Spectral–Spatial Weighted Sparse Regression for Hyperspectral Image Unmixing

Shaoquan Zhang, *Student Member, IEEE*, Jun Li, *Senior Member, IEEE*, Heng-Chao Li, *Senior Member, IEEE*, Chengzhi Deng, and Antonio Plaza, *Fellow, IEEE*

***Abstract*— Spectral unmixing aims at estimating the fractional abundances of a set of pure spectral materials (endmembers) in each pixel of a hyperspectral image. The wide availability of large spectral libraries has fostered the role of sparse regression techniques in the task of characterizing mixed pixels in remotely sensed hyperspectral images. A general solution for sparse unmixing methods consists of using the ℓ_1 regularizer to control the sparsity, resulting in a very promising performance but also suffering from sensitivity to large and small sparse coefficients. A recent trend to address this issue is to introduce weighting factors to penalize the nonzero coefficients in the unmixing solution. While most methods for this purpose focus on analyzing the hyperspectral data by considering the pixels as independent entities, it is known that there exists a strong spatial correlation among features in hyperspectral images. This information can be naturally exploited in order to improve the representation of pixels in the scene. In order to take advantage of the spatial information for hyperspectral unmixing, in this paper, we develop a new spectral–spatial weighted sparse unmixing (S^2WSU) framework, which uses both spectral and spatial weighting factors, further imposing sparsity on the solution. Our experimental results, conducted using both simulated and real hyperspectral data sets, illustrate the good potential of the proposed S^2WSU, which can greatly improve the abundance estimation results when compared with other advanced spectral unmixing methods.**

***Index Terms*— Hyperspectral imaging, sparse unmixing, spatial information, spatially weighted unmixing.**

Manuscript received August 22, 2017; revised December 7, 2017 and January 10, 2018; accepted January 13, 2018. Date of publication February 9, 2018; date of current version May 21, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 61771496, Grant 61371165, and Grant 61362036, in part by the National Key Research and Development Program of China under Grant 2017YFB0502900, in part by the Guangdong Provincial Natural Science Foundation under Grant 2016A030313254, in part by the Open Research Fund of Jiangxi Province Key Laboratory of Water Information Cooperative Sensing and Intelligent Processing under Grant 2016WICSIP010, and in part by the Natural Science Foundation of Jiangxi China under Grant 20161BAB202040. *(Corresponding authors: Jun Li; Heng-Chao Li.)*

S. Zhang and J. Li are with the Guangdong Provincial Key Laboratory of Urbanization and Geo-Simulation, Center of Integrated Geographic Information Analysis, School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China (e-mail: lijun48@mail.sysu.edu.cn).

H.-C. Li is with the Sichuan Provincial Key Laboratory of Information Coding and Transmission, Southwest Jiaotong University, Chengdu 610031, China (e-mail: lihengchao_78@163.com).

C. Deng is with the Jiangxi Province Key Laboratory of Water Information Cooperative Sensing and Intelligent Processing, Department of Information Engineering, Nanchang Institute of Technology, Nanchang 330099, China.

A. Plaza is with the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, Escuela Politécnica, University of Extremadura, E-10071 Cáceres, Spain.

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.2018.2797200

I. Introduction

With the rapid development of the hyperspectral technology, hyperspectral remote sensing images have been widely used in environmental surveillance, target detection, mineral exploration, and so on [1]. However, due to insufficient spatial resolution and spatial complexity, pixels in remotely sensed hyperspectral images are likely to be formed by a mixture of pure spectral constituents (*endmembers*) rather than a single substance [2]. The existence of mixed pixels complicates the exploitation of hyperspectral images [3]. Spectral unmixing, aimed at estimating the fractional abundance of the pure spectral signatures or endmembers, was proposed to deal with the problem of spectral mixing and effectively identifies the components of the mixed spectra in each pixel [4].

In the past few years, linear spectral unmixing has been one of the most active research lines in hyperspectral imaging [2]. Depending on whether a spectral library is available or not, researchers have developed many different unsupervised and semisupervised unmixing algorithms. Unsupervised algorithms that extract the endmembers directly from the scene have faced several difficulties, mainly related with the unavailability of pure signatures in the image data [5]–[8]. Furthermore, endmember generation algorithms tend to obtain virtual endmembers without physical meaning [9]–[14]. In addition, these algorithms generally require the estimation of the number of endmembers in the scene, which is also a difficult problem.

Recently, due to the wide availability of spectral libraries, sparse unmixing [15] has been shown to be able to circumvent the drawbacks introduced by such *virtual* endmembers and the unavailability of pure pixels. Sparse unmixing is a semisupervised approach in which mixed pixels are expressed in the form of combinations of a number of pure spectral signatures from a large spectral library. The sparse unmixing approach exhibits significant advantages over unsupervised approaches, as it does not need to extract endmembers from the hyperspectral data or estimate the number of the endmembers. Another advantage of sparse unmixing is that it provides great potential for the accurate estimation of the fractional abundances, as all endmembers are normally represented in the library. These new perspectives introduced by sparse unmixing fostered advanced developments in the field [16]–[22].

The success of sparse unmixing relies on the fact that the unmixing solution is sparse, as the number of endmembers used to represent a mixed pixel is generally much smaller

than the number of spectral signatures in the library [15]. As a result, new algorithms have been developed to enforce the sparsity on the solution. The sparse unmixing algorithm via variable splitting and augmented Lagrangian (SUnSAL) [15] adopts the ℓ_1 regularizer on the abundance matrix, which aims at introducing sparsity through the spectral domain that a pixel is unlikely to be mixed by a high number of components. The introduction of SUnSAL opened new avenues and brought new insights into the concept of sparse unmixing. However, the real degree of sparsity is beyond the reach of the ℓ_1 regularizer due to the unbalance between the number of endmembers in the library and the number of components that generally participate in a mixed pixel. New algorithms have been developed in order to perform a better characterization of the degree of sparsity. Two main directions have been pursued. On the one hand, some techniques have focused on the introduction of new orders over the sparse regularizer. For instance, the collaborative SUnSAL introduces the $\ell_{2,1}$ regularizer to impose sparsity among the endmembers simultaneously (collaboratively) for all pixels [23]. In [24], the $\ell_{1/2}$ regularizer was introduced into nonnegative matrix factorization (NMF) for unmixing so as to enforce the sparsity of endmember abundances. In [25], the graph-regularized $\ell_{1/2}$-NMF was proposed by introducing the manifold regularization into the sparsity-constrained NMF, leading to a more desirable result than other sparse NMF algorithms. In [26], the ℓ_p $(0 < p < 1)$ regularizer was used. On the other hand, other algorithms have introduced weighting factors to penalize the nonzero coefficients on the sparse solution [27]. For example, in [28], a weighted ℓ_1 minimization is adopted to enhance the sparsity of the solution. In the double reweighted sparse unmixing (DRSU) [29] algorithm, double weights are used to improve the sparsity of the endmembers in the spectral library and also the sparsity of the abundance fractions of each endmember. Although these methods obtained promising results, they consider pixels in a hyperspectral data as independent entities, while the spatial-contextual information in the hyperspectral image is generally disregarded. Since hyperspectral images generally follow specific spatial arrangements by nature, it is important to consider spatial information for their characterization.

In fact, hyperspectral images exhibit rich spatial correlation that can be exploited for better estimating endmember abundances, thus making the obtained abundances more consistent [30]. Following this observation, several algorithms have focused on promoting the spatial correlation on the final solution. For instance, the sparse unmixing via variable splitting augmented Lagrangian and total variation (SUnSAL-TV) [31] represents one of the first attempts to include spatial information in sparse unmixing. It exploits the spatial information via a first-order pixel neighborhood system. Similar to SUnSAL, the SUnSAL-TV opened new avenues and brought new insights into the concept of spatial sparse unmixing, which is able to promote piecewise transitions in the estimated abundances. However, its performance strongly relies on the parameter settings [32]. At the same time, its model complexity results in a heavy computational cost, further limiting its practical application potential. New developments aimed at fully exploiting the spatial correlation among image features (and further imposing sparsity on the abundance matrix) and have been mainly developed along two directions. On the one hand, high-order neighborhood information over spatial regularizers has been introduced to reach this goal. For instance, nonlocal sparse unmixing (NLSU) adopts the nonlocal means as a spatial regularizer for sparse unmixing, exploiting similar patterns and structures in the abundance image [33]. Although NLSU can take advantage of high-order structural information, the neighborhood of the pixel changes randomly, thus limiting the continuity of spectral information. Another drawback of NLSU is that its model is more complex than that of SUnSAL-TV, which limits its practical application. On the other hand, spatially weighted factors (aimed at characterizing spatial information through the inclusion of a weight on the sparse regularizer) have also been used to account for the spatial information in sparse unmixing. For example, the local collaborative sparse unmixing (LCSU) uses a spatial weight to impose local collaborativity, thus addressing some of the issues observed in SUnSAL-TV (including oversmoothed boundaries and blurred abundance maps) [34]. With similar complexity as the SUnSAL-TV (mainly because the spatial weight can be precomputed beforehand), the LCSU exhibits similar unmixing performance as the SUnSAL-TV. This indicates the good potential of using spatial weights (as compared with spatial regularizers) in two main aspects: good unmixing performance and low computational complexity.

Inspired by the advantages observed by the use of spatial weights in LCSU, in this paper, we develop a new spectral–spatial weighted sparse unmixing (S^2WSU) framework for hyperspectral unmixing. The proposed S^2WSU aims at obtaining a sparse solution that is constrained simultaneously from the spectral and spatial domains under the ℓ_1 framework. The spectral weight, following previous developments [27], enforces the sparsity of nonzero rows corresponding to the true endmembers in the estimated abundances. On the other hand, the spatial weight incorporates the spatial correlation to introduce sparsity along a neighborhood system (which represents an entirely original contribution of this paper). An important aspect of our newly proposed S^2WSU framework is that, with its open structure, it is able to accept multiple types of spectral and spatial weighting factors, thus providing great flexibility for the exploration of different spatial scenarios, such as edge information, nonlocal similarity, and homogeneous neighborhood information. This is one of the main contributions of this paper. With respect to its computational complexity, the proposed S^2WSU can be efficiency solved via the alternating direction method of multipliers (ADMM) [35], with complexity levels linearly as the well-known SUnSAL method. It should be finally noted that, similar to the TV-based methods, the proposed approach also promotes piecewise transitions in the estimated abundances, however, with two main advantages. First of all, the proposed approach only needs one regularization parameter. Furthermore, the proposed approach exhibits much lower computational complexity.

