial structure information and the sparse property, SGSNMF can obtain better unmixing results than the other popular algorithms.

The rest of this paper is organized as follows. In Section II, the linear unmixing problem and the NMF technique are briefly introduced. Section III presents the spatial prior and the SGSNMF scheme in detail. The performance of SGSNMF is

then presented and analyzed in Sections IV and V, through the use of both simulated and real hyperspectral data, respectively. Section VI draws the conclusion and discusses our future work.

## II. NONNEGATIVE MATRIX FACTORIZATION

In this section, we briefly introduce the LMM and its relationship with NMF.

### A. Linear Mixing Model

The proposed algorithm performs HU based on the LMM, which is generally recognized to be a straightforward and efficient way to represent mixed pixels. The LMM models the ground as being flat and assumes that the incident sunlight radiation reflects from the surface only through a single reflection, without multiple scattering. Each pixel is then modeled as a positive linear combination of all the radiated spectra of the materials making up the pixel. We let  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N] \in \mathbb{R}^{L \times N}$  denote the hyperspectral data matrix that consists of  $N$  pixels with  $L$  bands. For the  $j$ th mixed pixel  $\mathbf{x}_j = [x_{1j}, \dots, x_{Lj}]^T \in \mathbb{R}^L$ , the LMM provides a mathematical representation as follows:

$$\mathbf{x}_j = \sum_{i=1}^M s_{ij} \mathbf{a}_i + \boldsymbol{\epsilon}_j = \mathbf{A}\mathbf{s}_j + \boldsymbol{\epsilon}_j \quad (1)$$

where  $M$  is the number of endmember signatures in the scene,  $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_M] \in \mathbb{R}^{L \times M}$  represents the mixing matrix that consists of  $M$  distinct endmember spectral signatures,  $\mathbf{s}_j = [s_{1j}, s_{2j}, \dots, s_{Mj}]^T \in \mathbb{R}^M$  is the abundance vector of  $\mathbf{x}_j$ , and  $\boldsymbol{\epsilon}_j \in \mathbb{R}^L$  is the additive noise term. Additionally, due to the conservation of energy principle, the fractional abundance should meet two physical constraints, which are referred to as the ASC  $\sum_{i=1}^M s_{ij} = 1$  and the ANC  $s_{ij} \geq 0$ , respectively.

In general, each pixel in a scene is assumed to share the same set of endmembers  $\mathbf{A}$ . The hyperspectral data  $\mathbf{X}$  that consist of  $N$  pixels can therefore be described in the form of a matrix

$$\mathbf{X} \approx \mathbf{AS} + \mathbf{N} \quad (2)$$

where  $\mathbf{S} = [\mathbf{s}_1, \dots, \mathbf{s}_N] \in \mathbb{R}^{M \times N}$  represents the abundance matrix and  $\mathbf{N} = [\boldsymbol{\epsilon}_1, \dots, \boldsymbol{\epsilon}_N] \in \mathbb{R}^{L \times N}$  represents the noise matrix. Notably, as only the hyperspectral data  $\mathbf{X}$  are known in advance, the decomposition of  $\mathbf{X}$  falls into the class of BSS problems.

### B. Nonnegative Matrix Factorization

NMF is a group of approaches for multivariate analysis and BSS, where a matrix  $\mathbf{Y}$  is factorized into two smaller matrices  $\mathbf{V}$  and  $\mathbf{H}$ , with the property that no negative elements exist in all three matrices. Since NMF enhances the interpretability for various applications, i.e., each signal can be explained as an additive linear sum of the nonnegative basis vectors, it has drawn much attention in fields such as computer vision, text mining, document clustering, and bioinformatics.

Let the observation matrix  $\mathbf{Y} \in \mathbb{R}^{L \times N}$  be the product of the mixing matrix  $\mathbf{V} \in \mathbb{R}^{L \times M}$  and the source matrix  $\mathbf{H} \in \mathbb{R}^{M \times N}$

$$\mathbf{Y} \approx \mathbf{V}\mathbf{H} + \mathbf{U} \quad (3)$$

where the matrix  $\mathbf{U}$  is the residual term, and the elements of  $\mathbf{U}$  are not forced to be nonnegative. In addition, note that the dimension of  $\mathbf{V}$  and  $\mathbf{H}$  must be lower than that of  $\mathbf{Y}$  ( $M < \min(L, N)$ ).

In order to find an approximate factorization of (3), we need to define cost functions to quantify the quality of the approximation, which can be constructed according to a distance metric. One of the most commonly used measures is the Frobenius norm, which leads to the following optimization problem:

$$\min_{\mathbf{V}, \mathbf{H}} f(\mathbf{V}, \mathbf{H}) = \|\mathbf{Y} - \mathbf{V}\mathbf{H}\|_F \quad \text{s.t. } \mathbf{V} \geq 0, \quad \mathbf{H} \geq 0 \quad (4)$$

where  $\|\cdot\|_F$  is the Frobenius norm that is defined as  $\|\mathbf{Y}\|_F = (\sum_{ij} |y_{ij}|^2)^{1/2}$ , and  $\geq$  denotes that the matrix is element-wise nonnegative. In virtue of the mathematical connection with the linear unmixing model, NMF is an ideal tool for nonnegative constrained unmixing.

To address this optimization problem, Lin [31] presented alternating nonnegative least squares (ANLS), where the optimization problem is decomposed into two subproblems as follows:

$$\begin{aligned} \mathbf{V}^{k+1} &\leftarrow \arg \min_{\mathbf{V} \geq 0} \|\mathbf{Y} - \mathbf{V}^k \mathbf{H}^k\|_F \\ &= \max(0, \mathbf{V}^k - \omega_k \nabla f_{\mathbf{V}}(\mathbf{V}^k, \mathbf{H}^k)) \end{aligned} \quad (5)$$

$$\begin{aligned} \mathbf{H}^{k+1} &\leftarrow \arg \min_{\mathbf{H} \geq 0} \|\mathbf{Y} - \mathbf{V}^{k+1} \mathbf{H}^k\|_F \\ &= \max(0, \mathbf{H}^k - \alpha_k \nabla f_{\mathbf{H}}(\mathbf{V}^{k+1}, \mathbf{H}^k)) \end{aligned} \quad (6)$$

where the positive parameters  $\omega_k$  and  $\alpha_k$  are the step sizes, and the nonnegative constraints can be satisfied through an effective function  $\max(0, x)$ . The gradient used at each subiteration can be calculated as follows:

$$\nabla f_{\mathbf{V}}(\mathbf{V}^k, \mathbf{H}^k) = (\mathbf{V}^k \mathbf{H}^k - \mathbf{Y})(\mathbf{H}^k)^T \quad (7)$$

$$\nabla f_{\mathbf{H}}(\mathbf{V}^{k+1}, \mathbf{H}^k) = (\mathbf{V}^{k+1})^T (\mathbf{V}^{k+1} \mathbf{H}^k - \mathbf{Y}). \quad (8)$$

Sequentially, one block of variables is minimized under the nonnegative constraints and the remaining blocks are fixed. This strategy is called block coordinate descent (BCD), and there are only two block variables  $\mathbf{V}$  and  $\mathbf{H}$  in this case. Although solving subproblems (5) and (6) per update can be more expensive than the multiplicative update rules, ANLS can accelerate the iteration rate.

It is noteworthy that all of the subproblems for  $\mathbf{V}$  or  $\mathbf{H}$  are convex, yet the general optimization problem is not jointly convex for  $\mathbf{W}$  and  $\mathbf{H}$ . For any nonnegative invertible matrices  $\mathbf{D} \in \mathbb{R}^{M \times M}$  and  $\mathbf{D}^{-1}$ , the following equation holds:

$$\mathbf{V}\mathbf{H} \approx (\mathbf{V}\mathbf{D})(\mathbf{D}^{-1}\mathbf{H}). \quad (9)$$

Therefore, there can be many possible nonnegative solutions, and it is necessary to enforce additional constraints to ensure the uniqueness of the factorization in the unmixing [32].

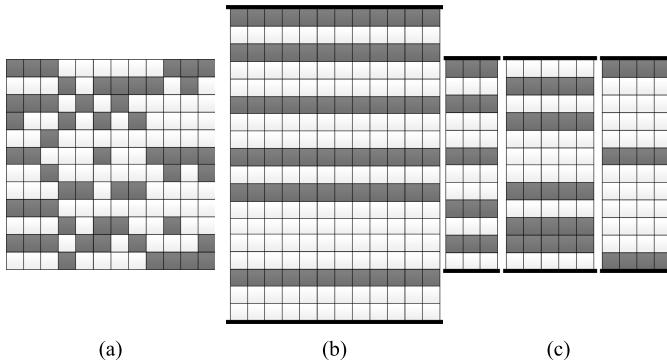


Fig. 1. Illustration of (a) sparsity, (b) collaborative sparsity, and (c) group sparsity. Each column represents the sparse pattern of a vector in the abundance matrix, and the white squares represent zero coefficients. In (a), the vectors are obtained with a simple sparsity-inducing regularizer such as the  $\ell_1$ -norm or the  $\ell_{1/2}$  quasi-norm. In (b), the collaborative sparsity is obtained by an  $\ell_2, 1$ -norm regularizer, where the gray rows denote the collaborative nonzero rows. In (c), the sparse patterns of the three column groups are induced based on the group sparsity penalty.

### C. Sparsity-Inducing Regularizers for NMF-Based HU

According to the sparse distribution of ground objects, sparsity is an important prior of the abundances, and has been explored in various NMF-based HU methods. Generally speaking, the sparsity is a helpful constraint to promote the uniqueness of the decomposition, and the optimization problem with sparsity constraint can be summarized as follows:

$$\min_{A \geq 0, S \geq 0} f(A, S) = \|X - AS\|_F + \lambda g(S) \quad (10)$$

where  $\lambda$  is used to control the tradeoff between the data fitting and the sparsity constraint, and  $g(\cdot)$  denotes the sparsity-inducing regularizer of the abundance matrix  $S$ . The  $\ell_1$ -norm is a widely used regularizer to promote a sparse representation in statistics [33], and for NMF-based HU, a typical example is DL [16], where the  $\ell_1$ -norm penalty term  $f(S) = \sum_{j=1}^N \sum_{i=1}^M |s_{ij}|$  is used to jointly learn a spectral dictionary  $A$  and a sparse representation, and the size of  $A$  is over-estimated and larger than the true number of endmembers. In SPICE [15], a reweighted  $\ell_1$ -norm regularizer term  $f(S) = \sum_{i=1}^M \gamma_i \|s^i\|_1$  is used to promote the sparsity of  $S$ , where  $s^i$  represents the  $i$ th row of the abundance matrix, and the weights  $\gamma_i$  are iteratively updated. With the same motivation, a nonconvex penalty  $\ell_{1/2}$  quasi-norm  $f(S) = \sum_{j=1}^N \sum_{i=1}^M |s_{ij}|^{1/2}$  is utilized to induce a sparser representation of  $S$  in  $L_{1/2}\text{NMF}$  [18]. The effect of the above regularizers is illustrated in Fig. 1(a), where the sparse structure of  $S$  varies pixel by pixel.

