

Sparse Unmixing for Hyperspectral Imagery via Comprehensive-Learning-Based Particle Swarm Optimization

Yapeng Miao¹ and Bin Yang¹, *Member, IEEE*

Abstract—Sparse unmixing methods have been extensively studied as a popular topic in hyperspectral image analysis for several years. Fundamental model-based unmixing problems can be better reformulated by exploiting sparse constraints in different forms. Gradient-based optimization approaches commonly serve for traditional sparse unmixing, but their limitations such as one-way search, often induce unsatisfactory local optimum, especially when the problems are nonconvex. Therefore, acceptable unmixing performance cannot always be guaranteed, and the sparsity of hyperspectral imagery may be incorrectly expressed. In this article, an unsupervised sparse unmixing method using comprehensive-learning-based particle swarm optimization (PSO) is proposed. Due to the basic PSO's premature convergence in dealing with high-dimensional problems, double swarms whose fitness functions are accordingly divided into a series of low-dimensional subproblems are constructed to search for optimal endmembers and abundances alternately, leading to the implementation of unmixing in refined solution spaces. Under this framework, two comprehensive learning strategies are introduced to promote and refine particles' mutual learning deeply at the element-level, through which the abundance sparsity in every local pixel and every endmember's global abundance sparsity can be better exploited and expressed. Experiments with both simulated datasets and real hyperspectral images are employed to validate the performance of the proposed method combined with different sparse constraints. In comparison with other state-of-the-art algorithms, the proposed method enables the achievement of better unmixing results.

Index Terms—Comprehensive learning, double swarms, hyperspectral imagery, particle swarm optimization, solution space partition, sparse regularization, spectral unmixing.

I. INTRODUCTION

HYPERSPECTRAL remote sensing based on imaging spectroscopy technology has contributed significantly to earth observation in the past few decades [1]–[3]. Since hyperspectral sensors can continuously sample the spectral information of the earth's surface in hundreds of narrow bands, each pixel of a hyperspectral image (HSI) can be characterized

by a unique spectral curve [4]. The substantial increase in spectral information enables the HSIs to be of great value in applications including geological survey, mineral and energy detection, environment and disaster monitoring, food safety, and military [5]–[7]. However, during the data collection process, the limited spatial resolution of the imaging spectrometer and the complex surface conditions usually result in the instantaneous field of view containing more than one material. Therefore, a large number of mixed pixels exist in the HSIs and affect the accuracy of remote sensing applications severely at the pixel scale [8]. Spectral unmixing technology solves this problem by extracting characteristic spectra of each kind of typical material and their corresponding fractional proportions from the HSIs, which are called endmembers and abundances, respectively [9]–[11]. The optimization problem of unmixing can be convex or nonconvex based on concrete tasks and applications. In the past decade, nonconvex unmixing frameworks have shown their effectiveness for automatically extracting favorable features, especially when the effects of various degradation mechanisms (e.g., nonlinear mixing effects and environmental conditions) should be considered [12].

Current spectral unmixing methods are mainly proposed based on the popular linear mixing model (LMM) or nonlinear mixing models (NLMMs) [4], [13], [14]. Taking light's multiple scatterings among ground covers into account, the NLMMs can explain the nonlinear mixing effects compared with the LMM only based on the single scattering assumption. Recently, nonlinear unmixing methods have attached more and more attention in dealing with complex scenarios. For instance, bilinear mixing models such as generalized bilinear model (GBM) [15] and polynomial postnonlinear model [16] explain the major nonlinearity with second-order scatterings, which greatly simplifies traditional physical models. Moreover, Wang *et al.* [17] used a plug-and-play prior technique based on the GBM to exploit the spatial correlation of abundance maps and nonlinear interaction maps, and obtained more accurate abundances. Resorting to the GBM, Gao *et al.* [18] implemented nonlinear unmixing with nonnegative tensor factorization, which retained the HSIs' spatial information and revealed the low-rank structures of abundance maps.

In contrast to the abovementioned complex nonlinear unmixing methods, linear unmixing methods usually have a simpler and acceptable explanation for the light propagation processes, which is meaningful in many situations. This article still focuses

Manuscript received August 1, 2021; revised September 8, 2021; accepted September 21, 2021. Date of publication September 24, 2021; date of current version October 8, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 62001098 and in part by the Fundamental Research Funds for the Central Universities under Grant 2232020D-33. (Yapeng Miao and Bin Yang contributed equally to this work.) (Corresponding author: Bin Yang.)

The authors are with the School of Computer Science and Technology, Donghua University, Shanghai 201620, China (e-mail: k13617412792@163.com; yangb19@dhu.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2021.3115177

on linear unmixing. Under the assumption of the LMM, each pixel is approximately expressed as the linearly weighted sum of endmembers according to their respective abundances [19]. As a critical step in unmixing, many endmember extraction algorithms have been proposed so far. They take advantage of the convex geometric properties of the LMM that the pixel points are located in a simplex with endmembers being vertices in the feature space. Endmember extraction can be transformed into the search of the simplex's vertices. Some methods repeatedly projected the pixel points and randomly selected the endpoints of projection vectors as the vertices in the feature space [20]. Moreover, a part of the methods tried to extract a subset of pixels corresponding to the largest simplex volume as the endmembers [21] or utilized orthogonal projection repeatedly to determine the vertices that expand the simplex, such as the popular vertex component analysis (VCA) [22]. Different variants based on these methods have also been recently researched [23]–[25]. With the use of the extracted endmembers, supervised abundance inversion algorithms such as the fully constrained least squares (FCLS) proposed by Heinz and Chang [26] can usually estimate satisfactory abundances.

It is worth noting that a pixel in a real hyperspectral image, often contains not all of the endmembers, and a kind of ground cover exists in a part of pixels. Thus, the HSIs have prominent sparse structures. Owing to this feature, a variety of sparse regularization terms including ℓ_0 , ℓ_1 , $\ell_{2,1}$, $\ell_{1/2}$, and $\ell_1 - \ell_2$ norms, have been introduced in the construction of unmixing problems. A series of sparse unmixing algorithms are proposed, which can produce unmixing results closer to the ground truth. Iordache *et al.* [27] solved the unmixing problem through sparse regression where sparse constraints such as ℓ_1 norm were enforced on the abundances and the spectral library for unmixing using variable separation augmented Lagrangian (SUnSAL). They further proposed an improved SUnSAL-TV [28] algorithm. A total variation (TV) regularizer was used to exploit spatial information and added to the previous optimization problem to make the attained abundances smoother. Furthermore, a collaborative sparse regression framework was proposed using the $\ell_{2,1}$ norm to accommodate the sparsity of each pixel [29]. Wang *et al.* [30] provided a detailed discussion about the effectiveness of $\ell_{2,p}$ regularization term in sparse unmixing. Sun *et al.* [31] built a novel sparsity $\ell_1 - \ell_2$ norm and combined it with the TV regularization term for hyperspectral unmixing. Moreover, among these mathematical forms of sparsity, the ℓ_0 norm can be used for unmixing through a given row hard threshold function [32]. Li *et al.* [33] adopted a reweighted $\ell_{1/2}$ norm to achieve sparse abundances. Some recently proposed methods exploited the HSIs' low-rank property in sparse unmixing to further improve the accuracy of abundance estimation [34], [35].

On the other hand, a reliable consensus is that insufficiently accurate endmembers could deteriorate the estimated abundances. Unfortunately, supervised sparse unmixing is dependent on the accuracy of the attained endmembers. Therefore, unsupervised unmixing becomes meaningful, which integrates endmember extraction with abundance inversion in a common framework. As an unsupervised feature extraction approach

whose optimization problem is nonconvex, nonnegative matrix factorization (NMF) has been often used for spectral unmixing. The introduction of proper constraints into the NMF framework has drawn considerable attention in the field of unmixing to alleviate the effect of local minima. Besides regularizers constructed based on the HSI's spatial information [36] and the LMM's geometric properties [37], sparse constraints enforced on abundances have also been widely studied in the constrained NMF unmixing framework.

For example, the $\ell_{1/2}$ norm was introduced into the NMF to promote the obtained abundances' sparsity [38]. In the case of a manifold regularizer being considered simultaneously, the sparse NMF unmixing algorithm could produce sparse and smooth abundances [39]. Considering the properties of HSIs, adaptive relationship preserving sparse NMF was developed to unite similarity learning and unmixing in an alternate optimization process to learn the best similarity matrix [40]. The global spatial similarity between the pixels of HSIs and the local spectral similarity between superpixels have also been exploited to improve the sparse NMF [41]. However, the original NMF's objective function is commonly sensitive to noises and outliers. In order to overcome this problem, self-paced NMF replaced the Euclidean distance loss function of the standard NMF with weighted least-squares losses and used a self-paced learning strategy to learn the weights adaptively [42]. Since the traditional single-layer NMF may produce incorrect unmixing results, Feng *et al.* [43] extended the single-layer NMF to deep NMF, which promoted the smooth segmentation of abundances. Besides, sparse constraints could be further enforced on each layer of the deep NMF [44]. The idea of subspace clustering has also been drawn into the NMF framework for unmixing [45], [46].