The remainder of this paper is organized as follows. The proposed spectral–spatial weighted sparse unmixing model is presented in Section II. Section III describes our experimental results with simulated hyperspectral data sets. Section IV

describes our experiments with real hyperspectral data sets. Finally, Section V draws some conclusions and hints at plausible future research lines.

II. Proposed Spectral–Spatial Weighted Sparse Unmixing Model

Let $\mathbf{Y} = [\mathbf{y}_1, \ldots, \mathbf{y}_n] \in \mathbb{R}^{d\times n}$ denote a hyperspectral image, where n is the number of pixel vectors and d is the number of bands. Let $\mathbf{A} \in \mathbb{R}^{d\times m}$ be a large spectral library, where m is the number of spectral signatures in $\mathbf{A}$, and $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_n]$ denotes the abundance maps corresponding to library $\mathbf{A}$ for the observed data $\mathbf{Y}$. With the aforementioned definitions in mind, sparse unmixing finds a linear combination of endmembers for $\mathbf{Y}$ from the spectral library $\mathbf{A}$

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{N} \quad \text{s.t.: } \mathbf{X} \geq 0 \tag{1}$$

where $\mathbf{N} \in \mathbb{R}^{d\times n}$ is the error, and $\mathbf{X} \geq 0$ is the so-called abundance nonnegativity constraint (ANC). It should be noted that we explicitly enforce the ANC constraint without the abundance sum-to-one constraint (ASC), due to some criticisms about the ASC in the literature [15].

As the number of endmembers involved in a mixed pixel is usually very small when compared with the size of the spectral library, the abundance matrix $\mathbf{X}$ is sparse. With these considerations in mind, the unmixing problem can be formulated as an $\ell_2 - \ell_0$ optimization problem

$$\min_{\mathbf{X}} \frac{1}{2}||\mathbf{A}\mathbf{X} - \mathbf{Y}||_F^2 + \lambda||\mathbf{X}||_0 \quad \text{s.t.: } \mathbf{X} \geq 0 \tag{2}$$

where $\|\cdot\|_F$ is the Frobenius norm, and λ is a regularization parameter. Problem (2) is nonconvex and difficult to solve [36], [37]. The SUnSAL alternatively uses the $\ell_2 - \ell_1$ norm to replace the $\ell_2 - \ell_0$ norm and solves the unmixing problem as follows [35]:

$$\min_{\mathbf{X}} \frac{1}{2}||\mathbf{A}\mathbf{X} - \mathbf{Y}||_F^2 + \lambda||\mathbf{X}||_{1,1} \quad \text{s.t.: } \mathbf{X} \geq 0 \tag{3}$$

where $||\mathbf{X}||_{1,1} = \sum_{j=1}^{n} ||\mathbf{x}_j||_1$ with $\mathbf{x}_j$ being the jth column of $\mathbf{X}$. The SUnSAL solves the optimization problem in (3) efficiently using the ADMM [35]. However, as stated before, the real degree of sparsity is generally beyond the reach of the ℓ_1 regularizer.

A. Double Reweighted Sparse Regression and Total Variation

Inspired by the success of weighted ℓ_1 minimization in sparse signal recovery, the DRSU and total variation (DRSU-TV) [38] was proposed to simultaneously exploit the spectral dual sparsity as well as the spatial smoothness of fractional abundances as follows:

$$\min_{\mathbf{X}} \frac{1}{2}||\mathbf{A}\mathbf{X} - \mathbf{Y}||_F^2 + \lambda||(\mathbf{W}_2\mathbf{W}_1) \odot \mathbf{X}||_{1,1} + \lambda_{\text{TV}}\text{TV}(\mathbf{X}), \quad \text{s.t.: } \mathbf{X} \geq 0 \tag{4}$$

where the operator $\odot$ denotes the elementwise multiplication of two variables. On the one hand, the first regularizer $\lambda||(\mathbf{W}_2\mathbf{W}_1) \odot \mathbf{X}||_{1,1}$ introduces a prior with spectral sparsity, where λ is the regularization parameter, $\mathbf{W}_1 = \{w_{1,ij} | i = 1, \ldots, m, j = 1, \ldots, n\} \in \mathbb{R}^{m\times n}$ and $\mathbf{W}_2 = \text{diag}(w_{2,11}, \ldots, w_{2,ii}, \ldots, w_{2,mm}) \in \mathbb{R}^{m\times m}$, for $i = 1, \ldots, m$, are the dual weights, with $\mathbf{W}_1$ being the original weight introduced in [27] aimed at penalizing the nonzero coefficients on the solution and $\mathbf{W}_2$ promoting nonzero row vectors. On the other hand, the later regularizer $\lambda_{\text{TV}}\text{TV}(\mathbf{X})$ exploits the spatial prior with λ_{TV} being the parameter controlling the degree of smoothness, where $\text{TV}(\mathbf{X}) \equiv \sum_{p,k\in\mathcal{N}} ||\mathbf{x}_p - \mathbf{x}_k||_1$, $\mathcal{N}$ represents the set of (horizontal and vertical) pixel neighbors in the image. It can be seen that DRSU-TV incorporates a TV-based regularizer to enforce the spatial smoothness of abundances compared with DRSU.

In [38], Problem (4) is optimized via the ADMM under an iterative scheme. The dual weights $\mathbf{W}_1$ and $\mathbf{W}_2$ are updated as follows, at iteration $t + 1$:

$$w_{1,ij}^{t+1} = \frac{1}{x_{ij}^t + \varepsilon} \tag{5}$$

where $\varepsilon > 0$ is a small positive value and

$$w_{2,ii}^{t+1} = \frac{1}{||\mathbf{X}^t(i,:)||_2 + \varepsilon} \tag{6}$$

where $\mathbf{X}^t(i,:)$ is the ith row in the estimated abundance of the tth iteration. Notice that, as shown in (5) and (6), it is suggested that large weights be used to discourage nonzero entries in the recovered signal, whereas small weights encourage nonzero entries. DRSU-TV, exploiting the spectral and spatial priors simultaneously under the sparse unmixing model, exhibits good potential in comparison with the ℓ_1- or TV-based methods. However, as an adaptation of the ℓ_1- and TV-based approach, the limitations of DRSU-TV are associated with the use of a regularizer-based spatial prior. That is, the computational complexity is similar to that of SUnSAL-TV. Such high computational complexity constrains the practical applications of DRSU-TV. Furthermore, the unmixing performance of the method is sensitive to the regularization parameter λ_{TV}.

B. Spectral–Spatial Weighted Sparse Unmixing

In this paper, aiming at exploiting the spatial information more efficiently for sparse unmixing purposes (and inspired by the generally good behavior of the weighted ℓ_1 optimization), we develop a new method called S^2WSU for hyperspectral unmixing. As opposed to the approaches that exploit a regularizer-based spatial prior (which have one additional parameter for the spatial regularizer and often exhibit high complexity), the proposed S^2WSU includes the spatial correlation via a weighting factor, resulting in good computational efficiency and less regularization parameters. Let $\mathbf{W}_{\text{spe}} \in \mathbb{R}^{m\times m}$ be the spectral weighting matrix and $\mathbf{W}_{\text{spa}} \in \mathbb{R}^{m\times n}$ be the spatial one. Following [27], the objective function of the proposed S^2WSU is given as follows:

$$\min_{\mathbf{X}} \frac{1}{2}||\mathbf{A}\mathbf{X} - \mathbf{Y}||_F^2 + \lambda||(\mathbf{W}_{\text{spe}}\mathbf{W}_{\text{spa}}) \odot \mathbf{X}||_{1,1}, \quad \text{s.t.: } \mathbf{X} \geq 0. \tag{7}$$

For the spectral weighting factor $\mathbf{W}_{\text{spe}}$, relying on the success of [23] and [29], we adopt row collaborativity to enforce joint sparsity among all the pixels. Similar to $\mathbf{W}_2$

in DRSU-TV, $\mathbf{W}_{\text{spe}}$ aims at enhancing the sparsity of the endmembers in the spectral library. In detail, at iteration $t+1$, it can be updated as

$$\mathbf{W}_{\text{spe}}^{t+1} = \text{diag}\left[\frac{1}{||\mathbf{X}^t(1,:)||_2+\varepsilon}, \ldots, \frac{1}{||\mathbf{X}^t(m,:)||_2+\varepsilon}\right]. \quad (8)$$

As mentioned before, due to the spatial arrangement of the pixels in an image, their corresponding abundances often show a pronounced spatial dependence. In other words, the rich spatial information that the image contains can be of great importance for estimating endmember abundances and also making the obtained abundances more consistent. With this information in mind, for the spatial weighting factor $\mathbf{W}_{\text{spa}}$, let $w_{\text{spa},ij}^{t+1}$ be the element of the ith line and jth row in $\mathbf{W}_{\text{spa}}$ at iteration $t+1$, we incorporate the neighboring information as follows:

$$w_{\text{spa},ij}^{t+1} = \frac{1}{f_{h\in\mathcal{N}(j)}(x_{ih}^t)+\varepsilon} \quad (9)$$

where $\mathcal{N}(j)$ denotes the neighboring set for element x_{ij}, and $f(\cdot)$ is a function explicitly exploiting the spatial correlations through the neighborhood system. Notice that the function $f(\cdot)$ can be linear or nonlinear; for example, the neighborhood coverage [39], the nonlocal similarity [40], the edge information representing the spatial discontinuity in [41], and so on. Specifically, in this paper, we use the neighboring coverage and importance to incorporate the spatial correlation as follows:

$$f(x_{ij}) = \frac{\sum_{h\in\mathcal{N}(j)} \epsilon_{ih} x_{ih}}{\sum_{h\in\mathcal{N}(j)} \epsilon_{ih}} \quad (10)$$

where $\mathcal{N}(j)$ corresponds to the neighboring coverage and ϵ represents the neighborhood importance. As mentioned before, the neighborhood coverage can be expressed in many forms, such as superpixel-based [42], first-order-based [43], and clustering-based [44]. For simplicity, we consider the 8-connected (3×3 window) and 24-connected (5×5 window) for algorithm design and experiments. They are defined as S^{2}WSU-W1 and S^{2}WSU-W2, respectively. With respect to the neighboring importance, for any two entries i and j, we compute it as follows:

$$\epsilon_{ij} = \frac{1}{\text{im}(i,j)} \quad (11)$$

where function im$(\cdot)$ is the importance measurement over the two elements x_i and x_j. Let (a,b) and (c,d) be the spatial coordinates of x_i and x_j. The European distance is specifically considered, that is, $\epsilon_{ij} = 1/((a-c)^2+(b-d)^2)^{1/2}$.