According to the similarities among the data items, there will also be some similarities among the corresponding sparse representation [34]. In recent years, in order to exploit the shared sparse pattern, mixed-norm regularizers have been introduced in statistics and machine learning [35], [36]. In NMF-based HU, a prime case is collaborative NMF (CoNMF) [37], where the number of endmembers  $\hat{M} > M$  is over-estimated, and a  $\ell_2, 1$ -norm regularizer  $f(S) = \sum_{i=1}^{\hat{M}} \|s^i\|_2$  is introduced to exploit the shared row sparsity of the whole abundance matrix  $S$ . As illustrated in Fig. 1(b),

TABLE I  
NOTATION USED IN THIS PAPER

$X = [x_1, \dots, x_N] \in \mathbb{R}^{L \times N}$	Hyperspectral data matrix
$X_r = (X^1, \dots, X^P) \in \mathbb{R}^{L \times N}$	Reordered data matrix based on segmentation
$X^i \in \mathbb{R}^{L \times n_i}$	Submatrix representing the $i$ -th superpixel
$A = [a_1, \dots, a_M] \in \mathbb{R}^{L \times M}$	Endmember matrix
$S = [s_1, \dots, s_N] \in \mathbb{R}^{M \times N}$	Abundance matrix
$S_r = (S^1, \dots, S^P) \in \mathbb{R}^{M \times N}$	Reordered abundance matrix for $X_r$
$S^i \in \mathbb{R}^{M \times n_i}$	Sparse pattern (abundance) of the $i$ -th superpixel
$C_i \in \mathbb{R}^{L+2}$	SLIC color space for HSIs
$\bar{x}^i = [x_{i1}, \dots, x_{iL}]^T \in \mathbb{R}^L$	Mean spectral signatures of the $i$ -th superpixel
$w_s$	Trade-off coefficient between the spatial and spectral similarity in SLIC
$w$	Average size of superpixels, which is used to control the number of superpixels $P$ in the scene
$W^i \in \mathbb{R}^{M \times M}$	Iteratively updated diagonal matrix controlling the sparse pattern of the $i$ -th superpixel
$c_j$	Confidence index of the $j$ -th pixel obtained during the segmentation stage
$\lambda$	Trade-off coefficient between the reconstruction and the spatial group sparsity regularizer

the  $\ell_2, 1$ -norm regularizer has the ability to promote the row sparsity of  $S$ .

Due to the similarities among the spatial neighbors, the HSI can be split into predefined spatial groups (superpixels)  $\{\vartheta_1, \dots, \vartheta_P\}$ , where the low-rank representations belonging to each group  $\vartheta$  are supposed to share the same sparse pattern. In order to exploit the group information, instead of the  $\ell_2, 1$ -norm regularizer, a spatial group sparsity regularizer  $f(S) = \sum_{p=1}^P \sum_{s_j \in \vartheta_p} c_j \|\mathbf{W}_p s_j\|_2$  [see (17) for details] is proposed in SGSNMF to promote the sparse representation in each group. As illustrated in Fig. 1(c), the sparsity pattern is explored at the level of spatial groups instead of the whole image. According to this, the local spatial information is naturally incorporated in the HU process. When compared with the  $\ell_2, 1$ -norm in CoNMF, in SGSNMF, the row sparsity-inducing mechanism is different, and it is the result of interaction between the group-wise weight matrix  $\mathbf{W}^p$  and the  $\ell_{1, 2}$ -mixed norm. Specifically, the  $\mathbf{W}^p$  controls the collaborative (row) representation, while the mixed norm promotes the sparsity of the representation. In addition, a pixel-wise confidence index  $c_j$  is defined to relax the group sparsity constraints of heterogeneous pixels, such as boundaries and small targets.

### III. SPATIAL GROUP SPARSITY REGULARIZED NMF

As mentioned in Section I, the assumptions of spatial autocorrelation and sparsity are two important priors for boosting the performance of HU. This section shows how to take the spatial and sparse priors into account by incorporating a spatial group sparsity regularizer into NMF-based HU. To help with the understanding of the following discussion, a summary of the notation is provided in Table I.

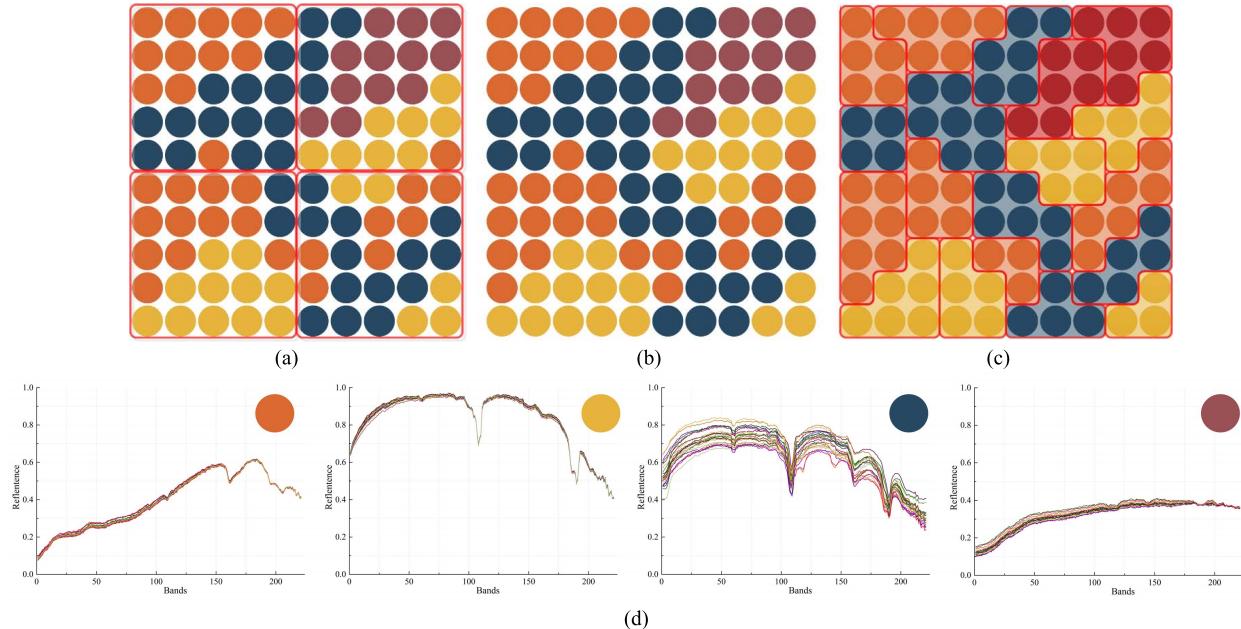


Fig. 2. Comparison between the different spatial neighborhoods in an HSI. (a) Pixel grids. (b) Pixels. (c) Superpixels. (d) Pixels in the same color show similar spectral properties.

### A. Data Preprocessing

In order to integrate the spatial and sparse priors into the HU process, image segmentation is used in the data preprocessing to generate the spatial groups used in SGSNMF. The image segmentation is aimed at labeling a set of spatially heterogeneous regions that collectively cover the whole scene. Instead of predefined pixel grids (windows), the spatial regions are defined with irregular shapes.

Let us first describe the spatial prior used in the proposed approach with an illustrative example. Fig. 2 shows the difference between superpixels and pixel grids, where the pixels that consist of an identical set of endmembers are in the same color. Fig. 2(a) shows the spatial regions that are defined by  $5 \times 5$  fixed-size windows, while Fig. 2(c) reveals the over-segmentation results. As shown in Fig. 2(a), the regular grids cannot perfectly preserve the spatial homogeneity. In contrast, superpixels, with the spectral features taken into account, can represent the spatial neighborhood explicitly and naturally, where the shape and size are adaptive and related to the spectral similarity of the neighboring pixels. In the proposed approach, the superpixels are obtained from a deliberate over-segmentation process, and as a result, compactness and homogeneity can be achieved.

In general, algorithms for image segmentation can be divided into two categories.

- 1) The graph-based methods, where each pixel is treated as a node in a graph and the superpixels are obtained by minimizing a graph-based function [38].
- 2) The gradient ascent-based methods, where the superpixels are generated by iteratively refining the clusters until convergence [39].

In this paper, a representative gradient ascent-based algorithm called SLIC [30] is introduced to generate the superpixels. This algorithm is simple in structure, fast, and contains

few parameters. More specifically, there are only two tunable parameters  $P$  and  $W_s$ , which are respectively used to control the desired number of superpixels and to weight the relative importance between the spectral and spatial similarities. However, SLIC is designed for color images in the CIELAB color space, and is not applicable to HSIs. Therefore, in this paper, an improved SLIC algorithm is put forward for generating the superpixels in the HSI. To avoid any repetition, only the differences with the original SLIC algorithm are described in this paper, and more details of SLIC can be found in [30]. The differences are summarized as follows.

- 1) Due to the increased number of spectral channels, there is a distinct change of the color space. In SGSNMF, the original spectral signatures are directly used in the clustering procedure without color space conversions, in order to minimize the loss of spectral information. For an HSI with  $L$  spectral bands, the cluster center of the  $i$ th superpixel is defined as follows:

$$\mathbf{C}_i = [x_{1i}, \dots, x_{Li}, m^i, n^i]^T \in \mathbb{R}^{L+2} \quad (11)$$

where  $\bar{\mathbf{x}}^i = [x_{1i}, \dots, x_{Li}]^T$  is the mean spectral reflectance of the  $i$ th superpixel, and  $[m^i, n^i]^T$  represents the spatial coordinates of the  $i$ th cluster center.

- 2) In the modified SLIC, superpixels correspond to clusters in the defined spectral-spatial space. To promote the homogeneity of each cluster, regular hexagonal grids [40] are introduced to generate the initial clusters, instead of the square grids used in the standard SLIC. As shown in Fig. 3, hexagons are six-sided polygons, where each edge and each corner are shared by two and three hexagons, respectively. Compared with square grids, the hexagonal grids have the following two advantages: one advantage is that each hexagon has more nondiagonal neighbors than a square, and the other is

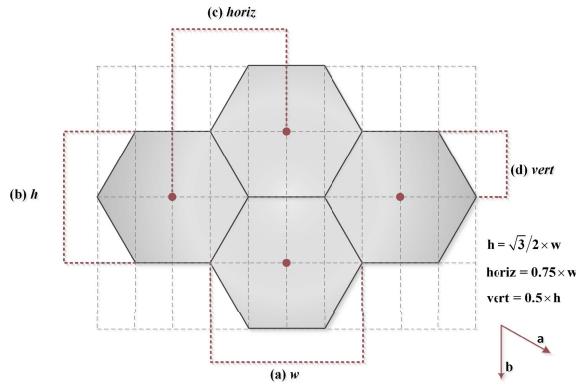


Fig. 3. Metrics of flat-topped hexagonal grids.

that hexagonal grids offer less distance distortion of boundary pixels. Due to the above advantages, more homogeneous spatial groups can be obtained by the modified SLIC with hexagonal grids as initial clusters. In addition, the nondiagonal neighbors facilitate any subsequent steps that seek to merge small superpixels. As shown in Fig. 3, the size of the hexagons can be described by the width  $w$  and height  $h$ , and the spacing of adjacent hexagons can be represented by the horizontal distance  $\text{horiz}$  and the vertical distance  $\text{vert}$ . Computing the face centers of hexagons is slightly trickier than that of squares, and can be represented as a simple matrix multiplication

$$\begin{bmatrix} m^i \\ n^i \end{bmatrix} = \begin{bmatrix} \text{horiz} & 0 \\ \text{vert} & h \end{bmatrix} \begin{bmatrix} \text{row}_i \\ \text{col}_i \end{bmatrix} \quad (12)$$

where  $[\text{row}_i, \text{col}_i]^T$  is the row and column indexes of the  $i$ th superpixel, and vectors  $\mathbf{a} = [\text{horiz} \text{ vert}]^T$  and  $\mathbf{b} = [0 \ h]^T$  consist of the hexagonal coordinates, for which the origin is at the upper left of the image. Note that the expected spatial extent of a superpixel is an approximately hexagonal region, but in the assignment step, the search for similar pixels is done in a region sized  $2w \times 2w$  around the superpixel center. Therefore, the corresponding computational complexity is still linear with regard to the number of pixels.

- 3) As the color space is changed, the corresponding spectral distance measure is updated to the spectral angle distance (SAD), which is an effective similarity measure for spectral signatures. The reason behind this is that the Euclidean distance can cause inconsistencies in clustering behavior for a varying number of spectral bands. The distance measures used in this case can be summarized as follows:

$$d_x = \cos^{-1} \left( \frac{\mathbf{x}_j^T \mathbf{x}^i}{\|\mathbf{x}_j\|_2 \|\mathbf{x}^i\|_2} \right) \quad (13)$$

$$d_{mn} = \sqrt{(m^i - m_j)^2 + (n^i - n_j)^2} \quad (14)$$

$$D_j = \sqrt{d_x^2 + \left( \frac{d_{mn}}{w} \right)^2 w_s^2} \quad (15)$$

where the spectral distance  $d_x$  and the spatial distance  $d_{mn}$  are defined by the SAD and the Euclidean distance, respectively, and  $D_j$  computes the distance between pixel  $\mathbf{x}_j$  and cluster center  $\mathbf{C}_i$ . Since the spatial distance  $d_{mn}$  varies significantly compared with  $d_x$ , the width of the hexagon  $w$  is used to normalize the spatial distance.