In the literature, the optimization of a majority of supervised and unsupervised sparse unmixing methods rely heavily on gradient information-driven approaches [47], [48], such as the multiplicative update rule and gradient descent algorithm. The gradients of the objective function should be derived in advance, and there is a high requirement for initial points and step size settings. For nonconvex constrained optimization problems, these algorithms are prone to fall into local optima, resulting in the inaccurate representation of the HSIs' sparsity. Artificial intelligence is considered a promising way to overcome such shortcomings of the traditional methods and has drawn much interest in the community. Recently, deep learning methods such as graph convolutional networks [49] have shown their power in improving hyperspectral classification performance. For unmixing, methods including cascaded autoencoders [50] and endmember-guided unmixing networks [51] have been proposed to yield more reasonable unmixing results. On the other hand, different from the neural networks' complex mutually connected structures, swarm computing intelligence addresses the issue of numerical optimization through the explicit cooperation and simple movement of multiple individuals. For instance, as a global optimization technique advocated by simulating the foraging behavior of birds, particle swarm optimization (PSO) [52] has the advantages of easy implementation and preferable convergence. Compared with the gradient-based optimization

methods, better solutions are more likely to be achieved by PSO, especially for the complex nonconvex constrained optimization problems.

Recently, PSO has opened a new perspective in the spectral unmixing field [53]–[55]. Zhang *et al.* [56] came up with a PSO-expectation maximization (PSO-EM) unmixing method based on the normal compositional model. Tong *et al.* [57] improved multiobjective discrete PSO by using a restart mechanism to enhance the diversity of the population and estimate more accurate endmembers. Du *et al.* [58] combined a quantum PSO algorithm with the collaborative update of the particles' best positions to extract endmembers. Xu *et al.* [59] designed a PSO-based unmixing strategy for highly mixed hyperspectral data. Although PSO has shown its good performance in some unmixing algorithms, it could be premature when the dimensionality of the problem is too high, generating undesirable unmixing results [60]. In consideration of this problem, Yang *et al.* [61] converted the original unmixing problem into a series of low-dimensional simple subproblems and processed them in parallel under the framework of PSO where divide-and-conquer dimensionality division facilitates the accurate implementation of unmixing.

In a word, although spectral unmixing has been studied a lot, it is still challenging and meaningful to improve the existing unsupervised sparse unmixing architecture in two aspects. First, in terms of improving the optimization robustness, instead of using the traditional gradient-based methods, we can reconstruct the PSO's structure accordingly to better solve the nonconvex sparse unmixing problem. Second, in terms of the correct expression for the abundances' inherent sparsity, it is promising to exploit the information exchange among the PSO's particles to better balance sparse regularization terms during unmixing. In this sense, abundances with both higher accuracy and correct sparsity can be obtained.

Therefore, in this article, we proposed an unsupervised sparse unmixing method using comprehensive-learning-based particle swarm optimization. To address the high-dimensional sparse unmixing problem and promote the diversity of PSO's swarms, we have adopted the idea of dimension division in [61]. First, an endmember swarm and an abundance swarm are constructed. Their optimization subproblems are correspondingly split according to bands or pixels. Consequently, the swarms' fitness functions can be vectorized, and the original feasible solution space is divided into low-dimensional slices where particles' positions can be finely determined. In brief, for example, the determination of the whole abundance matrix depends on the parallel search of the best abundance vectors of every pixel. Second, these two swarms are updated alternately by using the global best positions of each other, aiming at optimizing endmembers and abundances.

Moreover, in order to deeply refine the position search in every pixel's abundance vector, we have introduced a comprehensive learning strategy to promote the learning of the favorable abundances of every specific endmember. This improvement not only significantly enhances the diversity of the abundance swarm but also increases the probability of finding the optimal solution. In addition, another novel global comprehensive learning

mechanism focusing on the sparsity of abundances has been built through which both the abundance sparsity of each local pixel and the global abundance sparsity of each endmember can be better expressed in unmixing results.

The main contributions of this article can be summarized as follows.

- 1) A versatile PSO sparse unmixing framework that is suitable for different well-known sparse constraints is proposed. An existing double swarm PSO method based on dimension division is extended and reconstructed to cope with the task of unsupervised sparse unmixing.
- 2) Two comprehensive learning strategies are used to improve the PSO sparse unmixing framework. An advanced comprehensive learning strategy is modified to refine the searched particles' positions of an abundance swarm. In addition, another new comprehensive learning strategy is devised to improve the sparsity of the abundances.

The rest of this article is organized as follows. Section II gives a brief introduction to the linear mixture model, sparse unmixing framework, and PSO. Section III presents the description of the proposed algorithm. In Section IV, the effectiveness of the proposed algorithm is verified in experiments with simulated data and real hyperspectral images. Finally, Section V concludes this article.

II. RELATED WORKS

A. Linear Mixture Model

The LMM assumes that each pixel can be approximately written as the linear combination of endmembers according to their respective abundances. Let an HSI matrix be $\mathbf{X} \in \mathbb{R}^{m \times n}$, $\mathbf{E} \in \mathbb{R}^{m \times r}$ be the endmember matrix, and $\mathbf{A} \in \mathbb{R}^{r \times n}$ denote the abundance matrix, where m represents the number of spectral bands and n is the number of pixels. The LMM can be described as follows:

$$\mathbf{X} = \mathbf{E}\mathbf{A} + \mathbf{N} \quad (1)$$

where $\mathbf{N} \in \mathbb{R}^{m \times n}$ denotes the noise error. To be physically meaningful, the abundance must meet the nonnegative and sum-to-one constraints (ASC) in the following equation:

$$\mathbf{A} \geq 0, \mathbf{1}_r^T \mathbf{A} = \mathbf{1}_n^T \quad (2)$$

where $\mathbf{1}_r^T$ and $\mathbf{1}_n^T$ represent two column vectors of all 1s.

B. Sparse Unmixing Framework

If sparse constraints are enforced on the abundances, a typical sparse regression model for unmixing is expressed as

$$\begin{aligned} \min_{\mathbf{E}, \mathbf{A}} \quad & f(\mathbf{E}, \mathbf{A}) = \frac{1}{2} \|\mathbf{X} - \mathbf{E}\mathbf{A}\|_F^2 + \lambda \cdot \psi(\mathbf{A}) \\ \text{s.t.} \quad & \mathbf{A} \geq 0, \mathbf{1}_r^T \mathbf{A} = \mathbf{1}_n^T \end{aligned} \quad (3)$$

where $\psi(\mathbf{A})$ corresponds to the abundances' sparse constraint. It turns to $\|\mathbf{A}\|_1 = \sum_{i=1}^r \sum_{j=1}^n |\mathbf{A}_{i,j}|$, $\|\mathbf{A}\|_{1/2} = \sum_{i=1}^r \sum_{j=1}^n \sqrt{\mathbf{A}_{i,j}}$, and $\|\mathbf{A}\|_{2,1} = \sum_{j=1}^n \|\mathbf{A}_j\|_2$ (\mathbf{A}_j : represents the j th row of \mathbf{A}) when the ℓ_1 norm, $\ell_{1/2}$ norm, and $\ell_{2,1}$ norm are used to make \mathbf{A} sparse, respectively. These three sparse regularization terms are considered in this

article, and they have been popularly used in a variety of sparse unmixing methods, including SUnSAL [27], collaborative SUnSAL (CLSUnSAL) [29], and $\ell_{1/2}$ -NMF[38]. λ is a small nonnegative regularization parameter that controls the influence of the sparse constraints on the estimated abundances.

C. Particle Swarm Optimization

Particle swarm optimization algorithm originating from the study of birds' predatory behavior belongs to the swarm intelligence, which utilizes the collaboration of particles to search for optimal solutions to the problems. PSO's particles transmit position information through cooperation in every iteration until an optimal solution is found. Different from the traditional gradient-based methods, the particles' positions are initialized randomly and not unique, and the update of particles' velocities also has some degree of randomness. This makes it more likely to jump out of bad local optima and increase the probability of attaining the global optimal solution.

The basic PSO initializes a swarm with q particles. In the t th iteration, every particle has its current position $\mathbf{P}_\mu^t \in \mathbb{R}^n$ and velocity $\mathbf{V}_\mu^t \in \mathbb{R}^n$. The best position in history $\mathbf{P}_{\mu\text{pbest}}^t \in \mathbb{R}^n$, and the global best position $\mathbf{P}_{\text{gbest}}^t \in \mathbb{R}^n$ in the swarm need to be updated, $\mu = 1, \dots, q$. Every particle's position is actually a potential solution for the optimization problem. For example, given an objective function $S(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^n$ of a minimization optimization problem, it is regarded as the fitness function of PSO. The particles are compared with each other according to $S(\mathbf{x})$ so that the favorable position information can be exchanged and learned. In this sense, $\mathbf{P}_{\mu\text{pbest}}^t$ and $\mathbf{P}_{\text{gbest}}^t$ in the swarm can be calculated by (4). Resorting to $\mathbf{P}_{\mu\text{pbest}}^t$ and $\mathbf{P}_{\text{gbest}}^t$, a particle's velocity and its current position are updated using (5) and (6), respectively. This iterative process repeatedly runs until the convergence conditions are satisfied and $\mathbf{P}_{\text{gbest}}^t$ is the optimal solution for the optimization problem

$$\begin{cases} \mathbf{P}_{\mu\text{pbest}}^t = \mathbf{P}_{\mu\text{pbest}}^{t-1} & S(\mathbf{P}_{\mu\text{pbest}}^{t-1}) \leq S(\mathbf{P}_\mu^t) \\ \mathbf{P}_{\mu\text{pbest}}^t = \mathbf{P}_\mu^t & S(\mathbf{P}_{\mu\text{pbest}}^{t-1}) > S(\mathbf{P}_\mu^t) \\ \mathbf{P}_{\text{gbest}}^t = \arg \min \{S(\mathbf{P}_{1\text{pbest}}^t), S(\mathbf{P}_{2\text{pbest}}^t), \dots, S(\mathbf{P}_{q\text{pbest}}^t)\} \end{cases} \quad (4)$$

$$\mathbf{V}_\mu^{t+1} = \omega \cdot \mathbf{V}_\mu^t + c_1 \cdot r_1^t \cdot (\mathbf{P}_{\mu\text{pbest}}^t - \mathbf{P}_\mu^t) + c_2 \cdot r_2^t \cdot (\mathbf{P}_{\text{gbest}}^t - \mathbf{P}_\mu^t) \quad (5)$$

$$\mathbf{P}_\mu^{t+1} = \mathbf{P}_\mu^t + \mathbf{V}_\mu^{t+1}. \quad (6)$$

In (5), ω is an inertia weight used to control the swarm's exploration and development. c_1 and c_2 are mainly used to balance the influence of cognitive and social learning on \mathbf{V}_μ^t , while r_1 and r_2 are random numbers uniformly distributed in the range of (0,1).