It should be noted that, although in [29] and [38], the weighted ℓ_1 problem is optimized via the ADMM with very promising performance, the plug of the weights brings difficulty with respect to the convergence (i.e., there is no evidence that the ADMM can converge in this scenario). Nevertheless, we can perform an outer–inner looping scheme to iteratively solve this problem, as indicated in (7), with the inner loop corresponding to the update of the coefficients $\mathbf{X}$ via the ADMM and the outer loop corresponding to the update of the two weights, respectively. Specifically, for each inner loop, given $\mathbf{W}_{\text{spe}}$ and $\mathbf{W}_{\text{spa}}$, the coefficients $\mathbf{X}$ can be optimized via the ADMM as indicated in (3). In the following, we revisit the algorithm in detail to account for these enhancements.

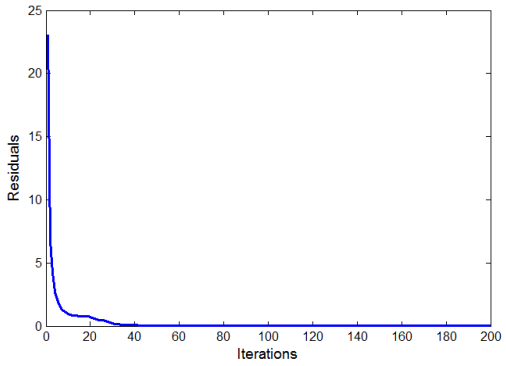


Fig. 1. Residual $||\mathbf{GU}^{(t)}+\mathbf{BV}^{(t)}||_F$ as a function of the number of iterations for the complete algorithm.

Let $\mathcal{L}(\mathbf{U},\mathbf{V},\ \mathbf{D}) \equiv g(\mathbf{U},\mathbf{V}) + (\mu/2)\|\mathbf{GU}+\mathbf{BV}-\mathbf{D}\|_F^2$ be the augmented Lagrangian for

$$\min_{\mathbf{U},\mathbf{V}}\ g(\mathbf{V}) \quad \text{s.t. } \mathbf{GU}+\mathbf{BV}=\mathbf{0} \quad (12)$$

where $\mu > 0$ is a positive constant, $\mathbf{U} = \mathbf{X}$, and $\mathbf{D} = (\mathbf{D_1},\mathbf{D_2},\mathbf{D_3})$ denote the Lagrange multipliers associated with the constraint $\mathbf{GU}+\mathbf{BV} = \mathbf{0}$; $\mathbf{V} \equiv (\mathbf{V_1},\mathbf{V_2},\mathbf{V_3})$, $\mathbf{G} = [\mathbf{A},\mathbf{I},\mathbf{I}]^{\mathbf{T}}$, $\mathbf{B} = \text{diag}(-\mathbf{I})$; $g(\mathbf{V}) \equiv (1/2)\|\mathbf{V}_1-\mathbf{Y}\|_F^2 + \lambda\|(\mathbf{W}_{\text{spe}}\mathbf{W}_{\text{spa}})\odot\mathbf{V}_2\|_{1,1} + \iota_{R+}(\mathbf{V}_3)$ is the cost function of the following optimization problem:

$$\min_{\mathbf{U},\mathbf{V}_1,\mathbf{V}_2,\mathbf{V}_3} \frac{1}{2}\|\mathbf{V}_1-\mathbf{Y}\|_F^2+\lambda\|(\mathbf{W}_{\text{spe}}\mathbf{W}_{\text{spa}})\odot\mathbf{V}_2\|_{1,1}+\iota_{R+}(\mathbf{V}_3)$$
$$\text{s.t.: } \mathbf{V}_1=\mathbf{AU},\quad \mathbf{V}_2=\mathbf{U},\ \mathbf{V}_3=\mathbf{U} \quad (13)$$

where $\iota_{R+}(\mathbf{X}) = \sum_{i=1}^n \iota_{R+}(\mathbf{x}_i)$ is the indicator function, and $\iota_{R+}(\mathbf{x}_i)$ is zero if $\mathbf{x}_i$ belongs to the nonnegative orthant and $+\infty$ otherwise. Now, we can implement the ADMM to solve the optimization problem involved in the S^{2}WSU algorithm as shown in Algorithm 1, where soft$(\cdot,\tau)$ denotes the componentwise application of the soft-threshold function $y \mapsto \text{sign}(y)\max\{|y|-\tau,0\}$, u and d_2 denote the elements of $\mathbf{U}$ and $\mathbf{D_2}$, respectively. In the outer loop of Algorithm 1, steps 4 and 5 update the spectral and spatial weights, respectively. In the inner loop, the Lagrange multipliers are updated via the ADMM.

It should be noted that, as mentioned before, we are aware that it is difficult to justify the convergence of Algorithm 1. However, the algorithm stops when the maximum iteration number is reached or $||\mathbf{GU}^{(t)}+\mathbf{BV}^{(t)}||_F \le$ *threshold* in practice, which we have empirically found out that is quite common. Fig. 1 illustrates the obtained residual, i.e., $||\mathbf{GU}^{(t)}+\mathbf{BV}^{(t)}||_F$, as a function of the number of iterations for the complete algorithm, where the number of iterations of the inner and outer loops is set to 5 and 200, respectively. It can be observed that the full scheme results in good convergence.

III. Experiments With Synthetic Data

In this section, we illustrate the unmixing performance of the proposed spectral–spatial weighted sparse unmix-

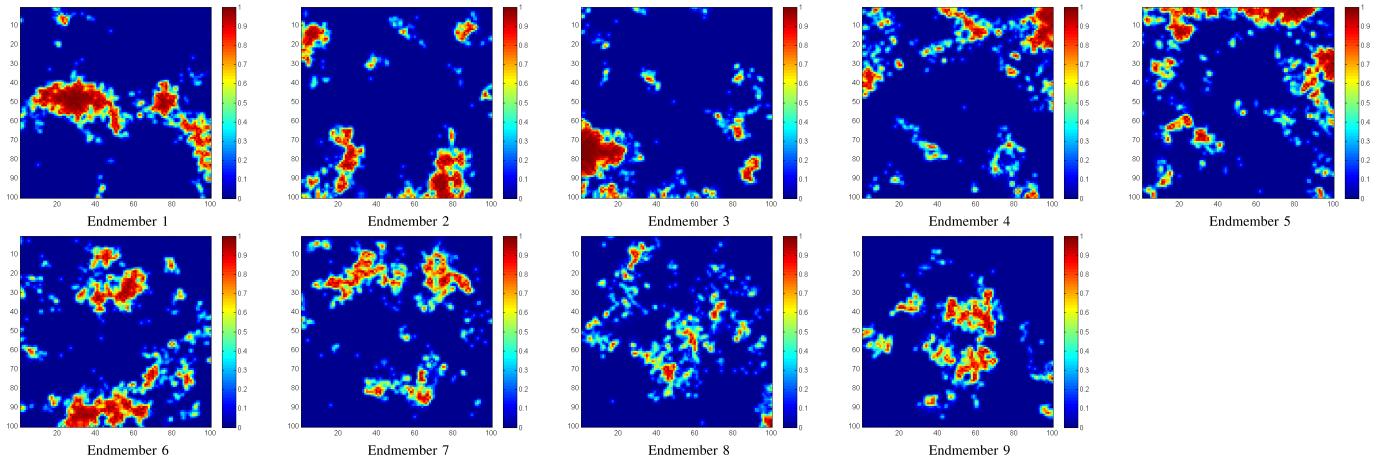


Fig. 2. True fractional abundances of the endmembers in the simulated data cube 1 (DC1).

Algorithm 1 Pseudocode of the S^{2}WSU Algorithm

Initialization:
set $k, t = 0$, choose $\mu, \lambda, \varepsilon > 0$, $\mathbf{U}^{(0)}, \mathbf{V}_1^{(0)}, \mathbf{V}_2^{(0)}, \mathbf{V}_3^{(0)}, \mathbf{D}_1^{(0)}, \mathbf{D}_2^{(0)}, \mathbf{D}_3^{(0)}$
Repeat:
$\mathbf{W}_{spe}^{(t)} \leftarrow \text{diag}\left[\frac{1}{||(\mathbf{U}^{(t)}-\mathbf{D}_2^{(t)})(1,:)||_2+\varepsilon}, \ldots, \frac{1}{||(\mathbf{U}^{(t)}-\mathbf{D}_2^{(t)})(m,:)||_2+\varepsilon}\right]$
$\mathbf{W}_{spa}^{(t)} \leftarrow \begin{bmatrix} w_{spa,11}^{(t)} & \cdots & w_{spa,1n}^{(t)} \\ \vdots & w_{spa,ij}^{(t)} & \vdots \\ w_{spa,m1}^{(t)} & \cdots & w_{spa,mn}^{(t)} \end{bmatrix}$,
where $w_{spa,ij}^{(t)} = \frac{1}{f_{h\in\mathcal{N}(j)}(u_{ih}^{(t)}-d_{2ih}^{(t)})+\varepsilon}$
Repeat:
$\mathbf{U}^{(k+1)} \leftarrow (\mathbf{A}^{\mathbf{T}}\mathbf{A}+2\mathbf{I})^{-1}(\mathbf{A}^{\mathbf{T}}(\mathbf{V}_1^{(k)}+\mathbf{D}_1^{(k)})+\mathbf{V}_2^{(k)}+\mathbf{D}_2^{(k)}+\mathbf{V}_3^{(k)}+\mathbf{D}_3^{(k)})$
$\mathbf{V}_1^{(k+1)} \leftarrow \frac{1}{1+\mu}[\mathbf{Y}+\mu(\mathbf{A}\mathbf{U}^{(k+1)}-\mathbf{D}_1^{(k)})]$
$\mathbf{V}_2^{(k+1)} \leftarrow \text{soft}(\mathbf{U}^{(k+1)}-\mathbf{D}_2^{(k)}, \frac{\lambda}{\mu}\mathbf{W}_{spe}^{(t)}\mathbf{W}_{spa}^{(t)})$
$\mathbf{V}_3^{(k+1)} \leftarrow \max(\mathbf{U}^{(k+1)}-\mathbf{D}_3^{(k)}, 0)$
Update Lagrange multipliers:
$\mathbf{D}_1^{(k+1)} \leftarrow \mathbf{D}_1^{(k)}-\mathbf{A}\mathbf{U}^{(k+1)}+\mathbf{V}_1^{(k+1)}$
$\mathbf{D}_2^{(k+1)} \leftarrow \mathbf{D}_2^{(k)}-\mathbf{U}^{(k+1)}+\mathbf{V}_2^{(k+1)}$
$\mathbf{D}_3^{(k+1)} \leftarrow \mathbf{D}_3^{(k)}-\mathbf{U}^{(k+1)}+\mathbf{V}_3^{(k+1)}$
Update iteration: $k \leftarrow k+1$
$\mathbf{U}^{(t+1)} \leftarrow \mathbf{U}^{(k+1)}$
$\mathbf{D}_2^{(t+1)} \leftarrow \mathbf{D}_2^{(k+1)}$
Update iteration: $t \leftarrow t+1$
until some stopping criterion is satisfied.