In the implementation of the modified SLIC, there are two parameters. The first parameter  $w_s$  weights the relative importance between the spatial and spectral similarities. When  $w_s$  is large, the spatial proximity is more important, and the resulting superpixels are more compact and have regular shape and size. When  $w_s$  is small, the spectral proximity is more important, and the resulting superpixels are more homogeneous but have less regular shape and size. When using the defined spectral–spatial space,  $w_s$  can be in the range  $[0.1, 1]$ . Since the spectral proximity is assumed to be more important in measuring the similarity, it was set to 0.3 in our experiments. The second parameter  $w$  is utilized to control the average size of the superpixels and the total number of spatial groups  $P$ . For a given hyperspectral scene and a superpixel size  $w$ ,  $P$  is easy to calculate and is roughly equal to the multiplication production of the maximum row and column indexes of the superpixels  $\text{row}_{\max} \times \text{col}_{\max}$ . In practice, in order to promote the homogeneity of each spatial group,  $w$  should be a small value in the range  $[3, 11]$ . In our experiments,  $w$  was set to 5.

### B. Spatial Group Sparsity Constraint

On the basis of the spatial groups, a new group sparsity regularizer is introduced into the NMF optimization, taking into account both spatial and sparse priors of the abundances. This regularizer is derived from the sparse feature of the abundances: the pixels that belong to a spatial group are expected to share the same sparsity pattern in the low-rank matrix. The chief aim of this regularizer is to reconstruct the observations (pixels) and interpret the sparsity of the abundances at the level of spatial objects instead of pixels.

Let us first reorder the pixels in  $\mathbf{X}$  according to the spatial groups generated by the SLIC over-segmentation with a matrix  $\mathbf{X}_r = (\mathbf{X}^1, \dots, \mathbf{X}^P) \in \mathbb{R}^{L \times N}$ , where  $P$  is the number of superpixels, and the submatrix  $\mathbf{X}^p \in \mathbb{R}^{L \times n_p}$  represents the pixels belonging to the  $p$ th superpixel. Accordingly, the abundance matrix is divided into  $P$  groups as  $\mathbf{S}_r = (\mathbf{S}^1, \dots, \mathbf{S}^P) \in \mathbb{R}^{M \times N}$ . Due to the spatial dependence, each spatial group is expected to share the same endmember assignment. Therefore, the pixels in each superpixel share the same nonzero structure in the abundance matrix. Moreover, the shared structure should be sparse, according to the fact that only a few endmembers play a dominant role in a continuous region.

Fig. 4 shows an example of  $\mathbf{X}_r$  and the group structure considered in our study. The light-gray rows of  $\mathbf{S}^p$  represent subvectors with zero abundance, while the other colored rows represent nonzero spatial sources. For example, the column vectors in  $\mathbf{S}^1$  share the same sparse structure, in that only the first and forth entries are nonzero, so that only the corresponding endmembers in  $\mathbf{A}$  are active in the reconstruction of  $\mathbf{X}^1$ . In other words, the connections between data items

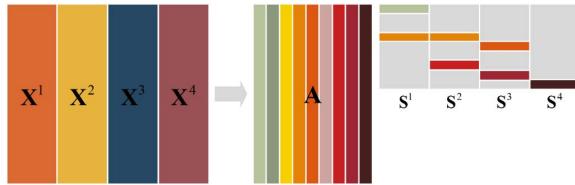


Fig. 4. Reordered observation matrix with four spatial groups and its factorization with group sparsity.

and endmembers are interpreted at the level of spatial groups instead of pixel items.

To promote the spatial group sparsity, we consider the following objective function by adding a spatial group sparsity regularization term for the abundance matrices:

$$\min_{\mathbf{A} \geq 0, \mathbf{S}_r \geq 0} f(\mathbf{A}, \mathbf{S}_r) = \frac{1}{2} \sum_{p=1}^P \|\mathbf{X}^p - \mathbf{AS}^p\|_F^2 + \lambda \sum_{p=1}^P \sum_{s_j \in \vartheta_p} c_j \|\mathbf{W}^p \mathbf{s}_j\|_2 \quad (16)$$

where  $\mathbf{S}^p = [\mathbf{s}_1, \dots, \mathbf{s}_{n_p}] \in \mathbb{R}^{M \times n_p}$  denotes the abundance matrix of spatial group  $\vartheta_p$ ,  $\lambda$  is a fixed parameter to control the tradeoff between the reconstruction and the proposed regularizer,  $c_j$  is the pixel-by-superpixel confidence index that weights the local similarity between a pixel and the superpixel it belongs to, and  $\mathbf{W}^p = \text{diag}(\mathbf{w}_1^p, \dots, \mathbf{w}_M^p) \in \mathbb{R}^{M \times M}$  is a diagonal matrix that is used to appropriately control the nonzero abundances of  $\vartheta_p$ .

According to the objective function (17), the spatial group sparsity regularization is a special case of the  $\ell_1, 2$  mixed norm with double weights. It behaves like a  $\ell_1$ -norm on vector  $\boldsymbol{\Omega} = [c_1 \|\mathbf{W}^p \mathbf{s}_1\|_2, \dots, c_{n_p} \|\mathbf{W}^p \mathbf{s}_{n_p}\|_2]^T \in \mathbb{R}^{n_p}$  [41], [42], so that the proposed regularizer can induce sparsity in the groups. The superpixel-wise weight matrix  $\mathbf{W}^p$  is designed in order to control the simultaneously relevant rows of  $\mathbf{S}^p$  that correspond to  $\vartheta_p$ . For each of  $p = 1, 2, \dots, P$ , the weights of the  $p$ th superpixel can be calculated as follows:

$$\bar{\mathbf{s}}^p \leftarrow \arg \min_{\bar{\mathbf{s}}^p} \frac{1}{2} \|\bar{\mathbf{x}}^p - \mathbf{A} \bar{\mathbf{s}}^p\|_F^2 \quad \text{s.t. } \bar{\mathbf{s}}^p \geq 0, \quad \mathbf{1}^T \bar{\mathbf{s}}^p = 1 \quad (17)$$

$$\mathbf{w}_i^p = \frac{1}{|\bar{\mathbf{s}}^p[i]| + \varepsilon} \quad (18)$$

where  $\bar{\mathbf{x}}^p$  is the mean spectral vector of the  $p$ th superpixel,  $\bar{\mathbf{s}}^p = [\bar{\mathbf{s}}^p[1], \dots, \bar{\mathbf{s}}^p[M]]^T$  is the average abundance vector, and  $\varepsilon$  is a positive parameter to stop the weights from approaching infinity, which should be set slightly smaller than the nonzero magnitudes of  $\bar{\mathbf{s}}^p$ . Note that  $\mathbf{W}^p$  is updated in every iteration step as the endmember matrix  $\mathbf{A}$  changes.

Compared to the superpixel-wise weight matrix  $\mathbf{W}^p$ , the confidence index  $c_j$  is a predefined positive weight in the segmentation process that is associated with each pixel. The confidence index  $c_j$  is inversely proportional to the spatial-spectral distance  $D_j^p$

$$c_j = \frac{1}{D_j^p}. \quad (19)$$

As the confidence index  $c_j$  increases,  $\|\mathbf{W}^p \mathbf{s}_j\|_2$  is more likely to be set to zero, and the opposite is also true. We therefore rescale the weights in order to avoid the loss of spatial details within a spatial group, such as small targets and boundary pixels of segments, which may not follow the designed nonzero pattern. For example, for unique small targets, the sparse structure is totally at variance with that of the surrounding pixels. A boundary pixel is also more likely to be a mixed pixel with dense abundances. Due to the spectral or spatial dissimilarity of these heterogeneous pixels, the corresponding confidence index values are relatively small. Therefore, the boundary effects and the information loss of small targets can be avoided, to some extent, in SGSNMF.

### C. Optimization Algorithm

With the spatial group sparsity regularization, the optimization problem is more difficult to solve than the standard one. In this section, we discuss the optimization algorithm for SGSNMF. As mentioned in Section II, BCD is a common way to approximately optimize the objective function with respect to one variable block at a time, while the others are kept fixed.

For the proposed objective function (16), the optimization consists of three main subproblems in the iteration: 1) update the superpixel-wise weight matrix  $\mathbf{W}^p$ ; 2) update the abundance matrix  $\mathbf{S}_r$ ; and 3) update the endmember matrix  $\mathbf{A}$ . For a given endmember matrix  $\mathbf{A}$ , subproblem (1) can easily be solved according to (17) and (18). Note that due to the different superpixel-wise weights, subproblem (2) consists of  $P$  independent ANLS problems, which can therefore be addressed in parallel. The extensions of the aforementioned update rules for each abundance group  $\mathbf{S}^P$  and endmember matrix  $\mathbf{A}$  are introduced as follows:

$$\mathbf{S}^P \leftarrow \arg \min_{\mathbf{S}^P \geq 0} f(\mathbf{A}, \mathbf{S}^P) = \frac{1}{2} \|\mathbf{X}^p - \mathbf{AS}^p\|_F^2 + \lambda \sum_{s_j \in \vartheta_p} c_j \|\mathbf{W}^p \mathbf{s}_j\|_2 \quad (20)$$

$$\mathbf{A} \leftarrow \arg \min_{\mathbf{A} \geq 0} f(\mathbf{A}, \mathbf{S}_r) = \frac{1}{2} \|\mathbf{X}_r - \mathbf{AS}_r\|_F^2. \quad (21)$$

To prevent the elements of  $\mathbf{A}$  and  $\mathbf{S}_r$  from growing arbitrarily large in the iteration, the ASC should be added in the learning of the abundances. In this paper, we adopt an effective method that is used in FCLS [9], where the endmember matrix  $\mathbf{A}$  and each spatial group  $\mathbf{X}^p$  are augmented before the learning of the abundance matrix as

$$\tilde{\mathbf{X}}^p = \begin{bmatrix} \mathbf{X}^p \\ \delta \mathbf{1}_{n_p}^T \end{bmatrix} \quad \tilde{\mathbf{A}} = \begin{bmatrix} \mathbf{A} \\ \delta \mathbf{1}_M^T \end{bmatrix} \quad (22)$$

where  $\delta$  is a positive value that is used to balance the estimate accuracy and additivity constraint. The sum of the abundance is forced to converge toward unity as  $\delta$  increases [13]. In the experiments,  $\delta$  was set to 15 in order to lead to a tradeoff.