III. PROPOSED METHOD

A. Dimension Division for Two Swarms in PSO

The optimization of the fundamental unsupervised sparse unmixing problem in (3) is resorting to PSO in this article. Two swarms containing q particles are generated to update the

endmember matrix \mathbf{E} and the abundance matrix \mathbf{A} alternately. Explicitly, the particles' positions represent the potential \mathbf{E} or \mathbf{A} that are updated when the global best position $\mathbf{P}_{\text{gbest}}^t$ is determined as the iterations progress.

As mentioned in Section II, PSO has strong global searching abilities, but it is sensitive to the problem's dimension in practice. Commonly, as the search space's dimension increases, the chances of the global optimal solution being found by PSO will decrease [62]. It becomes hard to control the particles' flying directions in the high-dimensional space, causing the whole swarm to lose its valuable diversity and converge early. However, it is noted that the problem in (3) is separable, which can be easily split into dozens of subproblems across the rows or columns of the hyperspectral data \mathbf{X} . Therefore, the strategy of dimension division devised in [61] can be directly extended here for the PSO-based sparse unmixing. In this sense, the objective functions of the swarms turn into two fitness vectors based on the band-wise or pixelwise division. Benefiting from the idea of divide and conquer, the swarms' global best particles' positions corresponding to more accurate endmembers and abundances can be found.

1) *Endmember Swarm*: The optimization subproblem for updating the endmembers can be written as follows:

$$\begin{aligned} \min_{\mathbf{E}} f(\mathbf{E}, \mathbf{A}) &= \|\mathbf{X} - \mathbf{E}\mathbf{A}\|_F^2 \\ \text{s.t. } \mathbf{E} &\geq 0 \end{aligned} \quad (7)$$

where the abundance matrix \mathbf{A} is a known variable. This problem is divided based on the *band* vector slices and transformed into the endmember swarm's fitness vector $\phi_E(\mathbf{E}_\mu)$, in (8), where $\mathbf{E}_\mu \in \mathbb{R}^{m \times r}$ is the μ th endmember particle's position, $\mu = 1, \dots, q$, and \mathbf{X}_{ij} denotes the value of the j th pixel's i th spectral band

$$\phi_E(\mathbf{E}_\mu) = (z_1, z_2, \dots, z_m)^T, z_i = \sum_{j=1}^n (\mathbf{X}_{ij} - (\mathbf{E}_\mu \mathbf{A})_{ij})^2. \quad (8)$$

According to the vectorized fitness function of the endmember swarm, each particle's best position in history and the swarm's global best position are determined as follows:

$$\begin{cases} \mathbf{E}_{\mu\text{pbest}}^t(i, :) \\ = \begin{cases} \mathbf{E}_{\mu\text{pbest}}^{t-1}(i, :), & \phi_E(\mathbf{E}_{\mu\text{pbest}}^{t-1})_i \leq \phi_E(\mathbf{E}_\mu^t)_i \\ \mathbf{E}_\mu^t(i, :), & \phi_E(\mathbf{E}_{\mu\text{pbest}}^{t-1})_i > \phi_E(\mathbf{E}_\mu^t)_i \end{cases} \\ \mathbf{E}_{\text{gbest}}^t(i, :) \\ = \arg \min (\phi_E(\mathbf{E}_{1\text{pbest}}^t)_i, \dots, \phi_E(\mathbf{E}_{q\text{pbest}}^t)_i) \end{cases} \quad (9)$$

where $\mathbf{E}_\mu^t(i, :)$ denotes the i th row of the μ th endmember particle's position, $i = 1, 2, \dots, m$.

2) *Abundance Swarm*: The objective function for the abundance optimization problem is based on (3). The endmember matrix \mathbf{E} is a known variable here. Similarly, we divide it based on the *pixel* vector slices and transform it into the corresponding abundance fitness vector $\phi_{A1}(\mathbf{A}_\mu)$ in (10). $\mathbf{A}_\mu \in \mathbb{R}^{r \times n}$ is the position of μ th abundance particle

$$\phi_{A1}(\mathbf{A}_\mu) = (b_1, b_2, \dots, b_n), \quad b_j$$

$$= \sum_{i=1}^m \left(\mathbf{X}_{ij} - (\mathbf{E}\mathbf{A}_\mu)_{ij} \right)^2 + \lambda \cdot \psi_j(\mathbf{A}_\mu). \quad (10)$$

In (10), $\psi_j(\mathbf{A}_\mu)$ is the j th additive subterm of sparse constraints with respect to the j th pixel. In the cases of $\ell_{1/2}$ norm and $\ell_{2,1}$ norm, it has the following expression:

$$\psi_j(\mathbf{A}_\mu) = \begin{cases} \sum_{k=1}^r \sqrt{(\mathbf{A}_\mu)_{kj}}, & \ell_{1/2}\text{norm} \\ \sqrt{\sum_{k=1}^r (\mathbf{A}_\mu)_{kj}^2}, & \ell_{2,1}\text{norm} \end{cases}. \quad (11)$$

For ℓ_1 norm, since abundances are normalized to satisfy the ASC here, the dimension division is not required for $\|\mathbf{A}_\mu\|_1$ in fact. Instead, an elementwise soft thresholding method [4] with a parameter α is applied. The determination of the particle's best position in history and the swarm's global best position of the abundance swarm can be achieved in

$$\begin{cases} \mathbf{A}_{\mu\text{pbest}}^t(:,j) \\ = \begin{cases} \mathbf{A}_{\mu\text{pbest}}^{t-1}(:,j) & \phi_{A1}(\mathbf{A}_{\mu\text{pbest}}^{t-1})_j \leq \phi_{A1}(\mathbf{A}_\mu^t)_j \\ \mathbf{A}_\mu^t(:,j) & \phi_{A1}(\mathbf{A}_{\mu\text{pbest}}^{t-1})_j > \phi_{A1}(\mathbf{A}_\mu^t)_j \end{cases} \\ \mathbf{A}_{\text{gbest}}^t(:,j) \\ = \arg \min(\phi_{A1}(\mathbf{A}_{1\text{pbest}}^t)_j, \dots, \phi_{A1}(\mathbf{A}_{q\text{best}}^t)_j) \end{cases}. \quad (12)$$

In (12), $\mathbf{A}_\mu^t(:,j)$ is the j th column of the μ th abundance particle's position, $j = 1, 2, \dots, n$.

B. Comprehensive Learning Strategies

In order to further improve the diversity of the abundance swarm and exploit the reasonable expression of abundances' sparsity exhaustively, we devise two comprehensive learning strategies under the framework of dimension division-driven PSO in this section.

1) *Comprehensive Learning for Position Search Refinement*: Dimension division indeed helps to reduce the impact of high-dimension by limiting the search scope of the abundance swarm to multiple pixelwise slices. It can be referred that a more accurate position can be possibly achieved if the deep elementwise dimension division can be further executed. To this end, the first comprehensive learning strategy is built referring to [63] to refine the position search for the previously divided pixel slices.

First, the μ th particle of the abundance swarm is assigned a learning probability P_{C_μ} defined in (13). Considering the i th row $\mathbf{A}_{i,\mu}^t$ of \mathbf{A}_μ^t , given a random number $\text{rand}_{i,\mu}$ drawn in $(0,1)$, if $P_{C_\mu} < \text{rand}_{i,\mu}$, the update of $\mathbf{A}_{i,\mu}^t$ relies on (5) and (6)

$$P_{C_\mu} = \frac{e^{\theta_\mu} - e^{\theta_1}}{2(e^{\theta_q} - e^{\theta_1})}, \quad \theta_\mu = \frac{5(\mu - 1)}{q - 1}, \quad \mu = 1, \dots, q. \quad (13)$$

Second, if $P_{C_\mu} \geq \text{rand}_{i,\mu}$, the best row position information in history of particles except for the μ th particle will be employed to improve the corresponding rows of $\mathbf{P}_{\mu\text{pbest}}^t$ in (5). By this means, better row information (i.e., each endmember's abundance row vector) of other particles has chances to be learned and shared across the abundance swarm. To be specific, for each row of \mathbf{A}_μ^t , two randomly selected particles except for \mathbf{A}_μ^t itself are compared by a tournament selection

method, according to the original scalar fitness values in (3) of their best positions in history. Next, for the winner particle, if the i th row of its best position in history $\mathbf{P}_{i,\text{winpbest}}^t$ is not equal to $\mathbf{A}_{i,\mu\text{pbest}}^t$, $\mathbf{P}_{i,\text{winpbest}}^t$ is used to replace $\mathbf{A}_{i,\mu\text{pbest}}^t$ directly. Otherwise, $\mathbf{A}_{i,\mu\text{pbest}}^t$ is replaced by a random row $\mathbf{P}_{\text{rand},\text{winpbest}}^t$ ($\text{rand} \neq i$) of the winner particle's best position in history, which happens rarely but may work as a perturbation to help to increase the population diversity if premature convergence occurs. The velocity and the position of every particle in the abundance swarm are updated using comprehensive learning improved (5) and (6).