ing method using simulated hyperspectral data sets. For quantitative analysis, the signal-to-reconstruction error (SRE) (measured in dB) is used to evaluate the unmixing accuracy. For comparative purposes, the results obtained by the SUnSAL [15], SUnSAL-TV [31], DRSU [29], and DRSU-TV [38] algorithms are also reported. Let $\widehat{\mathbf{x}}$ be the estimated abundance, and $\mathbf{x}$ be the true abundance. The SRE (dB) can be computed as follows:

$$\text{SRE(dB)} = 10 \cdot \log_{10}\left(E\left(||\mathbf{x}||_2^2\right)/E\left(||\mathbf{x}-\widehat{\mathbf{x}}||_2^2\right)\right) \quad (14)$$

where $E(\cdot)$ denotes the expectation function. Furthermore, we use another indicator, i.e., the probability of success p_s, which is an estimate of the probability that the relative error power be smaller than a certain threshold. It is formally defined as follows: $p_s \equiv P(\|\widehat{\mathbf{x}}-\mathbf{x}\|^2/\|\mathbf{x}\|^2 \leq threshold)$. In our case, the estimation result is considered successfully when $\|\widehat{\mathbf{x}}-\mathbf{x}\|^2/\|\mathbf{x}\|^2 \leq 3.16$ (5 dB). This threshold was demonstrated to be appropriate in [15]. The larger the SRE (dB) or p_s is, the more accurate the unmixing is. Furthermore, a new metric called *sparsity* is also introduced for validation. Specifically, the *sparsity* measures the proportion of the number of elements in $\widehat{\mathbf{x}}$ that are greater than 0.005 with regard to all elements. It is obvious that the smaller the *sparsity* is, the more sparse the unmixing solution is.

A. Simulated Data Sets

Two spectral libraries are considered in our experiments, which are dictionaries of minerals extracted from the United States Geological Survey (USGS) library.[1] The first library $\mathbf{A_1}$ contains $m = 240$ materials (different mineral types), which comprises spectral signatures with reflectance values given in $L = 224$ spectral bands and distributed uniformly in the interval 0.4–2.5 μm. The library $\mathbf{A_2}$ contains $m = 222$ materials (different mineral types), which comprises spectral signatures with reflectance values given in $L = 221$ spectral bands and distributed uniformly in the interval 0.4–2.5 μm. Two different data sets are simulated, in which the fractional abundances are subject to the ANC and ASC.

1) *Simulated Data Cube 1 (DC1):* DC1 is generated with 100×100 pixels and nine signatures, which are randomly chosen from the spectral library $\mathbf{A_1}$. The fractional abundances are piecewise smooth, i.e., they are smooth with sharp transitions. These data can reveal the spatial features quite well for the different unmixing algorithms. For illustrative purposes, Fig. 2 shows the true abundance maps of the endmembers. After generating the datacube, it was contaminated with independent identically distributed Gaussian noise, for three levels of the signal-to-noise (SNR) ratio: 30, 40, and 50 dB.

[1] Available online at http://speclab.cr.usgs.gov/spectral.lib06.

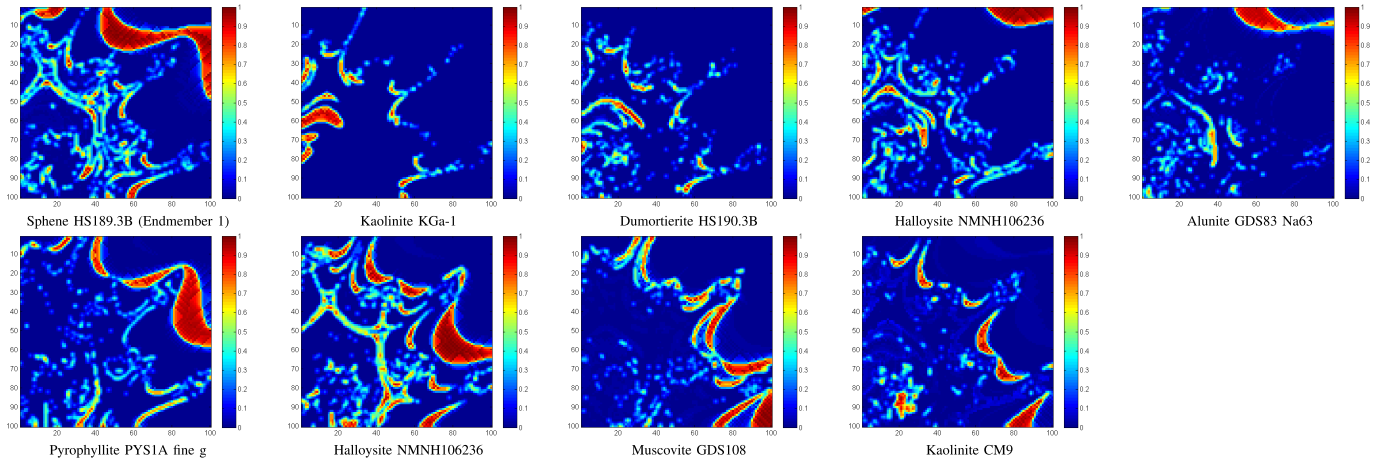


Fig. 3. True fractional abundances of the endmembers in the simulated data cube 2 (DC2).

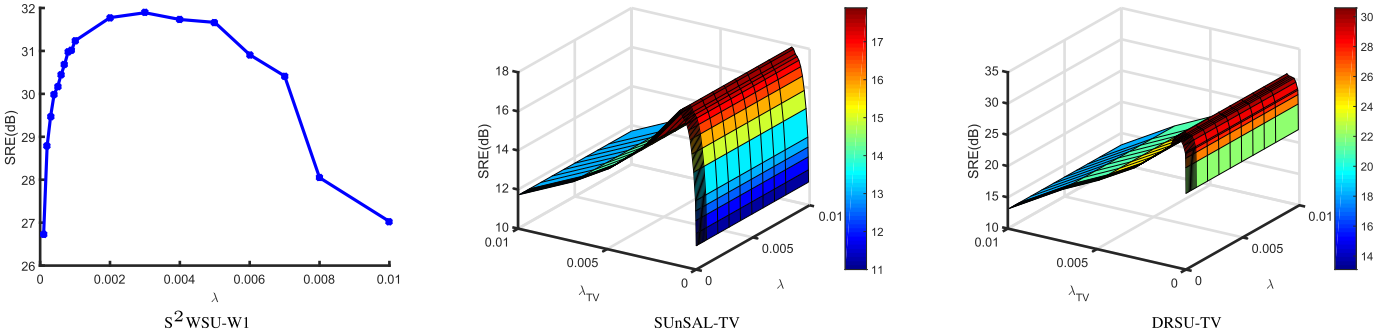


Fig. 4. SRE (dB) as functions of parameters λ or/and λ_{TV} for DC1 with SNR = 40 dB.

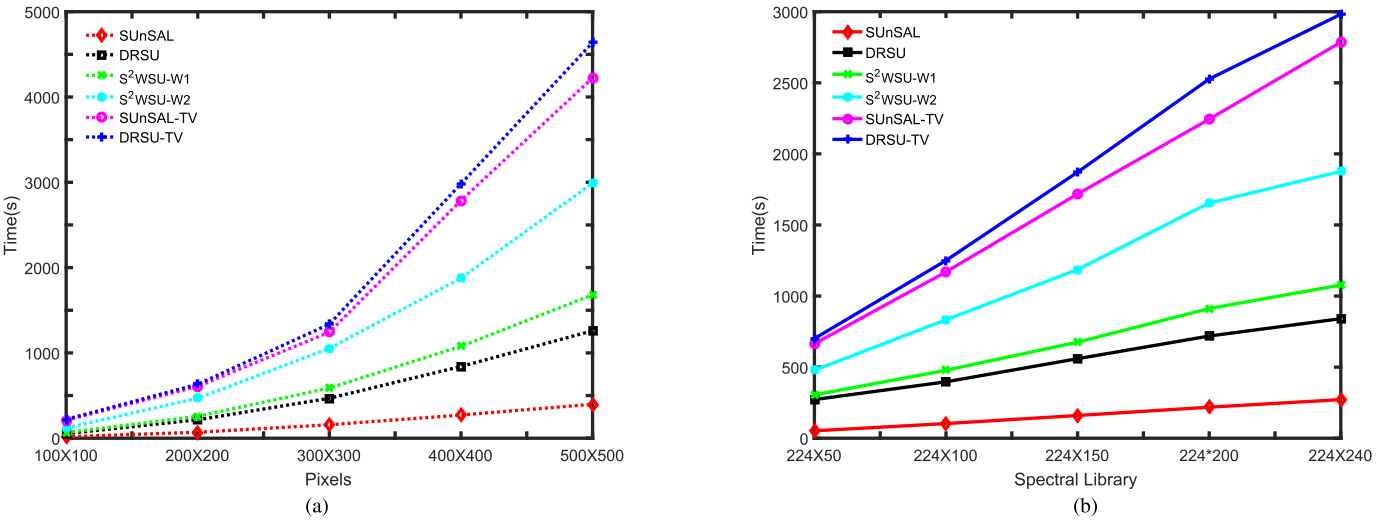


Fig. 5. Processing times (in seconds) of the considered methods. (a) Different number of pixels. (b) Different number of signatures in the spectral library.