To extend the update rules, we need to recalculate the partial derivatives to the elements in  $\mathbf{S}^P$  and  $\mathbf{A}$ . Because of the absence of constraints for the endmember matrix, the gradient of  $\mathbf{A}$  is easy to solve

$$\nabla f_{\mathbf{A}}(\mathbf{A}, \mathbf{S}_r) = (\mathbf{AS}_r - \mathbf{X}_r) \mathbf{S}_r^T. \quad (23)$$

In comparison, due to the spatial group sparsity constraint, the objective function of (21) consists of two terms

$$h_1(\mathbf{S}^p) = \frac{1}{2} \|\tilde{\mathbf{X}}^p - \tilde{\mathbf{A}}\mathbf{S}^p\|_F^2 \quad (24)$$

$$h_2(\mathbf{S}^p) = \lambda \sum_{\mathbf{s}_j \in \vartheta_p} c_j \|\mathbf{W}^p \mathbf{s}_j\|_2. \quad (25)$$

Note that the first term  $h_1(\mathbf{S}^p)$  is a convex differentiable function, and  $h_2(\mathbf{S}^p)$  is convex but nonsmooth. Specifically,  $h_2(\mathbf{S}^p)$  is not differentiable only at point  $\mathbf{W}^p \mathbf{s}_j = \mathbf{0}$ . However, the equation  $\mathbf{W}^p \mathbf{s}_j = \mathbf{0}$  fails to hold in this study, according to the ASC and the fact that  $\mathbf{W}^p > 0$ . Therefore, for  $\forall p, j$ , such that  $\|\mathbf{W}^p \mathbf{s}_j\|_2 \neq 0$ , we have

$$\nabla h_2(\mathbf{s}_j) = \lambda \frac{c_j}{\|\mathbf{W}^p \mathbf{s}_j\|_2} (\mathbf{W}^p)^T \mathbf{W}^p \mathbf{s}_j. \quad (26)$$

We then obtain

$$\nabla h_2(\mathbf{S}^p) = \lambda (\mathbf{W}^p)^T \mathbf{W}^p \mathbf{S}^p \mathbf{B}^p \quad (27)$$

$$\mathbf{B}^p = \text{diag} \left( \frac{c_1}{\|\mathbf{W}^p \mathbf{s}_1\|_2}, \frac{c_2}{\|\mathbf{W}^p \mathbf{s}_2\|_2}, \dots, \frac{c_{n_p}}{\|\mathbf{W}^p \mathbf{s}_{n_p}\|_2} \right) \quad (28)$$

where  $\mathbf{B}^p \in \mathbb{R}^{n_p \times n_p}$  is a diagonal matrix. Then, according to (8) and (27), the gradient of  $\mathbf{S}^p$  can be represented as follows:

$$\nabla f_{\mathbf{S}^p}(\tilde{\mathbf{A}}, \mathbf{S}^p) = \tilde{\mathbf{A}}^T (\tilde{\mathbf{A}}\mathbf{S}^p - \tilde{\mathbf{X}}^p) + \lambda (\mathbf{W}^p)^T \mathbf{W}^p \mathbf{S}^p \mathbf{B}^p. \quad (29)$$

The learning rules for the endmember matrix and each group abundance matrix can be represented as

$$\mathbf{A} = \max(0, \mathbf{A} - \omega_k \nabla f_{\mathbf{A}}(\mathbf{A}, \mathbf{S}_r)) \quad (30)$$

$$\mathbf{S}^p = \max(0, \mathbf{S}^p - \sigma_k \nabla f_{\mathbf{S}^p}(\tilde{\mathbf{A}}, \mathbf{S}^p)) \quad (31)$$

where the selection of step size parameters  $\omega_k$  and  $\sigma_k$  is based on the Armijo rule [43]. We let  $\omega_0$  denote the initial step size, and the step size  $\omega_k$  in (30) is iteratively updated by a scaling parameter  $\beta \in (0, 1)$ , such that  $\omega_k = \omega_0 \beta^{t_k}$ , where  $t_k$  is the first nonnegative integer  $t$  to meet the sufficient decrease condition

$$f(\mathbf{A}^{k+1}, \mathbf{S}_r) - f(\mathbf{A}^k, \mathbf{S}_r) \leq \partial \nabla f_{\mathbf{A}}(\mathbf{A}^k, \mathbf{S}_r)(\mathbf{A}^{k+1} - \mathbf{A}^k) \quad (32)$$

where  $\partial \in (0, 1)$  is the tolerance. Since  $\omega_{k-1}$  and  $\omega_k$  may be similar, a trick is to use  $\omega_{k-1}$  as the initial guess of  $\omega_k$ , and then either increase or decrease it to meet the Armijo rule. In the implementation, the step size  $\omega_k$  is selected according to [30, Algorithm 4], where the scaling parameter  $\beta$  is set to 0.1 to reduce the cost per iteration. The step size  $\sigma_k$  is updated based on the same procedure.

#### D. Implementation Issues

In this section, some implementation issues are discussed, including the initialization of the endmember matrix, the estimation of the number of endmembers, and the termination condition.

The first issue is to initialize  $\mathbf{A}$  and  $\mathbf{S}$  for NMF. Due to the fact that the joint optimization problem for NMF is nonconvex,

the unmixing result is sensitive to the initialization, to some extent, even in the constrained case. To minimize the distance between the initial point and the global optimum, EE methods are often used to generate the initial endmembers. In this paper, three initializations are studied: 1) random initialization; 2) VCA-FCLS initialization; and 3) region-based VCA-FCLS initialization. Random initialization of  $\mathbf{A}$  and  $\mathbf{S}$  was employed in the real-data experiment, where each random value was set in the interval  $[0, 1]$ , while VCA-FCLS initialization was used in the simulated experiments to speed up the convergence. In addition, we also initialized the endmember matrix based on the results of the segmentation and the region-based VCA [44]. More specifically, the mean spectra of the  $M$  most distinct superpixels are extracted as the initial endmembers. The main purpose of this is to fully exploit the spatial information. It is noteworthy that the region-based initialization is more robust to noise, and can provide a more accurate initial point for the HU.

For the iteration stopping rule, two commonly used criteria are adopted in SGSNMF: the maximum number of iterations and the error tolerance. The number of iterations is limited by a predefined maximum iteration number. In addition, if the gradient of the cost function (16) is less than  $10^{-3}$  of the initial value, the optimization stops.

Another issue is the estimation of the accurate number of endmembers, which can be regarded as an independent topic for HU in general. Therefore, in our experiments, the number of endmembers was given in advance. Note that the number of endmembers can also be automatically estimated by the Neyman–Pearson based thresholding method (HFC) [45] or hyperspectral signal subspace identification by minimum error (HySime) [46].

Finally, we summarize the proposed SGSNMF algorithm (with superpixel-based initialization) in Algorithm 1.

---

#### Algorithm 1 Summary of SGSNMF

---

##### *Input:*

$\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N] \in \mathbb{R}^{L \times N}$ ,  $M \in \mathbb{Z}_+$ , and  $\lambda, w, w_s \in \mathbb{R}_+$ .

##### *Pre-processing:*

- a) Generate  $\mathbf{P}$  spatial groups  $\mathbf{X}_r$  based on SLIC.
- b) Initialize  $\mathbf{A}$  and  $\mathbf{S}_r$  based on region-based EE.
- c) Calculate the confidence index  $c_j$  by (19).

##### *Repeat until convergence:*

Augment  $\mathbf{X}_r$  and  $\mathbf{A}$  to  $\tilde{\mathbf{X}}_r$  and  $\tilde{\mathbf{A}}$  by (22), respectively.

For each  $p = 1, 2, \dots, P$ :

- a) Update the weight matrix  $\mathbf{W}^p$  by (17) and (18).
- b) Update the abundance matrix  $\mathbf{S}^p$  by (29) and (30).
- end
- c) Update the endmember matrix  $\mathbf{A}$  by (30) and (31).

##### *Output:*

$\mathbf{A} = [\mathbf{a}_1, \dots, \mathbf{a}_M] \in \mathbb{R}^{L \times M}$ , and  $\mathbf{S} = [\mathbf{s}_1, \dots, \mathbf{s}_N] \in \mathbb{R}^{M \times N}$ .

---

## IV. EXPERIMENTS WITH SIMULATED DATA

In this section, a comprehensive performance comparison is given between SGSNMF and some state-of-the-art unmixing methods—VCA-FCLS, MVCNMF, ASSNMF, and

$L_{1/2}$ NMF—by the use of two simulated data sets. The goal is to analyze the effect of the proposed spatial group sparsity prior in both the endmember and abundance estimations.

In order to compare the unmixing performance, we introduce two indicators: the SAD and the root-mean-square error (RMSE). The SAD is used to measure the spectral angle between the estimated endmember  $\hat{\mathbf{a}}_m$  and the corresponding ground-truth endmember signature  $\mathbf{a}_m$

$$SAD_m = \cos^{-1} \frac{\hat{\mathbf{a}}_m^T \mathbf{a}_m}{\|\hat{\mathbf{a}}_m\|_2 \|\mathbf{a}_m\|_2} \quad (33)$$

where the units of the SAD values are in radians in the following tables and charts. The other metric, RMSE, is used to describe the difference between the estimated abundance  $\mathbf{S}_m$  and the corresponding ground truth  $\hat{\mathbf{S}}_m$ , and can be calculated as follows:

$$RMSE_m = \frac{1}{\sqrt{N}} \|\mathbf{S}_m - \hat{\mathbf{S}}_m\|_F. \quad (34)$$

The rest of this section is organized as follows. Section IV-A describes the generation and the properties of the two simulated data sets. In Section IV-B, the influences of noise, the initialization, the spatial group size, and the endmember number on the unmixing performance are discussed and analyzed.

#### A. Simulated Data Sets

1) *Simulated Data 1*: This simulated data set (Simu-1), with 221 spectral bands and a size of  $100 \times 100$  pixels, was generated according to the simulation procedure proposed by Hendrix *et al.* [47]. As shown in Fig. 5, Simu-1 consists of nine-endmember signatures selected from the USGS spectral library, namely, Kaolinite KGa-1, Dumortierite, Nontronite, Alunite, Sphene, Pyrobelite, Halloysite, Muscovite, and Kaolinite CM9. The corresponding abundance maps were generated by  $k$ -means clustering [48] and a Gaussian filter with both nonnegative and additive constraints. Simu-1 simulates the real-world distribution of objects; that is, the boundary pixels are more heavily mixed than those in the center of homogeneous regions.

In addition, to imitate the process of data acquisition, different noise levels were added to the simulated images, with respect to a specified signal-to-noise ratio (SNR)

$$SNR = 10 \log_{10} \frac{E[\mathbf{x}^T \mathbf{x}]}{E[\mathbf{n}^T \mathbf{n}]} \quad (35)$$

where  $\mathbf{x}$  and  $\mathbf{n}$  represent the original signal and the corresponding noise, respectively, and  $E[\cdot]$  is the expectation operator.

2) *Simulated Data 2*: The second simulated data set (Simu-2), with a size of  $128 \times 128$  pixels and 224 spectral channels, was generated using the Hyperspectral Imagery Synthesis toolbox<sup>1</sup> for MATLAB. This allows the user to choose appropriate parameters for the simulation procedure, such as the spatial dimensions, the distribution of the abundances, and the number of endmembers  $M$ , which provides considerable

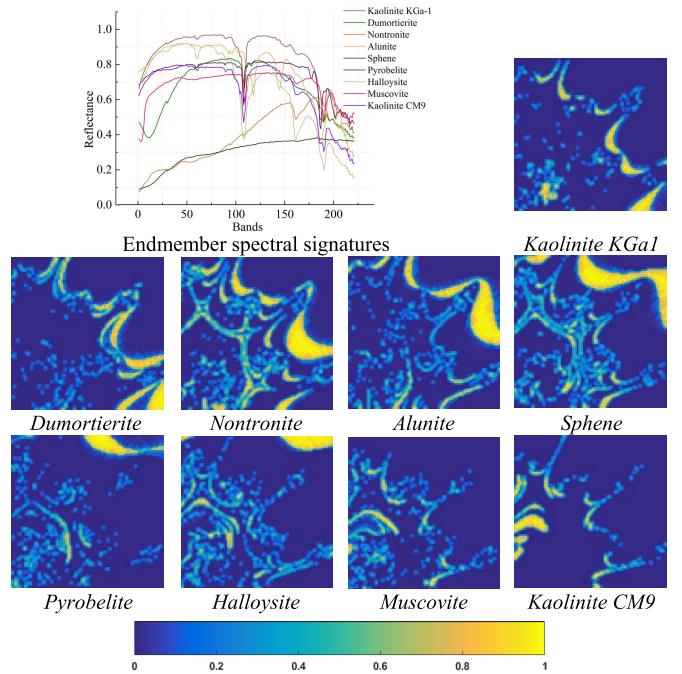


Fig. 5. Endmember signatures and abundance maps for Simu-1.

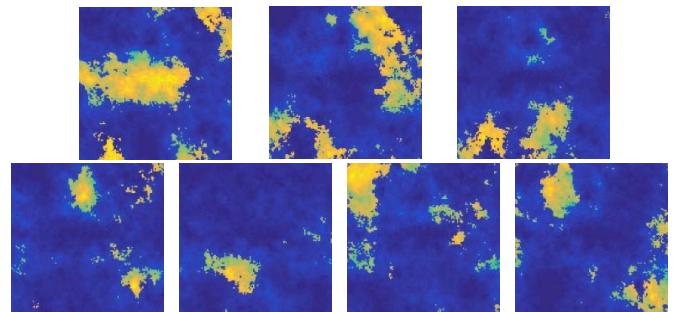


Fig. 6. Abundance maps for Simu-2 in the case of  $K = 7$ .