2) *Comprehensive Learning for Abundances' Sparsity*: The abovementioned dimension division for the abundance swarm depends on comparing the components of the fitness vectors (10) to determine each pixel vector slice in the abundance matrix. In this sense, the pixelwise local sparsity can be well learned, but the endmemberwise global sparsity may be not revealed correctly by the abundance swarm's particles. To address this issue, we devise another comprehensive learning strategy for improving the abundances' sparsity.

After the position of each abundance particle is updated using the first comprehensive learning strategy, $\mathbf{A}_{\text{gbest}}^t$ is obtained, and a new fitness function row vector $\phi_{A2}(\mathbf{A}_\mu)$ in (14) is formulated to generate an endmemberwise global best position $\mathbf{A}_{\text{gbest_row}}^t$

$$\begin{aligned} \phi_{A2}(\mathbf{A}_\mu) &= \left(\tilde{b}_1, \tilde{b}_2, \dots, \tilde{b}_r \right), \quad \tilde{b}_k \\ &= \left\| \mathbf{X} - \tilde{\mathbf{E}}_{\perp k} \tilde{\mathbf{A}}_{\perp k, \mu} \right\|_F^2 + \lambda_2 \cdot \tilde{\psi}_k(\mathbf{A}_\mu). \end{aligned} \quad (14)$$

In (14), $\tilde{\mathbf{E}}_{\perp k} \in \mathbb{R}^{m \times (r-1)}$ contains all but the k th column of \mathbf{E} , and $\tilde{\mathbf{A}}_{\perp k, \mu} \in \mathbb{R}^{(r-1) \times n}$ consists of all but the k th row of \mathbf{A}_μ^t . In most iterations, it is reliable that a small first term of \tilde{b}_k indicates that \mathbf{A}_μ^t is very likely to offer a good $\mathbf{A}_{k,\mu}^t$. λ_2 is a parameter controlling the abundances' row sparsity. $\tilde{\psi}_k(\mathbf{A}_\mu)$ can be written in (15) according to different sparse constraints

$$\tilde{\psi}_k(\mathbf{A}_\mu) = \begin{cases} \sum_{j=1}^n |(\mathbf{A}_\mu)_{kj}|, & \ell_1\text{norm} \\ \sum_{j=1}^n \sqrt{(\mathbf{A}_\mu)_{kj}}, & \ell_{1/2}\text{norm} \\ \sqrt{\sum_{j=1}^n (\mathbf{A}_\mu)_{kj}^2}, & \ell_{2,1}\text{norm} \end{cases}. \quad (15)$$

Next, $\mathbf{A}_{\text{gbest_row}}^t$ can be determined according to (14) and (15) using endmemberwise dimension division. In the iterations, it just works as an auxiliary for $\mathbf{A}_{\text{gbest}}^t$ to learn comprehensively the global sparsity of each endmember's abundances. Specifically, the sparse intensity $\rho_k(\mathbf{A}_{\text{gbest}}^t)$, $k = 1, \dots, r$, of each row of $\mathbf{A}_{\text{gbest}}^t$ is first calculated as follows:

$$\begin{aligned} \rho_k(\mathbf{A}_{\text{gbest}}^t) \\ = \frac{1 - \left(\sum_{j=1}^n (\mathbf{A}_{\text{gbest}}^t)_{kj} \right) / \left(\sum_{j=1}^n (\mathbf{A}_{\text{gbest}}^t)_{kj}^2 \right)}{1 - 1/n}. \end{aligned} \quad (16)$$

Base on (16), a row of $\mathbf{A}_{\text{gbest}}^t$ with lower sparse intensity is more likely to be selected by a roulette method [65]. Then, the

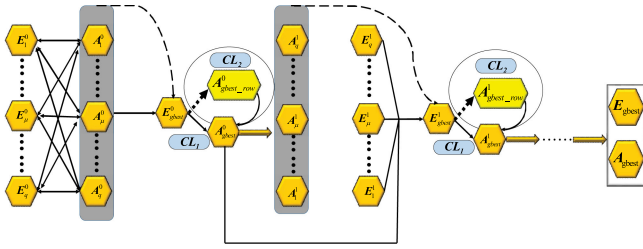


Fig. 1. Alternating update of dimension division based double swarm using comprehensive learning (CL) strategies CL_1 and CL_2 .

selected row of $\mathbf{A}_{\text{gbest}}^t$ learns from the corresponding row of $\mathbf{A}_{\text{gbest_row}}^t$ by replacing a part of its elements randomly with those of the latter. In this sense, the population diversity and the global sparsity of $\mathbf{A}_{\text{gbest}}^t$ can be improved. $\mathbf{A}_{\text{gbest_row}}^t$ is updated independently in each iteration according to the following operations:

$$\begin{cases} \mathbf{A}_{\mu\text{pbest_row}}^t(k, :) \\ = \begin{cases} \mathbf{A}_{\mu\text{pbest_row}}^{t-1}(k, :) \phi_{A2}(\mathbf{A}_{\mu\text{pbest_row}}^{t-1})_k \leq \phi_{A2}(\mathbf{A}_{\mu}^t)_k \\ \mathbf{A}_{\mu}^t(k, :) \phi_{A2}(\mathbf{A}_{\mu\text{pbest_row}}^{t-1})_k > \phi_{A2}(\mathbf{A}_{\mu}^t)_k \end{cases} \\ \mathbf{A}_{\text{gbest_row}}^t(k, :) \\ = \arg \min(\phi_{A2}(\mathbf{A}_{\text{lpbest_row}}^t)_k, \dots, \phi_{A2}(\mathbf{A}_{\text{qpbest_row}}^t)_k) \end{cases} \quad (17)$$

C. Alternating Update of Double Swarms for Unmixing

With the combination of dimension division and two comprehensive learning strategies, the endmembers and abundances are updated alternately in two collaborative swarms. As the iterations progress, the endmember swarm updates its particles and determine the global best position $\mathbf{E}_{\text{gbest}}^t$ according to (5), (6), (8), and (9), and then $\mathbf{E}_{\text{gbest}}^t$ serves as a known variable for the calculation of the particles' fitness values of the abundance swarm. In turn, the particles of the abundance swarm are updated using (5), (6), (10), (12) and two comprehensive learning strategies. $\mathbf{A}_{\text{gbest}}^t$ is then generated and employed by the endmember swarm. The abovementioned two subprocedures are executed alternately until the convergence conditions are satisfied. Fig. 1 illustrates the whole unmixing process of the proposed method.

In the initial step, the particles' positions of two swarms are randomly initialized. Their velocities are set as zeros at the beginning. Endmembers extracted by VCA and abundances estimated by FCLS are added into two swarms as elite particles to promote convergence. As shown in Fig. 1, the determination of initial $\mathbf{E}_{\mu\text{pbest}}^0$ and $\mathbf{E}_{\text{gbest}}^0$ depends on the cross computation of all the particles' positions of two swarms according to (8) and (9). Then, $\mathbf{E}_{\text{gbest}}^0$ is regarded as the given endmember matrix to produce $\mathbf{A}_{\mu\text{pbest}}^0$, $\mathbf{A}_{\text{gbest}}^0$, $\mathbf{A}_{\mu\text{pbest_row}}^0$, and $\mathbf{A}_{\text{gbest_row}}^0$ using (10), (12), (14), and (17), respectively. Moreover, some related PSO parameters should be set accordingly. 1) A damping and absorption boundary condition in [61] and [62] is imposed on the particles' positions to make the achieved endmembers and abundances nonnegative. By this means, if the particle position violates the boundary constraints in (3), it will be softly adjusted and go back to the specified range. Finally, the position of

Algorithm 1: Sparse Unmixing via Comprehensive learning based Particle Swarm Optimization (SUCPSO).

Input: Hyperspectral image $\mathbf{X} \in \mathbb{R}^{m \times n}$.

Output: Endmember matrix $\mathbf{E} \in \mathbb{R}^{m \times r}$ and abundance matrix $\mathbf{A} \in \mathbb{R}^{r \times n}$.

Step1: Setting $t = 0$ and generating an endmember swarm and an abundance swarm, initialize the positions of the two swarms and set the particle velocities as zeros, and determine $\mathbf{E}_{\text{gbest}}^0$, $\mathbf{A}_{\text{gbest}}^0$, $\mathbf{A}_{\text{gbest_row}}^0$.