2) *Simulated Data Cube 2 (DC2):* Following the work in [45] and [46], a database of 100 × 100-pixel synthetic hyperspectral scene has been created using fractals to generate distinct spatial patterns often found in nature. Nine spectral signatures are chosen from $\mathbf{A_2}$ to generate the synthetic hyperspectral images: Alunite GDS83 Na63, Dumortierite HS190.3B, Halloysite NMNH106236, Kaolinite CM9, Kaolinite KGa-1 (wxyl), Muscovite GDS108, Nontronite GDS41, Pyrophyllite PYS1A fine g, and Sphene HS189.3B. The fractional abundance maps associated with each signature in the construction of the aforementioned synthetic scene are shown in Fig. 3. Similar to DC1, the Gaussian noise with SNR = 30, 40, and 50 dB is included in the experiments.

B. Impact of the Regularization Parameters

In our first experiment, we analyze the impact of the regularization parameters involved in the considered methods. Fig. 4 illustrates the obtained SRE (dB) values as a function

TABLE I

SRE (dB), p_s, AND *Sparsity* SCORES ACHIEVED AFTER APPLYING DIFFERENT UNMIXING METHODS TO THE SIMULATED DATA CUBE 1 (THE OPTIMAL PARAMETERS FOR WHICH THE REPORTED VALUES WERE ACHIEVED ARE INDICATED IN THE PARENTHESES)

Algorithm	**SNR=30dB**			**SNR=40dB**			**SNR=50dB**		
	SRE(dB)	p_s	*sparsity*	SRE(dB)	p_s	*sparsity*	SRE(dB)	p_s	*sparsity*
SUnSAL	8.4373	0.7946	0.0497	15.1721	0.9886	0.0426	23.0894	1	0.0257
	($\lambda = 2e$-2)			($\lambda = 5e$-3)			($\lambda = 1e$-3)		
SUnSAL-TV	11.4304	0.9470	0.0552	17.7695	0.9998	0.0341	26.1655	1	0.0178
	($\lambda = 1e$-2; $\lambda_{TV} = 4e$-3)			($\lambda = 5e$-3; $\lambda_{TV} = 1e$-3)			($\lambda = 2e$-3; $\lambda_{TV} = 2e$-4)		
DRSU	14.9876	0.9745	0.0249	29.6861	1	0.0137	41.1967	1	0.0120
	($\lambda = 3e$-3)			($\lambda = 1e$-3)			($\lambda = 6e$-4)		
DRSU-TV	18.8630	0.9994	0.0241	30.9403	1	0.0128	41.1967	1	0.0120
	($\lambda = 2e$-3; $\lambda_{TV} = 2e$-3)			($\lambda = 2e$-3; $\lambda_{TV} = 4e$-4)			($\lambda = 6e$-4; $\lambda_{TV} = 0$)		
S^{2}WSU-W1	20.5709	0.9995	0.0201	31.9461	1	0.0126	41.4053	1	0.0120
	($\lambda = 5e$-3)			($\lambda = 3e$-3)			($\lambda = 6e$-4)		
S^{2}WSU-W2	20.0216	0.9995	0.0211	31.6869	1	0.0128	41.2838	1	0.0120
	($\lambda = 3e$-3)			($\lambda = 3e$-3)			($\lambda = 6e$-4)		

TABLE II

SRE (dB), p_s AND *Sparsity* SCORES ACHIEVED AFTER APPLYING DIFFERENT UNMIXING METHODS TO THE SIMULATED DATA CUBE 2 (THE OPTIMAL PARAMETERS FOR WHICH THE REPORTED VALUES WERE ACHIEVED ARE INDICATED IN THE PARENTHESES)

Algorithm	**SNR=30dB**			**SNR=40dB**			**SNR=50dB**		
	SRE(dB)	p_s	*sparsity*	SRE(dB)	p_s	*sparsity*	SRE(dB)	p_s	*sparsity*
SUnSAL	6.4259	0.6327	0.0653	11.5833	0.8877	0.0605	18.9987	0.9992	0.0444
	($\lambda = 8e$-3)			($\lambda = 2e$-3)			($\lambda = 3e$-4)		
SUnSAL-TV	9.0371	0.7829	0.0788	15.4514	0.9866	0.0415	25.3557	1	0.0280
	($\lambda = 4e$-3; $\lambda_{TV} = 2e$-3)			($\lambda = 6e$-5; $\lambda_{TV} = 9e$-4)			($\lambda = 5e$-5; $\lambda_{TV} = 9e$-5)		
DRSU	14.2770	0.9454	0.0279	26.0745	1	0.0216	34.5202	1	0.0210
	($\lambda = 2e$-3)			($\lambda = 6e$-4)			($\lambda = 1e$-4)		
DRSU-TV	19.1426	0.9926	0.0273	27.6676	1	0.0216	35.9889	1	0.0209
	($\lambda = 3e$-3; $\lambda_{TV} = 2e$-3)			($\lambda = 6e$-4; $\lambda_{TV} = 4e$-4)			($\lambda = 2e$-4; $\lambda_{TV} = 1e$-4)		
S^{2}WSU-W1	19.5999	0.9946	0.0226	27.9459	1	0.0216	36.5364	1	0.0209
	($\lambda = 3e$-3)			($\lambda = 7e$-4)			($\lambda = 2e$-4)		
S^{2}WSU-W2	19.3593	0.9932	0.0228	27.7186	1	0.0216	36.2337	1	0.0210
	($\lambda = 3e$-3)			($\lambda = 7e$-4)			($\lambda = 2e$-4)		

of parameters λ or/and λ_{TV} for DC1 with SNR $= 40$ dB for the proposed approach, taking S^{2}WSU-W1 as an example, and SUnSAL-TV and DRSU-TV. It can be observed from Fig. 4 that, for the proposed approach (with only one regularization parameter), it is very easy to reach a suboptimal setting. However, for the two competitors, the performance strongly depends on the setting of λ_{TV}. It should be noted that we only report the results obtained from the methods with spatial information. This is because the impact of the parameters involved in the other methods, i.e., S^{2}WSU-W2, SUnSAL, and DRSU, is very similar to that of the proposed approach.

C. Analysis of the Computational Cost

In this experiment, we analyze the computational cost of the considered methods. Fig. 5(a) and (b) shows the processing times (in seconds) as a function of a different number of pixels and a different number of signatures in the spectral library, respectively. The data sets are generated, similar to DC1, using 11 signatures that are randomly selected from the spectral library $\mathbf{A_1}$, varying the number of pixels and the number of signatures in the spectral library. All the algorithms were implemented using MATLAB R2016a on a desktop computer equipped with an Intel Core 7 Duo central processing unit (at 3.6 GHz) and 32 GB of RAM memory. It can be observed from Fig. 5 that, for the methods with spatial information, the proposed S^{2}WSU-W1 and S^{2}WSU-W2 methods are faster than SUnSAL-TV and DRSU-TV, especially with large image sizes or spectral libraries containing many signatures.

D. Results and Discussion

Tables I and II show the SRE (dB), p_s, and *sparsity* results achieved by the different tested methods with the two

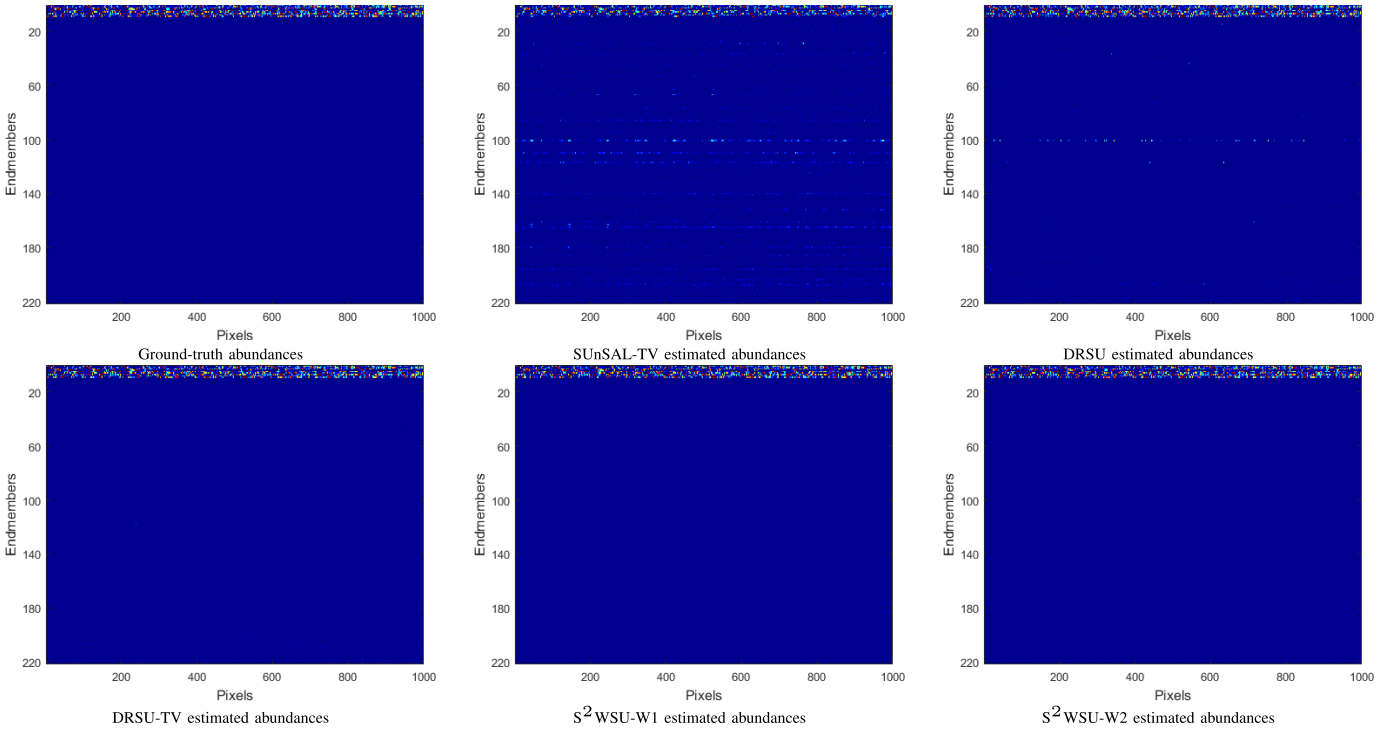


Fig. 6. Ground-truth and estimated abundances obtained for each endmember material in the spectral library for 1000 pixels in DC2 with SNR = 30 dB.