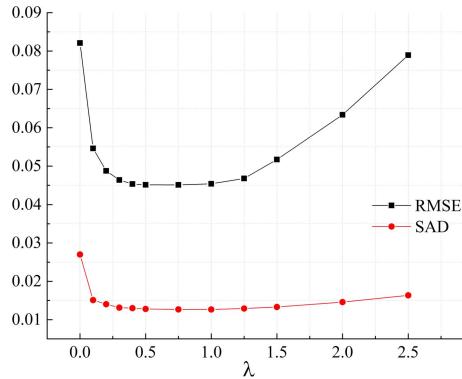
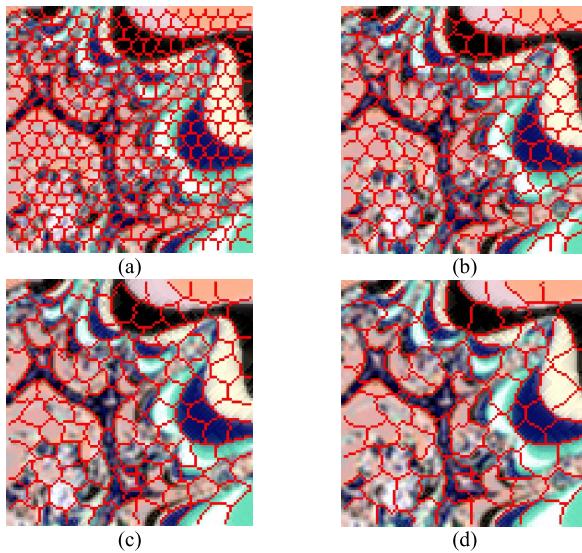
flexibility and choice for the user. Note that Simu-2 was only used in the sensitivity analysis of endmember numbers, where  $M$  was varied from 3 to 15 with an interval of 2. Specifically, the  $M$  endmembers were randomly selected from the USGS mineral spectra library, while the corresponding abundance maps were generated according to the spherical Gaussian field. Fig. 6 shows the ground-truth abundance maps of Simu-2 for the seven-endmember case.

#### B. Simulated Experiments and Analysis

This section describes the five simulated experiments, in order to analyze the unmixing performance from all aspects. Note that all of the experimental results in this paper are the average values obtained by 20 random tests. Due to the fact that the NMF-based methods are sensitive to the initial points, we adopted VCA-FCLS as the default initialization for all the NMF-based methods, unless noted otherwise. In addition, the adjustable parameters for the compared algorithms were consistent with the original references.

1) *Parameter Analysis*: In the first experiment, parameter  $\lambda$  was analyzed with Simu-1 in the condition of 20-dB

<sup>1</sup>[http://www.ehu.es/ccwintco/index.php/Hyperspectral\\_Imagery\\_Synthesis\\_tools\\_for\\_MATLAB](http://www.ehu.es/ccwintco/index.php/Hyperspectral_Imagery_Synthesis_tools_for_MATLAB)

Fig. 7. Performance analysis of SGSNMF with respect to  $\lambda$ .Fig. 8. Illustration of SLIC superpixels, where the group size  $w$  was varied from 5 to 11. (a)  $w = 5$ . (b)  $w = 7$ . (c)  $w = 9$ . (d)  $w = 11$ .

white Gaussian noise. In order to make better use of the spatial information, the initialization of SGSNMF was utilized. Fig. 7 illustrates the varying tendency of the unmixing performance as  $\lambda$  increases. As can be seen in Fig. 7, a stable and excellent performance is obtained when parameter  $\lambda$  varies from 0.2 to 1.2. Due to the tradeoff between the reconstruction and the group sparsity, when  $\lambda$  is larger than 1.5, the reconstruction term is relaxed, and there is a distinct downward trend in the abundance estimation accuracy. Thus, in order to obtain a stable and accurate HU solution, the optimal value of  $\lambda$  can be set in the range [0.2, 1.2]. In the following experiments, parameter  $\lambda$  was set to 0.3.

2) *Simulated Experiment With Different Kinds of Pixel Grids:* This experiment was initiated to check the validity of the proposed spatial prior. The analysis covers the spatial groups which were generated by both pixel grids and SLIC superpixels. In addition, the pixel-wise sparse CNMF method, L<sub>1/2</sub>NMF, was also introduced for comparison. The SLIC superpixels with different group sizes are illustrated in Fig. 8, where the RGB images are simulated with bands 50, 100, and 150.

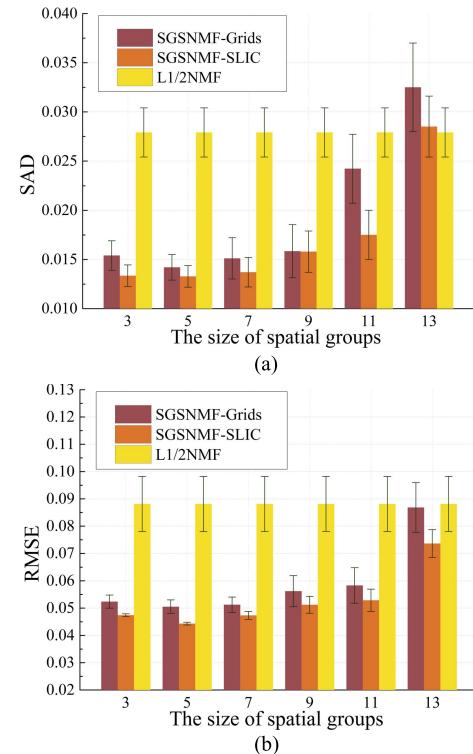


Fig. 9. Comparison between different spatial priors. The group size was varied from 3 to 13. (a) SAD. (b) RMSE.

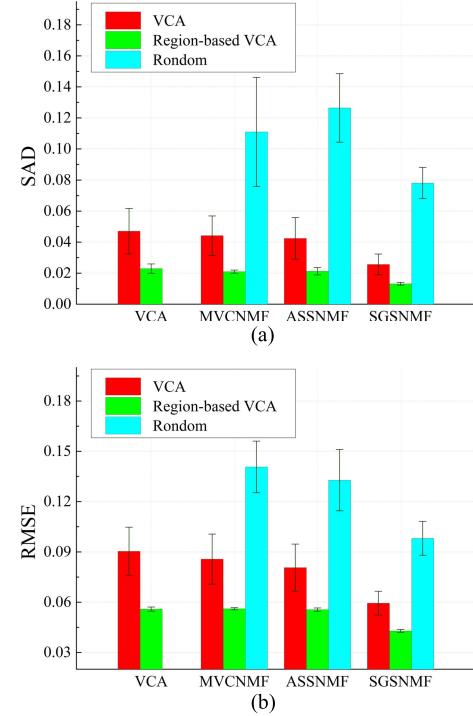


Fig. 10. Comparison between the BSS-based unmixing algorithms with different initializations. (a) SAD. (b) RMSE.

In this experiment, Simu-1 was polluted by 20-dB Gaussian white noise, and the maximum number of iterations for all the algorithms was set to 100. The size of the spatial groups was varied from 3 to 13 in order to meet the over-segmentation needs. Specifically, the size of the SLIC superpixels was

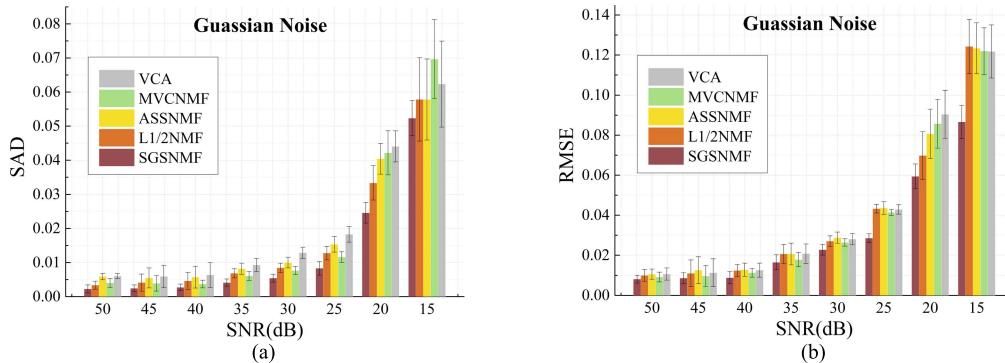


Fig. 11. Comparison between the algorithms with the Gaussian noise. (a) SAD. (b) RMSE.

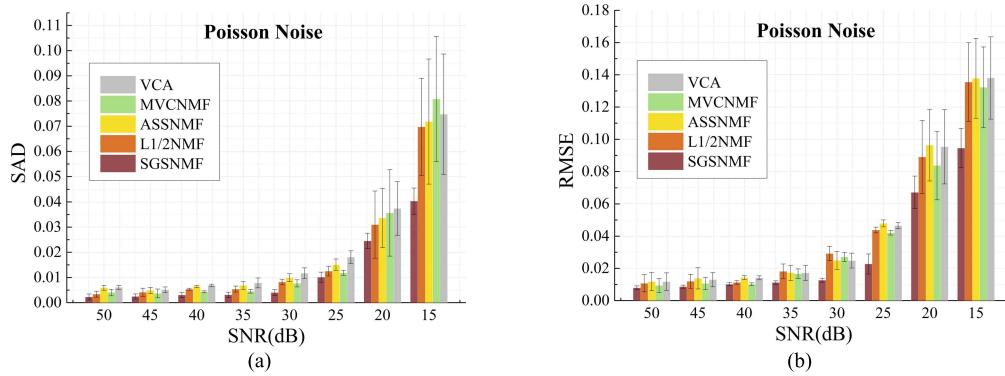


Fig. 12. Comparison between the algorithms with the Poisson noise. (a) SAD. (b) RMSE.

defined as the width of a hexagon. The average experimental results and the corresponding standard deviations are shown in Fig. 9. It is clear from Fig. 9 that SGSNMF shows an improved performance over  $L_{1/2}$ NMF, which proves the validity of the proposed group sparse regularization term. For the shape of the spatial groups, superpixels are superior to pixel grids. This is mainly because superpixels are the natural and adaptive representations of a scene, and the distribution of endmembers is sparser when compared with that of pixel grids. In addition, for the Simu-1 data set, the optimum group size is around 5, and the unmixing results are also acceptable for other group sizes, such as 3, 7, 9, and 11.

3) *Sensitivity Analysis to Different Initializations*: In this experiment, different initializations were introduced and compared in the aspect of unmixing performance and precision, including random initialization, VCA-FCLS, and region-based VCA-FCLS. More specifically, the candidate endmembers for the region-based VCA were the mean spectra of the SLIC superpixels. For the random initialization,  $A$  and  $S$  were both initialized by setting their entries to random values between 0 and 1. Fig. 10 shows the unmixing results obtained with different initializations in the condition of 20-dB white Gaussian noise (Simu-1). Since the VCA-FCLS initialization provides a more accurate initial point that is closer to the global optimal solution, the corresponding unmixing results outperform those obtained by the random initialization. Among the three initializations, the region-based method produces the highest accuracy, and shows strong stability and robustness with regard to noise. This is mainly because the accurate

SLIC segmentation results effectively enhance the SNR of the candidate endmembers, while the spatial property of the original data is also well preserved.

4) *Robustness Analysis to Noise*: We carried out this experiment to compare and analyze the robustness of the algorithms. In this experiment, the Simu-1 data set was polluted by the white Gaussian noise, as well as the Poisson noise with different noise levels, where the SNR was varied from 15 to 50 dB with an interval of 5 dB. The number of endmembers was known in advance for all the algorithms. Figs. 11 and 12 illustrate the average performances of the different algorithms in the conditions of the Gaussian noise and Poisson noise, respectively. Overall, it can be observed that the proposed SGSNMF favorably compares to the other algorithms, with regard to the SAD values and RMSE values in most SNR cases. Meanwhile, all the algorithms show a decreased unmixing accuracy as the SNR decreases. In detail, MVCNMF obtains the lowest accuracy at 15 dB. This is mainly because the minimum-volume constraint fails to treat the case of low SNR. However, the proposed method exhibits much better robustness in the abundance estimation, thanks to the spatial and sparse priors.