Step2: While $t < t_{\text{max}}$ and stopping criteria are not satisfied, do $t \leftarrow t + 1$

2a) Using $\mathbf{A}_{\text{gbest}}^{t-1}$ to calculate ϕ_E and obtain $\mathbf{E}_{\mu\text{pbest}}^t$ and $\mathbf{E}_{\text{gbest}}^t$ according to (8) and (9).

2b) Using $\mathbf{E}_{\text{gbest}}^t$ to calculate ϕ_{A1} according to (10) and ϕ_{A2} by (14), and obtain $\mathbf{A}_{\mu\text{pbest}}^t$, $\mathbf{A}_{\text{gbest}}^t$, $\mathbf{A}_{\mu\text{pbest_row}}^t$ and $\mathbf{A}_{\text{gbest_row}}^t$ by (12) and (17).

2c) Adjusting $\mathbf{A}_{\text{gbest}}^t$ using two different comprehensive learning strategies in turn.

2d) Updating \mathbf{E}_{μ}^t and \mathbf{A}_{μ}^t according to (5) and (6), and controlling the ranges of particles' positions and velocities. Normalizing the abundance particles' positions to satisfy the ASC.

end

Step3: Outputting $\mathbf{E}_{\text{gbest}}$ and $\mathbf{A}_{\text{gbest}}$.

each particle in the abundance swarm is processed by a soft threshold operation and is normalized to satisfy the ASC. 2) The inertia weight ω in the velocity update process of the endmember swarm, and the learning factors c_1, c_2 are set as follows:

$$\begin{cases} \omega = \omega_{\text{max}} - \frac{t}{t_{\text{max}}}(\omega_{\text{max}} - \omega_{\text{min}}) \omega \in [0.4, 0.95] \\ c_1 = 2.5 - \frac{t}{t_{\text{max}}}, c_2 = 1.5 + \frac{t}{t_{\text{max}}} \end{cases} \quad (18)$$

3) In the abundance swarm, the inertia weight ω in the velocity update process is consistent with (18), and $c_1 = c_2 = 1.49445$. The settings of every particle's learning probability $P_{C_{\mu}}$ in the abundance swarm are illustrated in (13) referred to [63].

In this article, sparse unmixing via comprehensive learning-based particle swarm optimization (SUCPSO) can be briefly summarized in Algorithm 1. Under the SUCPSO unmixing framework, three unmixing algorithms considering different sparse constraints, including SUCPSO_L1soft, SUCPSO_L2,1, and SUCPSO_L1/2 can be obtained. Due to the ASC should be satisfied, SUCPSO_L1soft uses a soft threshold and the regularizers' division in (15) to make the abundances sparse. SUCPSO_L1/2 and SUCPSO_L2,1 adopt the regularizers' division in (11) and (15) to calculate their fitness vectors.

IV. EXPERIMENTAL RESULTS

In this section, we exploit simulated datasets and real hyperspectral images to verify the effectiveness of the proposed method dealing with different sparse constraints, i.e., SUCPSO_L1soft, SUCPSO_L2,1, and SUCPSO_L1/2. Comparison is conducted with several state-of-the-art sparse unmixing algorithms. Specifically, these algorithms include the

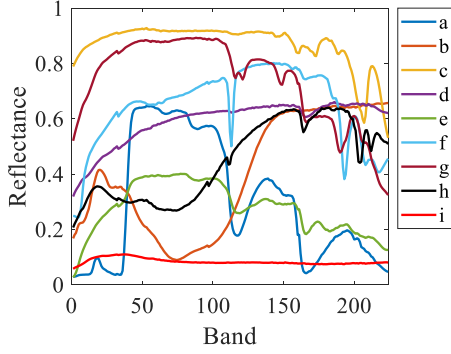


Fig. 2. Endmember spectra. (a) Maple_Leaves DW92-1. (b) Olivine GDS70.a Fo89 165 μm . (c) Calcite CO2004. (d) Quartz GDS74 Sand Ottawa. (e) Grass_dry.9+.1green AMX32. (f) Muscovite GDS107. (g) Alunite GDS82 Na82. (h) Uralite HS345.3B. (i) Pyrite HS35.3.

popular FCLS [26], two supervised sparse unmixing algorithms SUnSAL [27], CLSUnSAL [29], and two unsupervised sparse unmixing algorithms $\ell_{1/2}$ -NMF [38] and GLNMF [39]. In the following experiments, the regularization parameters of SUnSAL and CLSUnSAL are set as 10^{-4} according to their references. Equation (19) is used to generate the sparse regularization parameters λ for $\ell_{1/2}$ -NMF and GLNMF. In GLNMF, the graph regularization parameter is equal to 0.1

$$\lambda = \frac{1}{\sqrt{m}} \sum_{i=1}^m \frac{\sqrt{n} - \|\mathbf{X}_i\|_1 / \|\mathbf{X}_i\|_2}{\sqrt{n-1}}. \quad (19)$$

In (19), \mathbf{X}_i represents the i th row of the hyperspectral data matrix \mathbf{X} . The cross-validation approach is used to determine our algorithms' optimal parameters for the experiments with simulated data. In the experiments, VCA provided endmembers for supervised unmixing algorithms to estimate abundances. In the cases of unsupervised algorithms, endmembers and abundances were initialized by VCA and FCLS. The maximum number of iterations was set as 200. The algorithms were executed on MATLAB R2018b of a computer with a 3.80-GHz Intel Core i7-10700K CPU and 64 GBs of memory.

A. Simulated Data

After the parameter settings of the proposed method are first discussed, the robustness of the compared algorithms to noises, the number of endmembers, the mixing degrees, the number of pixels, and spectral variability are evaluated accordingly in the following experiments with several simulated datasets. Finally, the computational complexity of SUCPSO is analyzed, and the compared algorithms' running time is provided. In the generation of simulated data, endmembers were selected from the USGS spectral library in which each material spectrum consists of 224 bands. Their spectral curves can be clearly observed in Fig. 2. Abundances were randomly generated with the Gaussian random field method [66], and the largest abundance was 0.8 except for the second and the seventh experiments. Moreover, additive Gaussian white noises with different intensities were added to the simulated data.

In terms of quantitative evaluation, the mean spectral angle distance (MSAD) in (20) is employed to measure the similarity

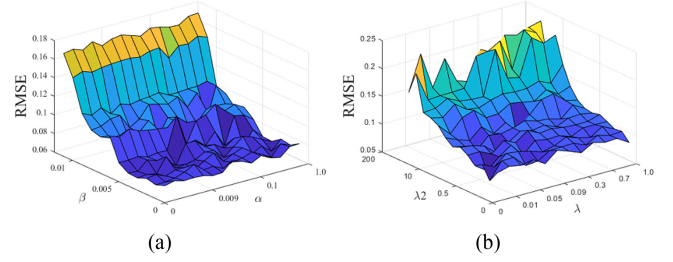


Fig. 3. Selection of parameters. (a) Threshold value α and the percentage of randomly replaced elements β . (b) Regularization parameters λ and λ_2 .

between true endmembers and estimated endmembers, and the root-mean-square error (RMSE) in (21) is adopted to measure the differences between true abundances and estimated abundances. Particularly, the sparseness of the estimated abundances is calculated by (22) referring to [64] and further quantitatively analyzed as the numbers of endmembers and pixels change. An absolute value of sparsity error (AVSE) in (23) is used to compare the estimated abundances' sparseness of the algorithms with the sparseness of true abundances

$$\text{MSAD} = \frac{1}{r} \sum_{i=1}^r \cos^{-1} \left(\frac{\mathbf{e}_i^T \hat{\mathbf{e}}_i}{\|\mathbf{e}_i\| \|\hat{\mathbf{e}}_i\|} \right) \quad (20)$$

$$\text{RMSE} = \sqrt{\frac{1}{rn} \sum_{i=1}^r \sum_{j=1}^n (\mathbf{A}_{ij} - \hat{\mathbf{A}}_{ij})^2} \quad (21)$$

$$\rho(\mathbf{A}) = \frac{1 - (\sum_{i=1}^r \sum_{j=1}^n \mathbf{A}_{ij}) / (\sum_{i=1}^r \sum_{j=1}^n (\mathbf{A}_{ij})^2)}{1 - 1/(rn)} \quad (22)$$

$$\text{AVSE} = \left| \rho(\hat{\mathbf{A}}) - \rho(\mathbf{A}) \right|. \quad (23)$$

In (20)–(23), $\mathbf{E} = [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_r]$ and $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n]$ denote the true endmember matrix and abundance matrix, respectively, while the estimated endmember matrix and abundance matrix are $\hat{\mathbf{E}} = [\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \dots, \hat{\mathbf{e}}_r]$ and $\hat{\mathbf{A}} = [\hat{\mathbf{a}}_1, \hat{\mathbf{a}}_2, \dots, \hat{\mathbf{a}}_n]$, respectively. Each compared algorithm had ten independent runs. Their unmixing results' averages and standard deviations are provided.