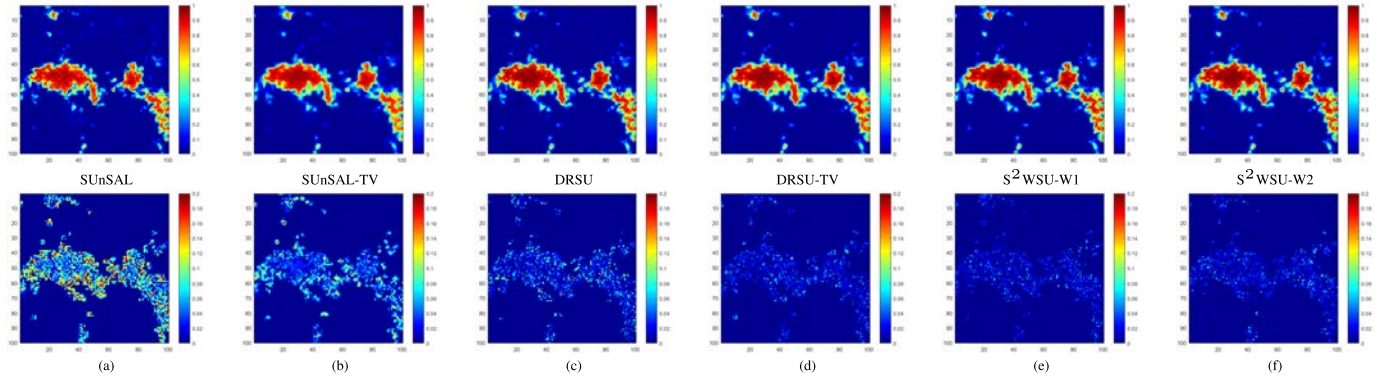


Fig. 7. Abundance maps obtained for the first endmember from DC1 under noise ratio of SNR = 30 dB. Difference map between the ground-truth abundances and the result obtained by (a) SUnSAL, (b) SUnSAL-TV, (c) DRSU, (d) DRSU-TV, (e) S^{2}WSU-W1, and (f) S^{2}WSU-W2.

considered simulated data sets, using all considered SNR levels. In Tables I and II, the best scores obtained across the considered parameter range (the optimal parameters for which the reported values were obtained are indicated in the parentheses) were reported.

From Tables I and II, we can see that the proposed S^{2}WSU-W1 and S^{2}WSU-W2 algorithms obtain better SRE (dB) results than other algorithms in all the cases. p_s obtained by the proposed approaches is also much better than those obtained by other algorithms in the case of low SNR values, which reveals that the inclusion of spatial information leads to high robustness. Furthermore, the proposed approaches can result in a substantial sparsity improvement. Note that, with respect to the degree of sparsity, the S^{2}WSU-W1 and S^{2}WSU-W2 algorithms achieved better or comparable results than the competitors in all the cases, which indicates that the inclusion of a spatial factor in the sparse regularizer can further enhance the sparsity of the solution. Based on this, we can conclude that the spectral–spatial weighted strategy offers the potential to improve unmixing performance in two different analysis scenarios. Finally, the results obtained by S^{2}WSU-W1 and S^{2}WSU-W2 are very similar, which indicates that different window sizes have a little impact on the quality of the final results.

To further illustrate the advantages of the proposed spectral–spatial weighted framework in improving the sparsity on the solution, we take DC2 with SNR = 30 dB as an example for the purpose of visual interpretation. Fig. 6 shows the corresponding fractional abundance estimations obtained for each endmember in $\mathbf{A_2}$, along with the ground-truth abundances, in which the line denotes the abundance of a certain endmember with 1000 randomly selected pixels. It should be noted that, due to space constraints and in view of the relatively poor estimation results of the SUnSAL, we did not show it in Fig. 6. It can be seen that the number of endmembers recovered by the proposed S^{2}WSU-W1 and S^{2}WSU-W2, and

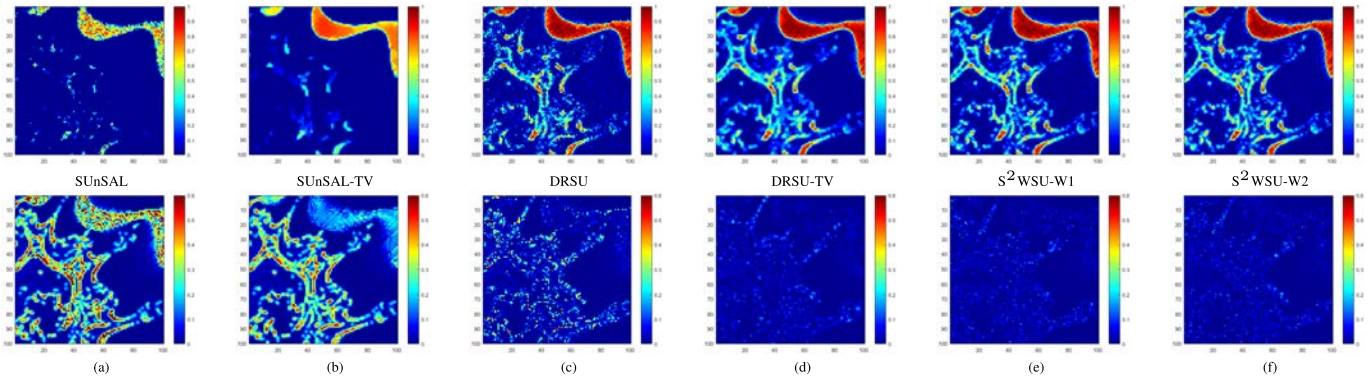


Fig. 8. Abundance maps obtained for the first endmember from DC2 under noise ratio of SNR = 30 dB. Difference map between the ground-truth abundances and the result obtained by (a) SUnSAL, (b) SUnSAL-TV, (c) DRSU, (d) DRSU-TV, (e) S^2WSU-W1, and (f) S^2WSU-W2.

DRSU-TV is very similar to the true number of endmembers, which is much smaller than those produced by SUnSAL-TV and DRSU. In other words, the proposed approaches obtained less abundance lines corresponding to false endmember signatures in the abundance matrix than both SUnSAL-TV and DRSU. The superiority of S^2WSU with respect to the other algorithms indicates that spectral–spatial weights can further improve the ability of the method for successfully identifying endmembers.

For illustrative purposes, Figs. 7 and 8 show a graphical comparison of the considered unmixing algorithms for the simulated problems with an SNR of 30 dB, in which only the abundance map of endmember 1 is presented, as the abundance maps estimated for all endmembers exhibited similar behavior. The difference maps between the estimated abundances and the real ones are computed. The results obtained by SUnSAL look noisy and/or inaccurate. SUnSAL-TV imposes spatial consistency to improve the quality of unmixing results, but the obtained abundance maps are oversmooth. Specially in Fig. 8, the results obtained by both SUnSAL and SUnSAL-TV are inaccurate. The unmixing results obtained by S^2WSU-W1 and S^2WSU-W2 are close to the ground-truth abundances, which reveals a great advantage over other methods, especially the SUnSAL and SUnSAL-TV algorithms. Furthermore, from those difference maps, we can observe that the results produced by the proposed approaches exhibit more details, preserving fine structures and texture, thus providing much better estimations than those obtained by the DRSU algorithm. Finally, the results obtained by S^2WSU methods are slightly better than those provided by DRSU-TV. This further indicates that the spatial weight can promote the spatial correlation on the solution and improve the unmixing performance.

IV. Experiments With Real Hyperspectral Data

In this section, we resort to the well-known Airborne Visible Infrared Imaging Spectrometer (AVIRIS) Cuprite data set for the evaluation of the proposed approach, which is a common benchmark for the validation of spectral unmixing algorithms. The data are available online in reflectance units.[2] The portion used in experiments corresponds to a 350×350 pixel subset of the scene, with 224 spectral bands in the range of 0.4–2.5 μm

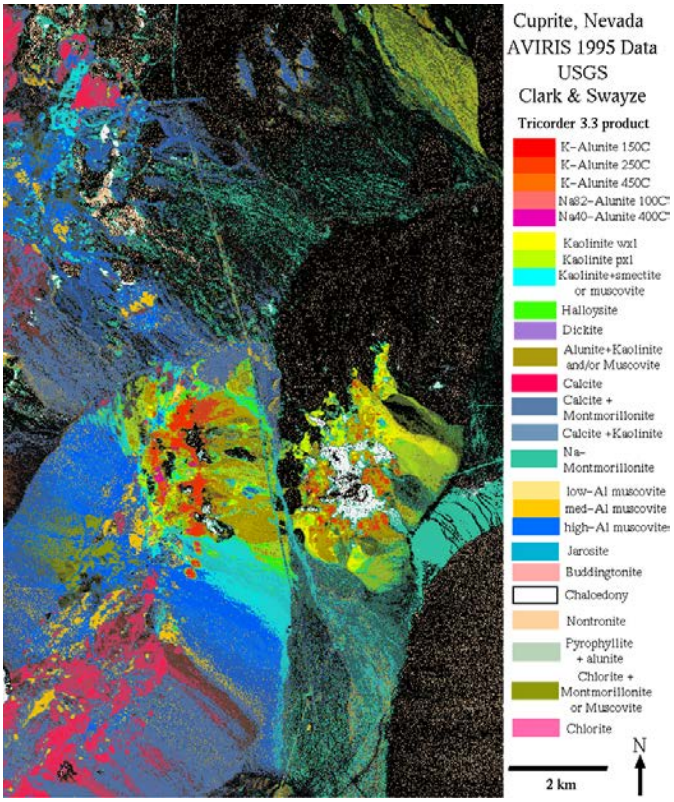


Fig. 9. USGS map showing the location of different minerals in the Cuprite mining district in Nevada.

and the nominal spectral resolution of 10 nm. Prior to the analysis, bands 1 and 2, 105–115, 150–170, and 223 and 224 were removed due to water absorption and low SNR, leaving a total of 188 spectral bands. The spectral library used in this experiment is the same library $\mathbf{A_1}$ used in our simulated experiments and the noisy bands are also removed from $\mathbf{A_1}$. The classification maps of these materials produced by Tricorder software[3] are also displayed. Fig. 9 shows a mineral map produced in 1995 by USGS, in which the Tricorder 3.3 software product [47] was used to map different minerals present in the Cuprite mining district. The USGS map serves as a good indicator for qualitative assessment of the fractional abundance maps produced by the different

[2]http://aviris.jpl.nasa.gov/html/aviris.freedata.html.