5) *Sensitivity Analysis to Endmember Numbers*: We conducted this experiment with the Simu-2 data set in order to verify the applicability of the proposed method with regard to different endmember numbers. This experiment was conducted under the condition of 20-dB white Gaussian noise. The number of endmembers  $M$  was varied from 3 to 15. In general, the difficulty of the unmixing increases as  $M$  increases. As can

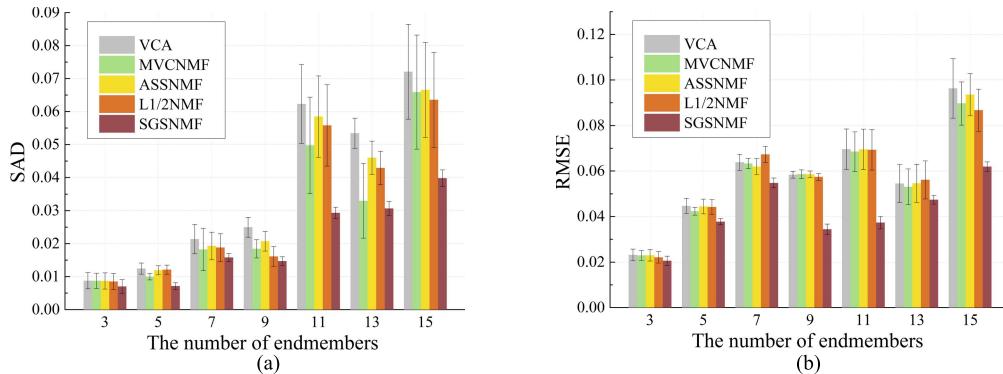


Fig. 13. Comparison between the algorithms with a different number of endmembers (from 3 to 15). (a) SAD. (b) RMSE.

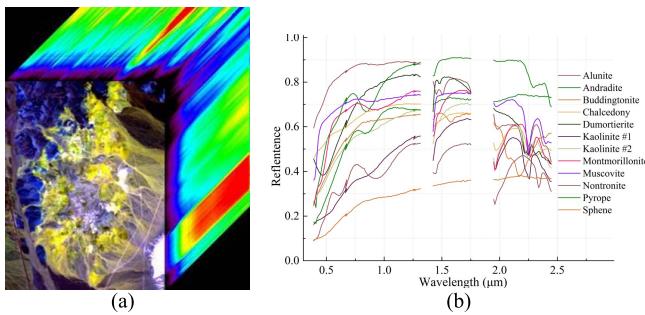


Fig. 14. AVIRIS hyperspectral data from Cuprite, NV, USA. (a) 3-D cube form of the Cuprite experiment area. (b) USGS library mineral spectra signatures.

be seen from Fig. 13, the overall unmixing accuracies are similar when  $M$  is relatively small, whereas the proposed method shows an obvious superiority as  $M$  increases. These differences can be attributed to the group constraint of the proposed method and the sparse structure of the abundances.

## V. REAL-DATA EXPERIMENTS

Two real HSIs were also used to evaluate the unmixing accuracy. The selected real hyperspectral data were obtained from the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) and an unmanned aerial vehicle (UAV)-borne hyperspectral sensor, with different spatial resolutions.

### A. Real-Data Experiments With Cuprite Data

The first data set was a  $250 \times 190$  subimage of the Cuprite data set captured by AVIRIS in June 1997 in NV, USA, with 224 bands covering the wavelength range of 370 to 2480 nm. After removing the water absorption bands (106–115 and 151–169) and low SNR bands (1–3 and 221–224), a subimage with 188 bands was considered in this experiment. The Cuprite data cube and the reference spectra are shown in Fig. 14. According to the slight spectral variations of the same type of minerals with different chemical compositions, 12 kinds of endmembers were selected for unmixing in this experiment: Alunite, Andradite, Buddingtonite, Dumortierite, Kaolinite\_1, Kaolinite\_2, Muscovite, Montmorillonite, Nontronite, Pyrope, Sphene, and

Chalcedony. Fig. 15 shows the abundance maps estimated by SGSNMF. The estimations correlate well with the geological maps of the Cuprite data set [49], and some endmembers with distinct distributions, such as Buddingtonite, Muscovite, and Montmorillonite, are well identified by the proposed method.

In order to evaluate the unmixing performance in the real hyperspectral scenes quantitatively, two comparisons were undertaken. The first was an SAD comparison between the estimated endmember signatures and the reference spectra from the USGS library. The SAD values of all the methods (in radians) are listed in Table II, where the results are the average values of 20 repeated experiments. The best numerical results are displayed in bold. It can be concluded from Table II that the endmember estimations of SGSNMF are better than those generated by the other methods.

Due to the differences between the image spectra and the field spectra from the USGS library, the SAD values are not a sufficient indicator of HU accuracy. Therefore, we calculated the reconstruction errors of each method to further evaluate the accuracy. If the estimations are accurate, only some small errors should be shown in the reconstruction maps. Table III shows the average reconstruction errors and the corresponding error maps of each method. The number of endmembers was set to  $M = 12$  and  $M = 14$  to minimize the impact of the unknown number of endmembers. As shown, it is clear that MVCNMF and SGSNMF outperform the other methods in the reconstruction. Even when  $M$  increases to 14, there are still some notable errors in the top right of the error maps of VCA, ASSNMF, and L<sub>1</sub>/2NMF. To summarize, the joint comparisons demonstrate that the proposed method can obtain improved unmixing results in a real hyperspectral data set.

### B. UAV-Borne Hyperspectral Data

The second real data set was a  $600 \times 400$  subimage of the “WHU-GX UAV” image captured by the authors with a Nano-Hyperspec<sup>2</sup> sensor in August 2016, in Guangxi province of China. Unlike airborne hyperspectral imaging, this sensor was integrated with a lightweight multirotor UAV known as the Aibot X6.<sup>3</sup> The scene comprises 270 spectral bands in the

<sup>2</sup><http://www.headwallphotronics.com/spectral-imaging/hyperspectral/nano-hyperspec>

<sup>3</sup><https://www.aibotix.com/>

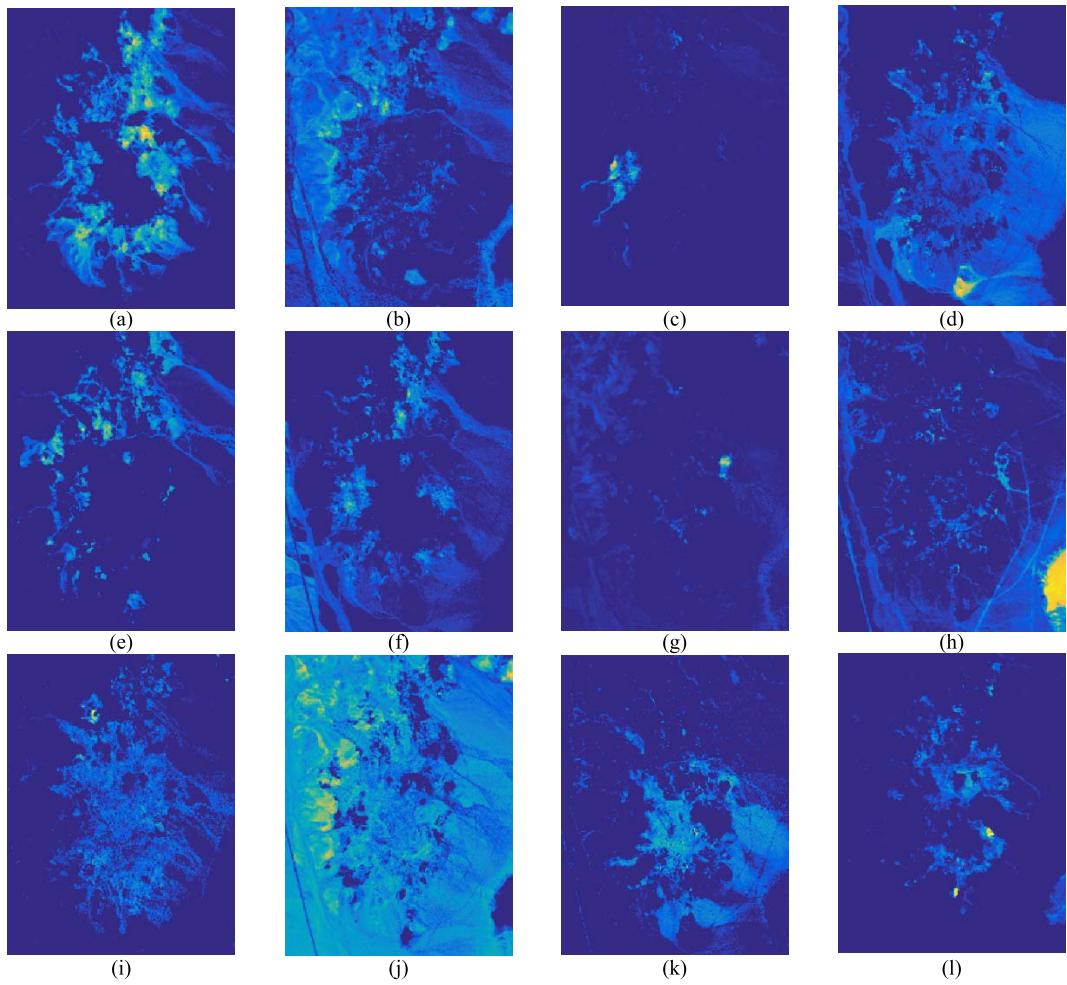


Fig. 15. Abundance maps estimated by SGSNMF. (a) Alunite'. (b) Andradite. (c) Buddingtonite. (d) Dumortierite. (e) Kaolinite\_1. (f) Kaolinite\_2. (g) Muscovite. (h) Montmorillonite. (i) Nontronite. (j) Pyrope. (k) Sphene. (l) Chalcedony.

TABLE II  
SAD COMPARISON FOR THE CUPRITE DATA SET

<i>Endmember</i>	<i>Method</i>	<i>VCA-FCLS</i>	<i>MVCNMF</i>	<i>L<sub>1/2</sub>NMF</i>	<i>ASSNMF</i>	<i>SGSNMF</i>
<i>Alunite</i>	<b>0.0792</b>	0.1329	0.0963	0.0897	0.1093	
<i>Andradite</i>	0.0736	<b>0.0812</b>	0.0865	0.0810	<b>0.0677</b>	
<i>Buddingtonite</i>	<b>0.0999</b>	0.0848	0.0808	0.0857	<b>0.0789</b>	
<i>Dumortierite</i>	<b>0.0720</b>	0.0916	0.0847	<b>0.0720</b>	0.0853	
<i>Kaolinite #1</i>	0.0819	0.0936	0.0748	0.0752	<b>0.0718</b>	
<i>Kaolinite #2</i>	<b>0.0658</b>	0.0675	0.0665	0.0699	0.0705	
<i>Montmorillonite</i>	0.1617	0.2605	0.1266	<b>0.1027</b>	0.1310	
<i>Muscovite</i>	0.0598	0.0671	0.0611	0.0586	<b>0.0579</b>	
<i>Nontronite</i>	<b>0.0665</b>	0.1589	0.0798	0.0689	0.0768	
<i>Pyrope</i>	0.1844	0.1249	0.1179	0.1502	<b>0.0884</b>	
<i>Sphene</i>	0.2298	<b>0.0731</b>	0.2002	0.1858	0.1428	
<i>Chalcedony</i>	0.0826	0.1068	<b>0.0775</b>	0.1031	0.1158	
<i>Average</i>	0.1048	0.1119	0.0961	0.0952	<b>0.0913</b>	

visible and near-infrared (400–1000 nm) spectral range, with a spatial resolution of 0.1 m per pixel. Before unmixing, the low SNR bands were removed, leaving 239 bands to be used in the experiment. Fig. 16 shows this UAV-borne hyperspectral data cube and the reference endmember spectra, including cement (road and canals), water, crops, and soil (path on the field).