1) *Parameter Settings*: In order to determine the proposed method's four parameters including two regularization parameters λ and λ_2 , a threshold value α , and the percentage of randomly replaced elements β in the second comprehensive learning strategy, we generated a simulated dataset containing 50×50 pixels using five endmembers for cross validation. The maximum value of abundances is set as 0.8, and signal to noise ratios ($\text{SNR} = 10 \log_{10}(E[\mathbf{x}^T \mathbf{x}] / E[\varepsilon^T \varepsilon])$) is 40 dB. α and β are first evaluated by using fixed λ and λ_2 , then the best λ and λ_2 are determined. RMSE was used as an indicator to compare the parameters. According to the results depicted in Fig. 3, α and β were set as 0.01 and 0.004, respectively, and $\lambda = 0.005$, $\lambda_2 = 0.8$ in the simulated data experiments.

2) *Noise Robustness Analysis*: In order to analyze the proposed algorithm's robustness to noises, we generated datasets with different SNRs including 20 dB, 30 dB, 40 dB, 50 dB, and

TABLE I

MSADs OF ENDMEMBERS EXTRACTED BY THE ALGORITHMS FOR SIMULATED DATA WITH DIFFERENT NOISE INTENSITIES

SNR(dB)	VCA	$\ell_{1/2}$ -NMF	GLNMF	SUCPSO_	SUCPSO_	SUCPSO_
				L1soft	L2,1	L1/2
20	5.041±1.109	3.983±0.532	3.981±0.532	4.036±0.914	4.225±1.002	3.798±1.070
30	4.714±0.555	3.681±0.433	3.702±0.440	3.535±0.797	3.793±0.669	3.406±0.802
40	4.004±0.909	3.425±0.551	3.407±0.568	3.466±0.794	3.710±0.611	3.183±0.841
50	4.003±0.903	3.288±0.838	3.282±0.840	3.079±0.839	3.426±0.596	2.944±1.005
INF	3.903±0.734	3.246±0.969	3.264±1.017	3.083±0.805	3.327±0.654	2.923±0.677

TABLE II

RMSES OF ABUNDANCES ESTIMATED BY THE ALGORITHMS FOR SIMULATED DATA WITH DIFFERENT NOISE INTENSITIES

SNR(dB)	FCLS	SUnSAL	CLSUnSAL	$\ell_{1/2}$ -NMF	GLNMF	SUCPSO_	SUCPSO_	SUCPSO_
						L1soft	L2,1	L1/2
20	0.438±0.044	0.457±0.045	0.461±0.047	0.470±0.044	0.470±0.044	0.153±0.066	0.162±0.066	0.149±0.070
30	0.373±0.135	0.349±0.142	0.349±0.142	0.379±0.139	0.379±0.138	0.110±0.058	0.132±0.072	0.106±0.061
40	0.344±0.138	0.342±0.141	0.342±0.141	0.339±0.137	0.338±0.137	0.100±0.047	0.106±0.048	0.098±0.048
50	0.320±0.167	0.318±0.176	0.318±0.176	0.331±0.175	0.330±0.175	0.101±0.072	0.085±0.042	0.088±0.043
INF	0.305±0.157	0.308±0.155	0.307±0.156	0.315±0.157	0.314±0.156	0.075±0.015	0.080±0.015	0.078±0.024

INF dB (i.e., no noises). The simulated data have 50×50 pixels constituted by five endmembers.

The MSADs of endmembers extracted by every algorithm are given in Table I. The best results are marked in bold. From this table, one can find that the MSADs corresponding to our methods are much smaller than the other algorithms' in most situations. Notably, two algorithms with $\ell_{1/2}$ norm provide better endmembers. It seems that $\ell_{1/2}$ norm may improve endmember extraction compared with other sparse constraints. SUCPSO_L1/2 provides the best MSADs as the SNR changes, proving the superiority of sparse unmixing under our improved PSO framework. In the case of the noise-free dataset, all the compared algorithms can obtain more accurate endmembers. SUCPSO_L1soft and SUCPSO_L2,1 also show their higher accuracies. On the other hand, from the RMSEs listed in Table II, it can be observed that our method with different sparse constraints can achieve considerably smaller values among these algorithms. Particularly, the abundances estimated by SUCPSO_L2,1 approach true abundances with minor errors. A reasonable explanation can be derived that the introduction of dimension division and comprehensive learning in PSO greatly facilitate unsupervised sparse unmixing to find better abundances. From these experiments, we can infer that the proposed method is robust to noises and can produce accurate unmixing results.

3) *Sensitivity Analysis to the Number of Endmembers*: In this experiment, we analyze the sensitivity of different algorithms to the number of endmembers. The size of the dataset is 50×50 -pixel, SNR equals 40 dB, and the number of endmembers was set as 3, 5, 7, and 9, respectively. Tables III and IV display the MSADs and RMSEs of the unmixing results, respectively. It can be seen that although the unmixing results of our algorithms are slightly affected by the number of endmembers, the estimated abundances are always closer to the references. In the cases of the number of endmembers being 3 or 5 when the data's sparsity may be not prominent, the

TABLE III

MSADs OF ENDMEMBERS EXTRACTED BY THE ALGORITHMS FOR SIMULATED DATA WITH DIFFERENT NUMBERS OF ENDMEMBERS

Number of endmembers	VCA	$\ell_{1/2}$ -NMF	GLNMF	SUCPSO_	SUCPSO_	SUCPSO_
				L1soft	L2,1	L1/2
3	7.659±0.149	4.497±1.103	4.451±1.101	3.393±0.965	3.520±0.926	3.387±0.889
5	4.004±0.909	3.425±0.551	3.407±0.568	3.466±0.794	3.710±0.611	3.183±0.841
7	4.502±0.560	4.109±0.634	4.112±0.638	4.085±0.897	4.268±0.784	4.303±0.679
9	4.527±0.120	4.540±0.110	4.546±0.110	4.482±0.411	4.467±0.316	4.566±0.113

TABLE IV

RMSEs OF ABUNDANCES ESTIMATED BY THE ALGORITHMS FOR SIMULATED DATA WITH DIFFERENT NUMBERS OF ENDMEMBERS

Number of endmembers	FCLS	SUnSAL	CLSUnSAL	$\ell_{1/2}$ -NMF	GLNMF	SUCPSO_	SUCPSO_	SUCPSO_
						L1soft	L2,1	L1/2
3	0.309±0.220	0.342±0.231	0.342±0.231	0.311±0.222	0.311±0.221	0.041±0.014	0.043±0.013	0.042±0.015
5	0.344±0.138	0.342±0.141	0.342±0.141	0.339±0.137	0.338±0.137	0.100±0.047	0.106±0.048	0.098±0.048
7	0.339±0.086	0.343±0.088	0.343±0.088	0.345±0.088	0.345±0.088	0.110±0.035	0.110±0.037	0.125±0.044
9	0.357±0.045	0.357±0.045	0.357±0.045	0.364±0.043	0.364±0.043	0.105±0.023	0.112±0.037	0.100±0.014

TABLE V

MSADs OF ENDMEMBERS EXTRACTED BY THE ALGORITHMS FOR SIMULATED DATA WITH DIFFERENT MIXING DEGREES

Maximum abundance	VCA	$\ell_{1/2}$ -NMF	GLNMF	SUCPSO_	SUCPSO_	SUCPSO_
				L1soft	L2,1	L1/2
0.7	6.084±0.516	5.544±1.094	5.540±1.100	4.381±1.005	4.246±1.026	4.072±0.767
0.8	4.004±0.909	3.425±0.551	3.407±0.568	3.466±0.794	3.710±0.611	3.183±0.841
0.9	1.977±0.297	2.315±0.443	2.311±0.439	1.313±0.507	1.403±0.326	1.154±0.655
1.0	0.146±0.023	1.156±0.370	1.148±0.367	0.245±0.194	0.241±0.264	0.102±0.036

TABLE VI

RMSEs OF ABUNDANCES ESTIMATED BY THE ALGORITHMS FOR SIMULATED DATA WITH DIFFERENT MIXING DEGREES

Maximum abundance	FCLS	SUnSAL	CLSUnSAL	$\ell_{1/2}$ -NMF	GLNMF	SUCPSO_	SUCPSO_	SUCPSO_
						L1soft	L2,1	L1/2
0.7	0.358±0.072	0.375±0.068	0.375±0.068	0.363±0.085	0.364±0.084	0.148±0.037	0.169±0.049	0.155±0.049
0.8	0.344±0.138	0.342±0.141	0.342±0.141	0.339±0.137	0.338±0.137	0.100±0.047	0.106±0.048	0.098±0.048
0.9	0.306±0.219	0.308±0.221	0.308±0.221	0.322±0.221	0.322±0.221	0.042±0.005	0.047±0.009	0.040±0.005
1.0	0.289±0.148	0.286±0.149	0.286±0.149	0.301±0.151	0.301±0.151	0.013±0.003	0.013±0.004	0.012±0.003

best MSADs can be attained by SUCPSO_L1/2. SUCPSO_L2,1 performs better when the number of endmembers is 9. It is indicated that different sparse constraints may suffer from the number of endmembers. In terms of abundances' accuracy, the proposed method can obtain the best results due to the use of comprehensive learning strategies, which can effectively deal with complex data with more endmembers.

4) *Robustness Analysis to the Mixing Degrees*: The sensitivity of each algorithm to different mixing degrees of data is discussed in this experiment. The simulated dataset is composed of five endmembers, and contains 50×50 pixels. SNR is 40 dB, and the maximum value of abundances is set as 0.7, 0.8, 0.9, and 1.0, respectively. The description of unmixing accuracies is provided in Tables V and VI.