[3]http://speclab.cr.usgs.gov/PAPER/tetracorder.

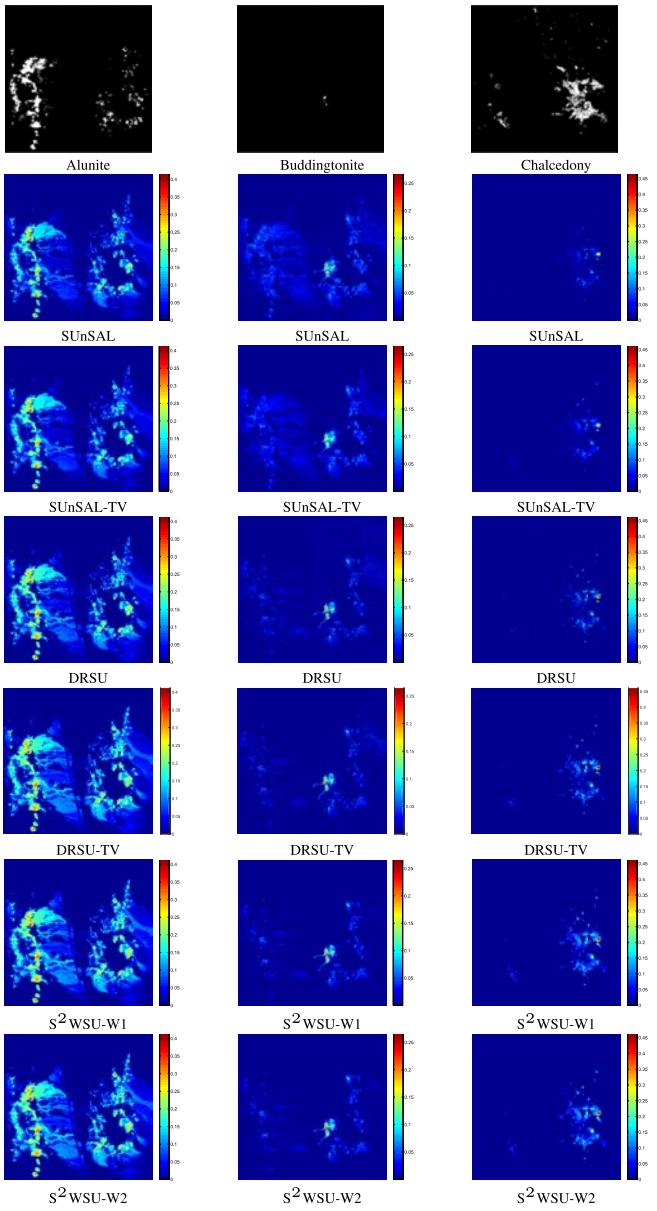


Fig. 10. Fractional abundance maps estimated by SUnSAL, SUnSAL-TV, DRSU, DRSU-TV, S^2WSU-W1, and S^2WSU-W2 as compared with the classification maps produced by USGS Tricorder software for the considered 350×350 pixel subset of the AVIRIS Cuprite scene.

unmixing algorithms. Note that the publicly available AVIRIS Cuprite data were collected in 1997 but the Tricorder map was produced in 1995. In addition, the true abundances of the real hyperspectral data are unavailable. Thus, we can only make a qualitative analysis of the performances of different sparse unmixing algorithms by comparing their estimated abundances with the minerals map.

Fig. 10 conducts a qualitative comparison between the classification maps produced by the USGS Tricorder algorithm and the fractional abundances estimated by the SUnSAL, SUnSAL-TV, DRSU, DRSU-TV, S^2WSU-W1, and S^2WSU-W2 algorithms for three highly representative minerals in the Cuprite mining district (Alunite, Buddingtonite, and Chalcedony). In this experiment, the regularization parameters used for SUnSAL, DRSU, S^2WSU-W1, and S^2WSU-W2 were empirically set to $\lambda = 0.001$, $\lambda = 0.0001$, $\lambda = 0.002$, and $\lambda = 0.002$, respectively, whereas the parameters for SUnSAL-TV and DRSU-TV were set to $\lambda = 0.001$ and $\lambda_{\text{TV}} = 0.001$ and $\lambda = 0.002$ and $\lambda_{\text{TV}} = 0.0001$, respectively. As shown in Fig. 10, all the algorithms obtained reasonable unmixing results, with high abundances for the pixels showing the presence of the considered minerals. This indicates that the sparse unmixing algorithms can lead to good interpretation of the considered hyperspectral data set. However, it can be seen that some of the abundance maps (e.g., Buddingtonite mineral) estimated by SUnSAL and SUnSAL-TV look noisy and the results obtained by SUnSAL-TV are oversmoothed. In addition, DRSU yields abundance maps without good spatial consistency of the minerals of interest (e.g., Chalcedony mineral), and we can also find that the abundances estimated by S^2WSU algorithms are generally comparable or higher in the regions classified as respective minerals in comparison with DRSU. Finally, the *sparsity* obtained by SUnSAL, SUnSAL-TV, DRSU, DRSU-TV, S^2WSU-W1, and S^2WSU-W2 are 0.0682, 0.0743, 0.0430, 0.0423, 0.0420, and 0.0422, respectively. These small differences lead to the conclusion that the proposed approaches use a smaller number of elements to explain the data, thus obtaining higher sparsity. Therefore, from a qualitative viewpoint, we can conclude that the newly developed S^2WSU method exhibits good potential to improve the results obtained by other algorithms in real analysis scenarios.

V. Conclusion and Future Work

In this paper, we have introduced a new spectral–spatial weighted sparse unmixing (S^2WSU) framework for an enhanced hyperspectral data analysis. The proposed S^2WSU simultaneously exploits the spectral and spatial information contained in hyperspectral images via weighting factors, aiming at enhancing the sparsity of the solution. The underlying optimization problem is iteratively solved by an outer–inner looping scheme, where the inner loop updates the unmixing coefficients via the ADMM and the outer loop updates the spectral and spatial weights. Our experiments with both simulated and real hyperspectral data reveal that the S^2WSU consistently achieves a better unmixing performance than other advanced spectral unmixing algorithms. In addition, the proposed method exhibits comparable or lower computational complexity than available techniques. Although the experimental results obtained in this paper are very encouraging, further experiments with additional data sets should be conducted in the future developments to further evaluate the performance of S^2WSU. Furthermore, in future work, we will adopt new techniques, such as low-rank representation [48], [49] and parallel computing [50], [51], to further optimize our new spectral–spatial weighted model for hyperspectral unmixing.

Acknowledgment

The authors would like to thank the associate editor and the anonymous reviewers for their outstanding comments and suggestions, which greatly helped to improve the technical quality and presentation of this paper.

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] J. M. Bioucas-Dias *et al.*, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Observat. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.

[3] A. Plaza, Q. Du, J. M. Bioucas-Dias, X. Jia, and F. A. Kruse, "Foreword to the special issue on spectral unmixing of remotely sensed data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4103–4110, Nov. 2011.

[4] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.

[5] 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.

[6] J. W. Boardman, F. Kruse, and R. Green, "Mapping target signatures via partial unmixing of AVIRIS data," in *Proc. JPL Airborne Earth Sci. Workshop*, 1995, pp. 23–26.

[7] 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.

[8] 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.

[9] A. Ifarraguerri and C. I. Chang, "Multispectral and hyperspectral image analysis with convex cones," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 2, pp. 756–770, Mar. 1999.

[10] 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.

[11] 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.

[12] J. Li, A. Agathos, D. Zaharie, J. M. Bioucas-Dias, A. Plaza, and X. Li, "Minimum volume simplex analysis: A fast algorithm for linear hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 9, pp. 5067–5082, Sep. 2015.

[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] S. Zhang, A. Agathos, and J. Li, "Robust minimum volume simplex analysis for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 11, pp. 6431–6439, Nov. 2017.

[15] M.-D. Iordache, J. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2014–2039, Jun. 2011.

[16] 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.

[17] F. Chen and Y. Zhang, "Sparse hyperspectral unmixing based on constrained $lp - l_2$ optimization," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 5, pp. 1142–1146, Sep. 2013.

[18] Z. Shi, W. Tang, Z. Duren, and Z. Jiang, "Subspace matching pursuit for sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3256–3274, Jun. 2014.

[19] K. E. Themelis, A. Rontogiannis, and K. D. Koutroumbas, "A novel hierarchical Bayesian approach for sparse semisupervised hyperspectral unmixing," *IEEE Trans. Signal Process.*, vol. 60, no. 2, pp. 585–599, Feb. 2012.

[20] W. Tang, Z. Shi, Y. Wu, and C. Zhang, "Sparse unmixing of hyperspectral data using spectral *a priori* information," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 2, pp. 770–783, Feb. 2015.

[21] F. Zhu, Y. Wang, S. Xiang, B. Fan, and C. Pan, "Structured sparse method for hyperspectral unmixing," *ISPRS J. Photogramm. Remote Sens.*, vol. 88, pp. 101–118, Feb. 2014.

[22] 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.

[23] M.-D. 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, Jan. 2014.

[24] 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.

[25] 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.

[26] L. Sun, Z. Wu, L. Xiao, J. Liu, Z. Wei, and F. Dang, "A novel $l_{1/2}$ sparse regression method for hyperspectral unmixing," *Int. J. Remote Sens.*, vol. 34, no. 20, pp. 6983–7001, Oct. 2013.