As shown in Fig. 16(b), the distribution of the four end-members is quite simple; that is, water is found in the canal in the lower right of the image, the road and the edges of the two canals are made up of cement, the soil is located on the H-shaped path in the field and the boundaries, and crops cover the other major parts. Although the spatial resolution of the UAV image is relatively high, there are still some

TABLE III  
RECONSTRUCTION ERROR COMPARISON FOR THE CUPRITE DATA SET

<i>Method</i>	<i>VCA-FCLS</i>	<i>MVCNMF</i>	<i>ASSNMF</i>	<i>L<sub>1/2</sub>NMF</i>	<i>SGSNMF</i>
<i>M=12</i>					
<i>RMSE</i>	0.1418	<b>0.0572</b>	0.1111	0.0914	0.0591
<i>M=14</i>					
<i>RMSE</i>	0.0922	0.0524	0.0752	0.0729	<b>0.0511</b>

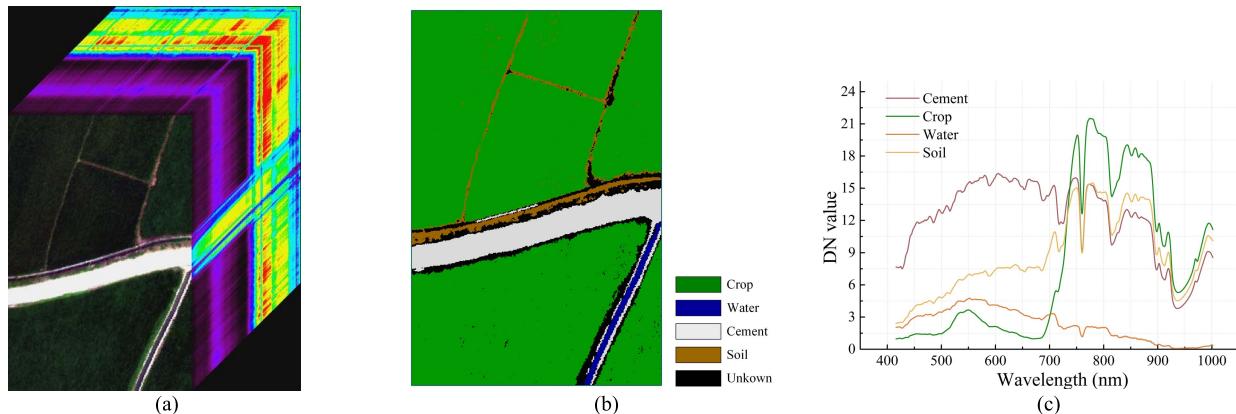


Fig. 16. UAV-borne hyperspectral data from Guangxi, China. (a) 3-D cube form of the experiment area. (b) Ground truth. (c) Reference spectral signatures of four endmembers.

mixed pixels located in the boundaries and the path in the field.

The abundance maps estimated by each algorithm are shown in Table IV. Overall, the abundances estimated by SGSNMF are sparser and agree with the ground truth very well. As shown in the ground-truth map, the ground truth of the path in the field is soil, and some crops are found in the boundaries. Due to the shading effect on the path, the abundances on the narrow path are difficult to estimate, and were estimated as some low-reflectance endmembers (e.g., water) by the compared methods. In contrast, the major abundance estimation of the path is soil for SGSNMF. This is mainly because the spatial group sparsity leads to a tradeoff, as a result, the ASC is slightly relaxed and the estimation is more accurate. However, it should be noted that there

are still some inaccurate abundances (e.g., cement) in the path estimated to fit the reconstruction errors of SGSNMF. Due to the width of the path being less than  $w$ , for some heterogeneous boundary pixels, the confidence index  $c_j$  is small, the reconstruction term plays an important role, and therefore the sparse pattern is not forced to be followed. To evaluate the unmixing performance quantitatively, the SAD values between the estimated and reference signatures are compared in Table V. The SAD results agree with the finding in the abundance maps, in that the overall performance of SGSNMF is much better than that of the compared methods.

### C. Efficiency Analysis With Real Hyperspectral Data

The effectiveness of the proposed algorithm was verified by the use of two real hyperspectral data sets, using

TABLE IV  
COMPARISON BETWEEN THE ABUNDANCE MAPS FOR THE UAV-BORNE HYPERSPECTRAL DATA

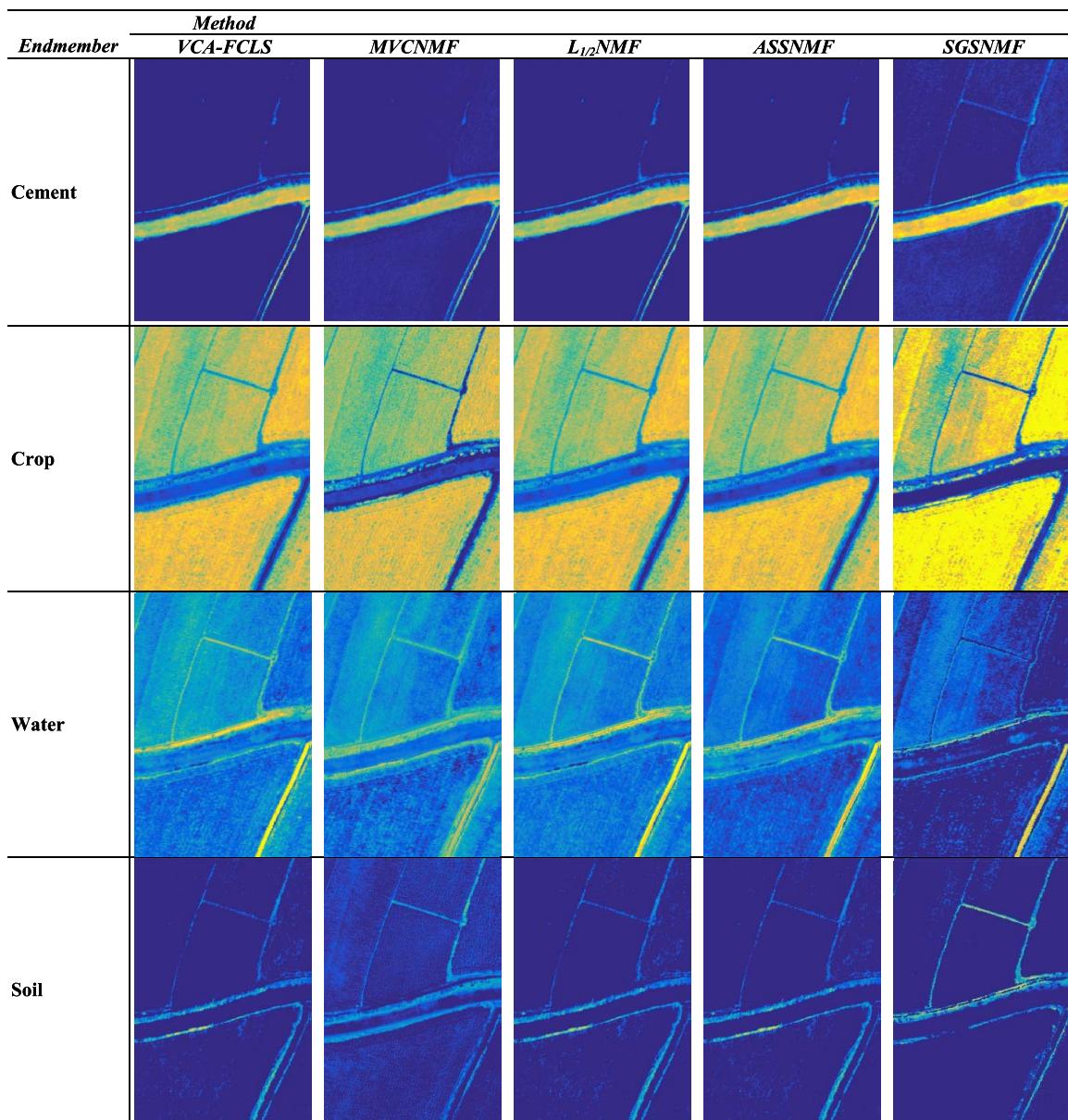


TABLE V  
SAD COMPARISON FOR THE UAV-BORNE HYPERSPECTRAL DATA

Method	Endmember				
	Cement	Crop	Water	Soil	Average
VCA-FCLS	0.0872	0.0385	0.1860	0.3123	0.1560
MVCNMF	0.1868	0.0613	0.2909	0.3086	0.2119
$L_{1/2}NMF$	0.1089	0.0448	0.1985	0.3048	0.1642
ASSNMF	<b>0.0841</b>	0.0348	0.1847	0.3189	0.1556
SGSNMF	0.0890	<b>0.0247</b>	<b>0.1615</b>	<b>0.1988</b>	<b>0.1185</b>

MATLAB 2016a. In this experiment, all the algorithms were compared on a personal computer equipped with an Intel Xeon CPU at 3.50 GHz and 16-GB RAM. The efficiency comparison with the two real hyperspectral data sets is provided

in Table VI, where we present the average time (in seconds), the number of iterations, and the objective values, which are equal to  $\|\mathbf{X} - \mathbf{AS}\|_F^2$ . The running time of SGSNMF is divided into two parts: the first part is the time for the SGSNMF decomposition and the time in brackets is the time for the SLIC segmentation.

As shown in Table VI, because of the alternative iterative process, the running times of the NMF-based algorithms are large, compared with VCA-FCLS, but the accuracy errors of the reconstruction are much lower. The optimization rules of the NMF-based algorithms are different: a multiplicative update rule is used in ASSNMF and  $L_{1/2}NMF$ , while the ANLS-projected gradient method is used in MVCNMF and SGSNMF. The results show that the algorithms with ANLS-projected gradient optimization have the merits of

TABLE VI  
EFFICIENCY COMPARISON FOR THE REAL HYPERSPECTRAL DATA

<i>Method</i>	<i>Cuprite</i>			<i>WHU-GX UAV</i>		
	<i>Running time (s)</i>	<i>Iterations</i>	<i>Objective value</i>	<i>Running time</i>	<i>Iterations</i>	<i>Objective value</i>
<i>VCA-FCLS</i>	<b>65.77</b>	-	1016.05	<b>68.16</b>	-	76951.59
<i>MVCNMF</i>	<b>169.19</b>	<b>26.3</b>	<b>256.14</b>	213.26	17.2	<b>38669.61</b>
<i>L1/2NMF</i>	<b>195.46</b>	791.5	749.90	299.14	273.3	40362.67
<i>ASSNMF</i>	<b>368.67</b>	127.6	810.74	<b>137.56</b>	121.6	56783.46
<i>SGSNMF</i>	72.908 (3.74)	36	371.90	83.84 (24.17)	<b>16.2</b>	40121.06

quick convergence and high reconstruction precision. Due to the low convergence rate of the multiplicative update rule, the number of iterations of ASSNMF and  $L_{1/2}$ NMF is greater. In contrast, the proposed SGSNMF algorithm has the fastest final convergence. Note that, due to the SLIC segmentation preprocessing, for SGSNMF, the overall running time may be larger than that of the other methods, particularly in the case of the image size being large. This is mainly because the running time of the modified SLIC is linear with regard to the number of pixels in the scene, which can be verified by comparing the SLIC running times of hyperspectral data with a different numbers of pixels.

## VI. CONCLUSION

In this paper, we have developed a novel blind unmixing algorithm, namely SGSNMF, for remotely sensed hyperspectral data. SLIC segmentation is introduced to the hyperspectral data to generate spatial groups, and the initial endmembers which are closer to the global minimum are estimated based on this spatial prior. In addition, the sparse structures of the spatial groups are exploited to describe the local distribution of the endmembers. The spatial prior and the sparsity of the abundance are integrated as a new regularization into the unmixing process, where the pixels within a spatial group are expected to share the same sparse pattern. The experimental results obtained with simulated data as well as real hyperspectral data with different spatial resolutions confirm the potential of the proposed algorithm. The proposed algorithm has several benefits when compared to some of the current state-of-the-art methods. In particular, SGSNMF is more robust to noise, as a result of the spatial prior, and performs better than the other methods as the number of endmembers increases.