Table V shows the MSADs of all the algorithms under different mixing degrees. It is clear that the endmembers' accuracies of the algorithms decrease as the data become highly mixed. However, endmembers extracted by SUCPSO_L1/2 are more accurate than the other algorithms' in most situations.

TABLE VII
MSADs OF ENDMEMBERS EXTRACTED BY THE ALGORITHMS FOR SIMULATED DATA WITH DIFFERENT NUMBERS OF PIXELS

Number of pixels	VCA	$\ell_{1/2}$ -NMF	GLNMF	SUCPSO_ _{L1soft}	SUCPSO_ _{L2,1}	SUCPSO_ _{L1/2}
900	4.173±0.489	4.074±1.002	4.074±0.947	3.284±0.928	3.972±1.657	3.069±1.064
2500	4.004±0.909	3.425±0.551	3.407±0.568	3.466±0.794	3.710±0.611	3.183±0.841
3600	3.960±0.623	3.418±0.596	3.414±0.630	2.301±0.455	2.594±0.446	2.133±0.493
4900	4.369±0.467	3.580±0.152	3.601±0.148	2.842±0.537	2.919±0.523	2.714±0.414

TABLE VIII
RMSEs OF ABUNDANCES ESTIMATED BY THE ALGORITHMS FOR SIMULATED DATA WITH DIFFERENT NUMBERS OF PIXELS

Number of pixels	FCLS	SUnSAL	CLSUnSAL	$\ell_{1/2}$ -NMF	GLNMF	SUCPSO_ _{L1soft}	SUCPSO_ _{L2,1}	SUCPSO_ _{L1/2}
900	0.296±0.102	0.298±0.099	0.298±0.099	0.297±0.107	0.297±0.106	0.080±0.016	0.100±0.019	0.088±0.005
2500	0.344±0.138	0.342±0.141	0.342±0.141	0.339±0.137	0.338±0.137	0.100±0.047	0.106±0.048	0.098±0.048
3600	0.326±0.125	0.328±0.121	0.328±0.121	0.333±0.127	0.332±0.127	0.070±0.017	0.076±0.017	0.072±0.019
4900	0.335±0.113	0.339±0.131	0.339±0.131	0.348±0.122	0.348±0.122	0.098±0.015	0.102±0.013	0.095±0.019

In Table VI, the overall RMSEs of SUCPSO_L1/2's abundances are also the smallest. Similar to the first experiment, the differences among sparse regularization terms in unmixing are present in this experiment. Remarkably, under the same sparse unmixing frameworks, the proposed method enables more accurate abundances to be achieved through the optimization implemented by the comprehensive-learning-based double swarm PSO. The proposed method's robustness to the mixing degree of data is verified in this experiment.

5) *Sensitivity Analysis to the Number of Pixels*: The goal of this experiment is to study the influence of the data's size on the algorithms. Five endmembers were used to construct several datasets consisting of 30×30 , 50×50 , 60×60 , and 70×70 pixels, respectively. SNR is 40 dB. Tables VII and VIII provide the unmixing results of every algorithm. It can be observed that our method always outperforms other sparse unmixing algorithms in endmember extraction and abundance inversion. Moreover, the proposed method considering different sparse constraints is not significantly affected by the number of pixels. SUCPSO_L1/2 gives the best unmixing results among all the algorithms. MSADs and RMSEs of SUCPSO_L1soft and SUCPSO_L2,1 are also smaller than those of gradient methods optimized sparse unmixing algorithms. It can be concluded that PSO has superior optimization ability, and reasonable sparse constraints and learning strategies make it more suitable for unmixing. In a word, the proposed method is not sensitive to the changes of the number of pixels and has outstanding unmixing performance compared with other typical sparse unmixing algorithms.

6) *Quantitative Evaluation of Abundances' Sparseness*: RMSE is a straightforward metric to illustrate the accuracy of the estimated abundances. However, in terms of sparse unmixing, there is also a requirement that the sparseness of abundances should be evaluated. To this end, AVSE in (23) is employed to conduct a deep comparison of the sparse unmixing algorithms. It is noted that the abundances with the highest degree of sparsity may be not the best solution. The difference between

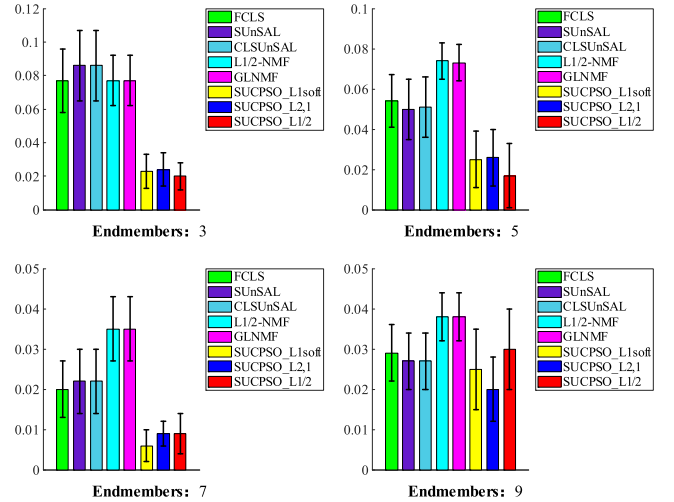


Fig. 4. Comparison of AVSEs for datasets with different number of endmembers.

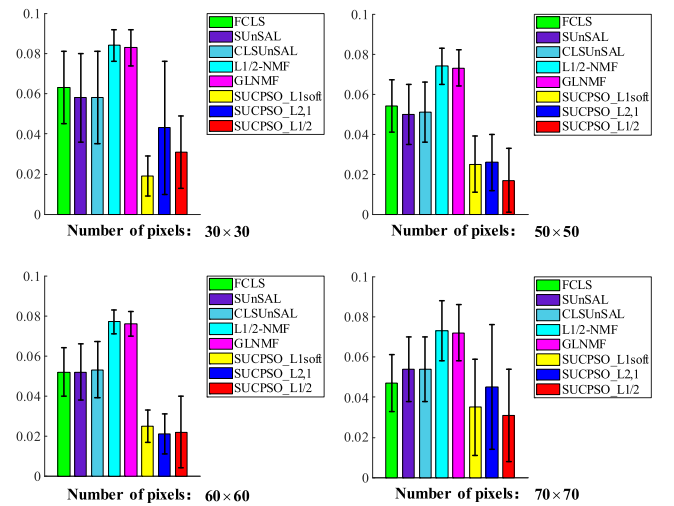


Fig. 5. Comparison of AVSEs for datasets with different number of pixels.

the sparseness of the estimated abundances and true abundances indicated by AVSE can better evaluate the abundances.

Since the sparsity of the abundances can be related to the number of endmembers and pixels, the AVSEs corresponding to each algorithm are analyzed as the number of endmembers and the number of pixels changes. In Figs. 4 and 5, the AVSEs of the algorithms are compared. From these two figures, one can easily conclude that the abundances generated by the proposed method are very likely to have the similar sparseness of true abundances. All the algorithms have similarly large AVSEs in the case of three endmembers. This is because the sparseness seems to be not apparent when endmembers are few. When endmembers become more, AVSEs of SUCPSO_L1soft and SUCPSO_L2,1 are smaller than the other algorithms', indicating that their abundances are more consistent with true abundances. Moreover, when the number of pixels changes, the AVSE values of our method are still smaller than the gradient methods optimized sparse unmixing algorithms. It can be inferred that two comprehensive learning strategies can express the abundances'

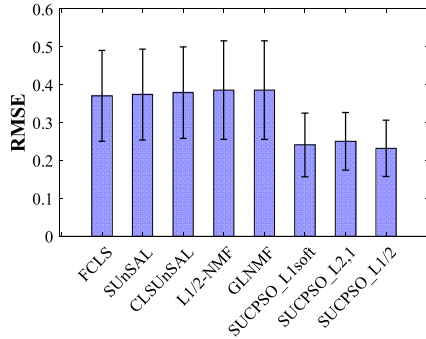


Fig. 6. Comparison of the algorithms' abundance estimation accuracy for simulated data with spectral variability.

inherent sparsity correctly. This advantage enables the proposed algorithm to have acceptable superiority and stability in the implementation of sparse unmixing.

7) *Sensitivity Analysis to Spectral Variability*: It is known that the existence of noises, nonlinear mixing effects, and endmembers' spectral variabilities in hyperspectral images may affect the unmixing process simultaneously. Therefore, we further generated simulated data with endmember perturbation to give a possible study on how the spectral variabilities influence unsupervised sparse unmixing and the proposed method. The simulated data were constructed referring to [67]. Specifically, five selected endmembers constitute a reference endmember set. Each reference endmember multiplied a parameter randomly drawn from the interval $[0.75, 1.25]$ for each pixel, and the scaled endmembers were further deteriorated by a 25 dB white Gaussian noise. Abundances were generated like the above experiments and the largest abundance is set as 1. Then, the simulated data were obtained following the LMM, and SNR is set as 40 dB.

Due to the existence of spectral variability, we have only compared the RMSEs of the algorithms' estimated abundances in Fig. 6. In contrast to the results listed in Table VI, it can be observed that the overall unmixing accuracies of all the algorithms decrease a lot even if pure pixels exist in this case. Remarkably, the proposed method SUCPSO can still provide the most accurate abundances. However, it also seems to suffer from the data complexity induced by spectral variability, indicating that the research of spectral variability in sparse unmixing can be meaningful in practice.