[27] E. J. Candés, M. B. Wakin, and S. P. Boyd, "Enhancing sparsity by reweighted ℓ_1 minimization," *J. Fourier Anal. Appl.*, vol. 14, nos. 5–6, pp. 877–905, 2008.

[28] C. Y. Zheng, H. Li, Q. Wang, and C. L. P. Chen, "Reweighted sparse regression for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 479–488, Jan. 2016.

[29] R. Wang, H.-C. Li, W. Liao, and A. Pižurica, "Double reweighted sparse regression for hyperspectral unmixing," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2016, pp. 6986–6989.

[30] C. Shi and L. Wang, "Incorporating spatial information in spectral unmixing: A review," *Remote Sens. Environ.*, vol. 149, pp. 70–87, Jun. 2014.

[31] M.-D. Iordache, J. 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.

[32] X.-L. Zhao, F. Wang, T.-Z. Huang, M. K. Ng, and R. J. Plemmons, "Deblurring and sparse unmixing for hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 4045–4058, Jul. 2013.

[33] Y. Zhong, R. Feng, and L. Zhang, "Non-local sparse unmixing for hyperspectral remote sensing imagery," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 1889–1909, Jun. 2014.

[34] S. Zhang, J. Li, K. Liu, C. Deng, L. Liu, and A. Plaza, "Hyperspectral unmixing based on local collaborative sparse regression," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 5, pp. 631–635, May 2016.

[35] J. Bioucas-Dias and M. A. T. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *Proc. 2nd Workshop Hyperspectral Image Signal Process., Evol. Remote Sens. (WHISPERS)*, Reykjavík, Iceland, Jun. 2010, pp. 1–4.

[36] E. J. Candés and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[37] E. J. Candés and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[38] R. Wang, H.-C. Li, A. Pizurica, J. Li, A. Plaza, and W. J. Emery, "Hyperspectral unmixing using double reweighted sparse regression and total variation," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 7, pp. 1146–1150, Jul. 2017.

[39] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Hyperspectral image classification using dictionary-based sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 3973–3985, Oct. 2011.

[40] Y. Zhao, J. Yang, and J. C.-W. Chan, "Hyperspectral imagery super-resolution by spatial–spectral joint nonlocal similarity," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2671–2679, Jun. 2014.

[41] J. Li, M. Khodadadzadeh, A. Plaza, X. Jia, and J. M. Bioucas-Dias, "A discontinuity preserving relaxation scheme for spectral–spatial hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 2, pp. 625–639, Feb. 2016.

[42] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "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.

[43] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Spectral–spatial classification of hyperspectral data using loopy belief propagation and active learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 2, pp. 844–856, Feb. 2013.

[44] J. Hou, H. Gao, and X. Li, "DSets-DBSCAN: A parameter-free clustering algorithm," *IEEE Trans. Image Process.*, vol. 25, no. 7, pp. 3182–3193, Jul. 2016.

[45] G. Martin and A. Plaza, "Region-based spatial preprocessing for endmember extraction and spectral unmixing," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 745–749, Jul. 2011.

[46] G. Martin and A. Plaza, "Spatial-spectral preprocessing prior to endmember identification and unmixing of remotely sensed hyperspectral data," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 380–395, Apr. 2012.

[47] R. Clark *et al.*, "Imaging spectroscopy: Earth and planetary remote sensing with the USGS Tetracorder and expert systems," *J. Geophys. Res.*, vol. 108, no. E12, pp. 5131–5135, Dec. 2003.

[48] X. Lu, Y. Wang, and Y. Yuan, "Graph-regularized low-rank representation for destriping of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 7, pp. 4009–4018, Jul. 2013.

[49] Y. Xu, Z. Wu, J. Li, A. Plaza, and Z. Wei, "Anomaly detection in hyperspectral images based on low-rank and sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, pp. 1990–2000, Apr. 2016.
[50] Z. Wu *et al.*, "GPU parallel implementation of spatially adaptive hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, to be published.
[51] Z. Wu, Y. Li, A. Plaza, J. Li, F. Xiao, and Z. Wei, "Parallel and distributed dimensionality reduction of hyperspectral data on cloud computing architectures," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 6, pp. 2270–2278, Jun. 2016.

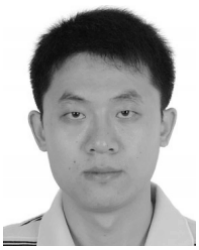
Shaoquan Zhang (S'16) received the B.S. and M.E. degrees from the Nanchang Institute of Technology, Nanchang, China, in 2012 and 2015, respectively. He is currently pursuing the Ph.D. degree with the School of Geography and Planning, Sun Yat-sen University, Guangzhou, China.

His research interests include hyperspectral unmixing, sparse representation, and machine learning.

Jun Li (M'12–SM'16) received the B.S. degree in geographic information systems from Hunan Normal University, Changsha, China, in 2004, the M.E. degree in remote sensing from Peking University, Beijing, China, in 2007, and the Ph.D. degree in electrical engineering from the Instituto de Telecomunicações, Instituto Superior Técnico (IST), Universidade Técnica de Lisboa, Lisbon, Portugal, in 2011.

From 2007 to 2011, she was a Marie Curie Research Fellow with the Departamento de Engenharia Electrotcnica e de Computadores and the Instituto de Telecomunicações, IST, Universidade Técnica de Lisboa, in the framework of the European Doctorate for Signal Processing. She has been actively involved in the Hyperspectral Imaging Network, a Marie Curie research training network involving 15 partners in 12 countries and intended to foster research, training, and cooperation on hyperspectral imaging at the European level. Since 2011, she has been a Post-Doctoral Researcher with the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, Escuela Politécnica, University of Extremadura, Cáceres, Spain. She is currently a Professor with Sun Yat-Sen University, Guangzhou, China. Her research interests include hyperspectral image classification and segmentation, spectral unmixing, signal processing, and remote sensing.

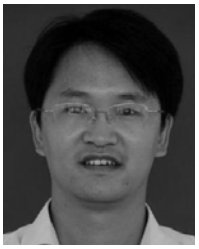
Dr. Li has received several important awards and distinctions, including the IEEE Geoscience and Remote Sensing Society Early Career Award in 2017 due to her outstanding contributions to remotely sensed hyperspectral and synthetic aperture radar data processing. She has been an Associate Editor of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING since 2014. She has been a Guest Editor for several journals, including the PROCEEDINGS OF THE IEEE and *ISPRS Journal of Photogrammetry and Remote Sensing*. She has also been an active Reviewer for several journals, including the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, and the IEEE TRANSACTIONS ON IMAGE PROCESSING.



Heng-Chao Li (S'06–M'08–SM'14) received the B.Sc. and M.Sc. degrees from Southwest Jiaotong University, Chengdu, China, in 2001 and 2004, respectively, and the Ph.D. degree from the Graduate University of Chinese Academy of Sciences, Beijing, China, in 2008, all in information and communication engineering.

From 2013 to 2014, he was a Visiting Scholar with Prof. W. J. Emery at the University of Colorado at Boulder, Boulder, CO, USA. He is currently a Professor with the Sichuan Provincial Key Laboratory of Information Coding and Transmission, Southwest Jiaotong University, Chengdu. His research interests include the statistical analysis of synthetic aperture radar images, remote sensing image processing, and signal processing in communications.

Dr. Li received several scholarships or awards, especially including the Special Grade of the Financial Support from China Postdoctoral Science Foundation in 2009 and the New Century Excellent Talents in University from the Ministry of Education of China in 2011. He has also been a Reviewer for several international journals and conferences, such as the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, the IEEE TRANSACTIONS ON IMAGE PROCESSING, the *IET Radar, Sonar and Navigation*, and the *Canadian Journal of Remote Sensing*. He is currently serving as an Associate Editor of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING.



Chengzhi Deng received the B.S. and M.S. degrees from Jiangxi Normal University, Nanchang, China, in 2002 and 2005, respectively, and the Ph.D. degree from the Huazhong University of Science and Technology, Wuhan, China, in 2008.

He is currently a Professor with the Department of Information Engineering, Nanchang Institute of Technology, Nanchang, where he is also the Deputy Director of the Jiangxi Province Key Laboratory of Water Information Cooperative Sensing and Intelligent Processing. His research interests include hyperspectral image processing, machine learning, and remote sensing of water environment.

Antonio Plaza (M'05–SM'07–F'15) received the M.Sc. and Ph.D. degrees in computer engineering from the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, University of Extremadura, Cáceres, Spain, in 1999 and 2002, respectively.

He is currently the Head of the Hyperspectral Computing Laboratory, Department of Technology of Computers and Communications, University of Extremadura. He has authored or co-authored over 600 publications, including over 200 *Journal of Controlled Release* journal papers (over 150 in IEEE journals), 24 book chapters, and over 300 peer-reviewed conference proceeding papers. His research interests include hyperspectral data processing and parallel computing of remote sensing data.

Dr. Plaza is a fellow of IEEE for contributions to hyperspectral data processing and parallel computing of Earth observation data. He was a member of the Steering Committee of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING. He was a recipient of the recognition of best reviewers of the IEEE GEOSCIENCE AND REMOTE SENSING LETTERS in 2009 and a recipient of the recognition of best reviewers of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING (TGRS) in 2010, for which he served as an Associate Editor in 2007–2012. He is also an Associate Editor of the IEEE ACCESS and was a member of the Editorial Board of the IEEE GEOSCIENCE AND REMOTE SENSING NEWSLETTER in 2011–2012 and the IEEE GEOSCIENCE AND REMOTE SENSING MAGAZINE in 2013. He was a recipient of the Best Column Award of the *IEEE Signal Processing Magazine* in 2015, the 2013 Best Paper Award of the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, and the most highly cited paper in 2005–2010 in the *Journal of Parallel and Distributed Computing*. He received the best paper awards at the IEEE International Conference on Space Technology and the IEEE Symposium on Signal Processing and Information Technology. He served as the Director of Education Activities for the IEEE Geoscience and Remote Sensing Society (GRSS) in 2011–2012 and a President of the Spanish Chapter of the IEEE GRSS in 2012–2016. He has served as the Editor-in-Chief of the IEEE TGRS from 2013 to 2017. He has guest edited ten special issues on hyperspectral remote sensing for different journals. He has reviewed more than 500 papers for over 50 different journals.