In practical applications, given hyperspectral data and the number of endmembers  $M$ , the proposed blind HU method, SGSNMF, is fully automated. SGSNMF is not greatly dependent on the parameters, and the experimental results show that a large range of values for some of the important parameters, such as  $\lambda$  and  $w$ , give realistic endmember and abundance estimates. If a higher accuracy is needed, these parameters can be adjusted by the user referring to the given range and the original data. A relatively small parameter  $\lambda = 0.3$  is suggested to weight the proposed regularizer, while in a case with low SNR, it could be larger. For the SLIC segmentation, we suggest that the average superpixel size  $w$  and weight  $w_s$

are set as small values, to guarantee the homogeneous assumption of the spatial groups. In our future work, in addition to the application of the group sparsity idea to other data sets, we plan to further exploit the framework by considering the connections and differences between the different spatial groups. In addition, adaptive parameter estimation will also be studied to promote the practicality of the proposed method.

## ACKNOWLEDGMENT

The authors would like to thank the editor, associate editor, and anonymous reviewers for their helpful comments and suggestions to improve this paper, the research group supervised by Prof. J. M. Bioucas-Dias and Prof. A. Plaza for sharing their simulated data for comparison purposes, and the NBL Imaging System Ltd., for its assist during the UAV data collection.

## REFERENCES

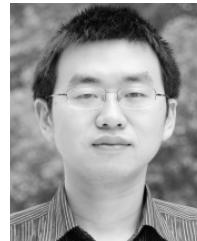
- [1] J. M. Bioucas-Dias, A. Plaza, G. Camps-Valls, P. Scheunders, N. M. Nasrabadi, and J. Chanussot, "Hyperspectral remote sensing data analysis and future challenges," *IEEE Geosci. Remote Sens. Mag.*, vol. 1, no. 2, pp. 6–36, Jun. 2013.
- [2] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.
- [3] J. M. Bioucas-Dias *et al.*, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.
- [4] J. W. Boardman, F. A. Kruse, and R. O. Green, "Mapping target signatures via partial unmixing of AVIRIS data," in *Proc. 5th Annu. JPL Airborne Earth Sci. Workshop*, vol. 1. Pasadena, CA, USA, 1995, pp. 23–26.
- [5] M. E. Winter, "N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data," *Proc. SPIE*, vol. 3753, pp. 266–275, Oct. 1999.
- [6] J. M. P. Nascimento and J. M. Bioucas-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.
- [7] C.-I. Chang, C.-C. Wu, W.-M. Liu, and Y.-C. Ouyang, "A new growing method for simplex-based endmember extraction algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 10, pp. 2804–2819, Oct. 2006.
- [8] T.-H. Chan, C.-Y. Chi, Y.-M. Huang, and W.-K. Ma, "A convex analysis-based minimum-volume enclosing simplex algorithm for hyperspectral unmixing," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4418–4432, Nov. 2009.
- [9] D. C. Heinz and C.-I. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, pp. 529–545, Mar. 2001.
- [10] J. M. P. Nascimento and J. M. B. Dias, "Does independent component analysis play a role in unmixing hyperspectral data?" *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 1, pp. 175–187, Jan. 2005.
- [11] J. Wang and C.-I. Chang, "Applications of independent component analysis in endmember extraction and abundance quantification for hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 9, pp. 2601–2616, Sep. 2006.
- [12] D. D. Lee and H. S. Seung, "Algorithms for non-negative matrix factorization," in *Proc. Adv. Neural Inf. Process. Syst.*, 2001, pp. 556–562.
- [13] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 3, pp. 765–777, Mar. 2007.
- [14] M. Berman, H. Kiiveri, R. Lagerstrom, A. Ernst, R. Dunne, and J. F. Huntington, "ICE: A statistical approach to identifying endmembers in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 10, pp. 2085–2095, Oct. 2004.
- [15] A. Zare and P. Gader, "Sparsity promoting iterated constrained end-member detection in hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 4, no. 3, pp. 446–450, Jul. 2007.

- [16] A. S. Charles, B. A. Olshausen, and C. J. Rozell, "Learning sparse codes for hyperspectral imagery," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 5, pp. 963–978, Sep. 2011.
- [17] Y. Zhong, X. Wang, L. Zhao, R. Feng, L. Zhang, and Y. Xu, "Blind spectral unmixing based on sparse component analysis for hyperspectral remote sensing imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 119, pp. 49–63, Sep. 2016.
- [18] Y. Qian, S. Jia, J. Zhou, and A. Robles-Kelly, "Hyperspectral unmixing via  $L_{1/2}$  sparsity-constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4282–4297, Nov. 2011.
- [19] X. Lu, H. Wu, Y. Yuan, P. Yan, and X. Li, "Manifold regularized sparse NMF for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 5, pp. 2815–2826, May 2013.
- [20] X. Lu, H. Wu, and Y. Yuan, "Double constrained NMF for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2746–2758, May 2014.
- [21] D. A. Roberts, M. Gardner, R. Church, S. Ustin, G. Scheer, and R. O. Green, "Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models," *Remote Sens. Environ.*, vol. 65, no. 3, pp. 267–279, Sep. 1998.
- [22] 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.
- [23] A. Zare and K. C. 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.
- [24] C. Shi and L. Wang, "Incorporating spatial information in spectral unmixing: A review," *Remote Sens. Environ.*, vol. 149, pp. 70–87, Jun. 2014.
- [25] A. Plaza, P. Martinez, R. Perez, and J. Plaza, "Spatial/spectral endmember extraction by multidimensional morphological operations," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 9, pp. 2025–2041, Sep. 2002.
- [26] S. Mei, M. He, Z. Wang, and D. Feng, "Spatial purity based endmember extraction for spectral mixture analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 9, pp. 3434–3445, Sep. 2010.
- [27] M. Zortea and A. Plaza, "Spatial preprocessing for endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2679–2693, Aug. 2009.
- [28] 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, Nov. 2012.
- [29] X. Liu, W. Xia, B. Wang, and L. Zhang, "An approach based on constrained nonnegative matrix factorization to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 757–772, Feb. 2011.
- [30] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.
- [31] C.-J. Lin, "Projected gradient methods for nonnegative matrix factorization," *Neural Comput.*, vol. 19, no. 10, pp. 2756–2779, 2007.
- [32] W.-K. Ma et al., "A signal processing perspective on hyperspectral unmixing: Insights from remote sensing," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 67–81, Jan. 2014.
- [33] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Statist. Soc. B (Statist. Methodol.)*, vol. 58, no. 1, pp. 267–288, 1996.
- [34] J. Kim, R. D. C. Monteiro, and H. Park, "Group sparsity in nonnegative matrix factorization," in *Proc. SIAM Int. Conf. Data Mining*, 2012, pp. 851–862.
- [35] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *J. Roy. Statist. Soc. B (Statist. Methodol.)*, vol. 68, no. 1, pp. 49–67, 2006.
- [36] L. Meier, S. Van De Geer, and P. Bühlmann, "The group lasso for logistic regression," *J. Roy. Statist. Soc. B (Statist. Methodol.)*, vol. 70, no. 1, pp. 53–71, 2008.
- [37] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Collaborative nonnegative matrix factorization for remotely sensed hyperspectral unmixing," in *Proc. Int. Geosci. Remote Sens. Symp. (IGARSS)*, 2012, pp. 3078–3081.
- [38] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *Int. J. Comput. Vis.*, vol. 59, no. 2, pp. 167–181, Sep. 2004.
- [39] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, pp. 603–619, May 2002.
- [40] E. Luczak and A. Rosenfeld, "Distance on a hexagonal grid," *IEEE Trans. Comput.*, vol. C-25, no. 5, pp. 532–533, May 1976.
- [41] M. Kowalski, "Sparse regression using mixed norms," *Appl. Comput. Harmon. Anal.*, vol. 27, no. 3, pp. 303–324, 2009.
- [42] F. Bach, R. Jenatton, J. Mairal, and G. Obozinski, "Convex optimization with sparsity-inducing norms," in *Optimization for Machine Learning*, vol. 5. Cambridge, MA, USA: MIT Press, 2011.
- [43] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MD, USA: Athena Scientific, 1999.
- [44] D. R. Thompson, L. Mandrake, M. S. Gilmore, and R. Castano, "Superpixel endmember detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 11, pp. 4023–4033, Nov. 2010.
- [45] C.-I. Chang and Q. Du, "Estimation of number of spectrally distinct signal sources in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, pp. 608–619, Mar. 2004.
- [46] J. M. Bioucas-Dias and J. M. P. Nascimento, "Hyperspectral subspace identification," *IEEE Trans. Geosci. Remote Sensing*, vol. 46, no. 8, pp. 2435–2445, Aug. 2008.
- [47] E. M. T. Hendrix, I. Garcia, J. Plaza, G. Martin, and A. Plaza, "A new minimum-volume enclosing algorithm for endmember identification and abundance estimation in hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 7, pp. 2744–2757, Jul. 2012.
- [48] J. A. Hartigan and M. A. Wong, "Algorithm AS 136: A K-means clustering algorithm," *J. Roy. Statist. Soc. C (Appl. Statist.)*, vol. 28, no. 1, pp. 100–108, 1979.
- [49] G. A. Swayze, "The hydrothermal and structural history of the Cuprite mining district, southwestern Nevada: An integrated geological and geochemical approach," Ph.D. dissertation, Dept. Geol. Sci., Univ. Colorado Boulder, Boulder, CO, USA, 1997.



**Xinyu Wang** received the B.S. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2014, where he is currently pursuing the Ph.D. degree in communication and information systems with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing.

His research interests include the blind source separation, hyperspectral unmixing, and infrared target detection.



**Yanfei Zhong** (M'11–SM'15) received the B.S. degree in information engineering and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2002 and 2007, respectively.

He has been a Teacher with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, since 2007, and was promoted to an Associate Professor and a Full Professor in 2008 and 2010, respectively. He has authored

more than 120 research papers, including more than 60 peer-reviewed articles in international journals such as the *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, the *IEEE TRANSACTIONS ON IMAGE PROCESSING*, the *IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING*, and *Pattern Recognition*. His research interests include hyperspectral remote sensing information processing, high-resolution remote sensing image understanding, and geoscience interpretation for multisource remote sensing data.

Dr. Zhong was a recipient of the Excellent Young Scientist Foundation of China, the National Excellent Doctoral Dissertation Award of China, and the New Century Excellent Talents in the University of China. He was also a referee of more than 30 international journals. He is an Associate Editor of the *IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING*, the *International Journal of Remote Sensing*, and *Remote Sensing*.



**Liangpei Zhang** (M'06–SM'08) received the B.S. degree in physics from Hunan Normal University, Changsha, China, in 1982, the M.S. degree in optics from the Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, China, in 1988, and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 1998.

He is currently the Head of the Remote Sensing Division, State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, he is also a “Chang-Jiang Scholar” Chair Professor appointed by the Ministry of Education of China, Beijing, China. He was a Principal Scientist for the China State Key Basic Research Project during 2011–2016 appointed by the Ministry of National Science and Technology of China, Beijing, to lead the remote sensing program in China. He has authored more than 450 research papers and five books, and holds 15 patents. He has edited several conference proceedings, issues, and geoinformatics symposiums. His research interests include hyperspectral remote sensing, high-resolution remote sensing, image processing, and artificial intelligence.

Dr. Zhang is a fellow of the Institution of Engineering and Technology. He received the 2010 Best Paper Boeing Award and the 2013 Best Paper ERDAS Award from the American Society of Photogrammetry and Remote Sensing. He regularly serves as a Co-Chair of the series SPIE conferences on multispectral image processing and pattern recognition, conference on Asia remote sensing, and many other conferences. He also serves as an Associate Editor of the *International Journal of Ambient Computing and Intelligence*, the *International Journal of Image and Graphics*, the *International Journal of Digital Multimedia Broadcasting*, the *Journal of Geospatial Information Science*, and the *Journal of Remote Sensing*, and a Guest Editor of the *Journal of Applied Remote Sensing* and the *Journal of Sensors*. He is currently serving as an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.



**Yanyan Xu** received the B.S. degree in mechanics from the Southwest University of Science and Technology, Mianyang, China, in 1995, the M.S. degree in electronics from the Hubei University of Technology, Wuhan, China, in 2000, and the Ph.D. degree in communication and information system from Wuhan University, Wuhan, in 2007.

She is currently a Professor with the State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University. She has authored more than 40 research papers and one book. Her research interests include multimedia communication systems and information security.