8) *Computation Complexity and Execution Time*: Since three proposed unmixing algorithms with different sparse constraints are based on Algorithm 1, they have very similar computational complexity. We can directly analyze the computational cost of Algorithm 1. During the process of unmixing, the floating-point calculation in each iteration mainly comes from the update of particles' positions and the computation and comparison of particles' fitness values. Given a hyperspectral image containing m bands, n pixels, and r endmembers, each swarm contains q particles, in terms of worst-case complexities, the computational complexity of two swarms' above three basic procedures is $O(qmnr + q^2(m + n))$. In addition, the time complexity of two comprehensive learning strategies for the abundance swarm is

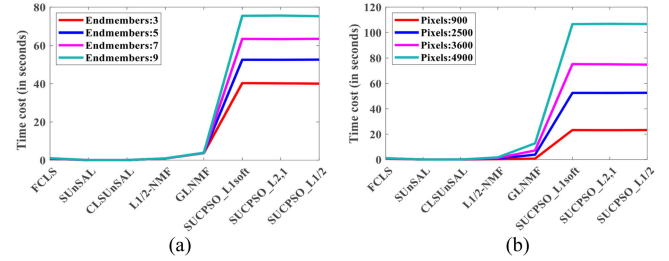


Fig. 7. Comparison of the algorithms' execution time. Results of simulated data with (a) different numbers of endmembers, and (b) different numbers of pixels.

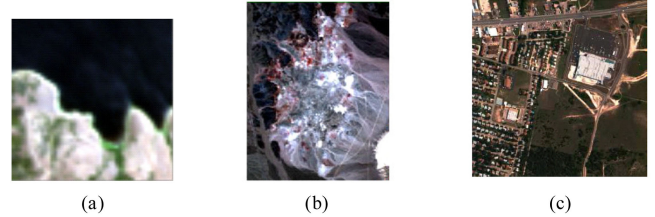


Fig. 8. Real hyperspectral images for experiments. (a) Moffett Field. (b) Cuprite. (c) Urban.

TABLE IX
COMPARISON OF SADS OF ENDMEMBERS EXTRACTED BY THE ALGORITHMS FOR MOFFETT FIELD DATA

Endmember	VCA	$\ell_{1/2}$ -NMF	GLNMF	SUCPSO_	SUCPSO_	SUCPSO_
				L1soft	L2,1	L1/2
Vegetation	4.731	9.352	9.351	4.563	4.575	10.619
Water	25.934	15.000	14.993	16.428	15.567	14.879
Soil	2.520	3.562	3.554	0.966	0.972	1.594
Average	11.062	9.305	9.299	7.319	7.038	9.031

$O(qmnr + q^2r)$. As a result, the total computational complexity of SUCPSO turns to be $O(qmnr + q^2(m + n + r))$. The computational complexities of the other compared algorithms can be found in their corresponding references.

Moreover, the execution time is provided in Fig. 7 when the simulated data contain different numbers of endmembers and pixels. We can observe that the proposed method should spend much more time getting the unmixing results than the other compared algorithms. It implies that the improvement of SUCPSO's efficiency is an essential step to achieve higher computational performance.

B. Real Hyperspectral Images

Resorting to the ideal simulated data, the abovementioned experiments have provided convincing quantitative evaluation for the sparse unmixing algorithms. In this section, experiments based on three real hyperspectral images are further carried out to complement the validation of the proposed method in practice. Fig. 8 shows these three images, which are popularly used for the evaluation of unmixing algorithms in the literature [68]–[70]. Therefore, partial ground truth and some related unmixing results can be referred to.

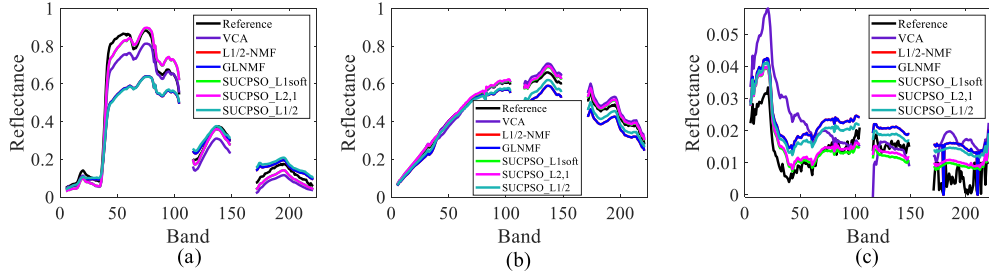


Fig. 9. Spectral curves of endmembers extracted by the compared algorithms for the Moffett Field data. (a) Vegetation. (b) Soil. (c) Water.

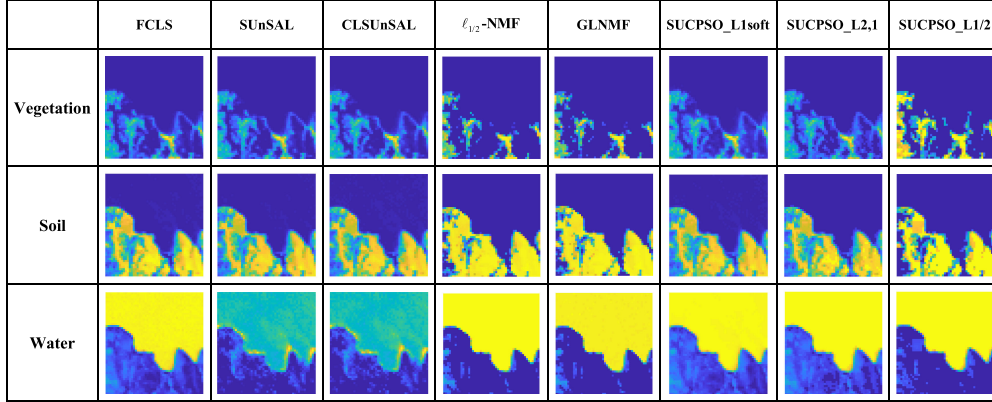


Fig. 10. Abundance maps of three major ground covers in the Moffett Field data.

TABLE X
RES OF THE UNMIXING RESULTS OF THE ALGORITHMS FOR REAL HYPERSPECTRAL IMAGES

Dataset	FCLS	SUnSAL	CLSUnSAL	$\ell_{1/2}$ -NMF	GLNMF	SUCPSO_L1soft	SUCPSO_L2,1	SUCPSO_L1/2
Moffett Field	0.013	0.012	0.012	0.028	0.028	0.010	0.019	0.018
Cuprite	0.008	0.005	0.005	0.007	0.007	0.005	0.004	0.005
Urban	0.065	0.065	0.064	0.048	0.048	0.028	0.023	0.033

TABLE XI
RUN TIME (IN SECONDS) OF THE ALGORITHMS FOR REAL HYPERSPECTRAL IMAGES

Dataset	FCLS	SUnSAL	CLSUnSAL	$\ell_{1/2}$ -NMF	GLNMF	SUCPSO_L1soft	SUCPSO_L2,1	SUCPSO_L1/2
Moffett Field	0.140	0.050	0.023	0.747	3.146	34.903	35.072	35.248
Cuprite	8.179	0.706	0.393	16.769	739.730	1775.747	1772.003	1791.423
Urban	5.528	0.758	0.913	26.640	3878.112	1704.247	1702.237	1706.846

TABLE XII
COMPARISON OF SADs (IN DEGREES) OF ENDMEMBERS EXTRACTED BY THE ALGORITHMS FOR CUPRITE DATA

Endmember	VCA	$\ell_{1/2}$ -NMF	GLNMF	SUCPSO_L1soft	SUCPSO_L2,1	SUCPSO_L1/2
Alunite	18.505	16.437	16.536	13.745	13.315	14.455
Sphene	5.373	3.223	3.254	2.963	3.183	2.749
Kaolinite	8.706	6.594	6.580	6.992	6.956	6.975
Montmorillonite	3.516	3.037	3.034	3.117	2.872	2.960
Kaolinite#2	8.274	6.019	6.001	6.419	6.365	6.363
Buddingtonite	5.802	5.816	5.823	5.821	5.988	5.951
Pyrope	3.002	3.633	3.605	3.890	3.612	4.064
Nontronite	7.176	6.021	6.005	5.904	6.055	6.009
Muscovite	6.450	6.222	6.217	6.614	6.612	6.276
Halloysite	12.497	10.088	10.119	9.675	9.645	9.642
Chalcedony	4.941	2.997	2.985	3.302	3.236	3.313
Desert Varnish	8.620	6.463	6.489	5.620	5.742	5.966
Average	7.738	6.379	6.387	6.172	6.132	6.227

Since true abundances are not available here, RMSE of abundances cannot be obtained. Abundance maps generated by each algorithm are compared qualitatively. In addition, two performance metrics are adopted for quantitative evaluation. Spectral angular distance (SAD) [1] defined in (24) is used to measure the similarity between an estimated endmember and its corresponding reference endmember. Reconstruction error (RE) in (25) of unmixing results is also considered

$$\text{SAD} = \cos^{-1} \left(\frac{\mathbf{e}^T \hat{\mathbf{e}}}{\|\mathbf{e}\| \|\hat{\mathbf{e}}\|} \right) \quad (24)$$

$$\text{RE} = \sqrt{\frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n (\mathbf{X}_{ij} - \hat{\mathbf{X}}_{ij})^2}. \quad (25)$$

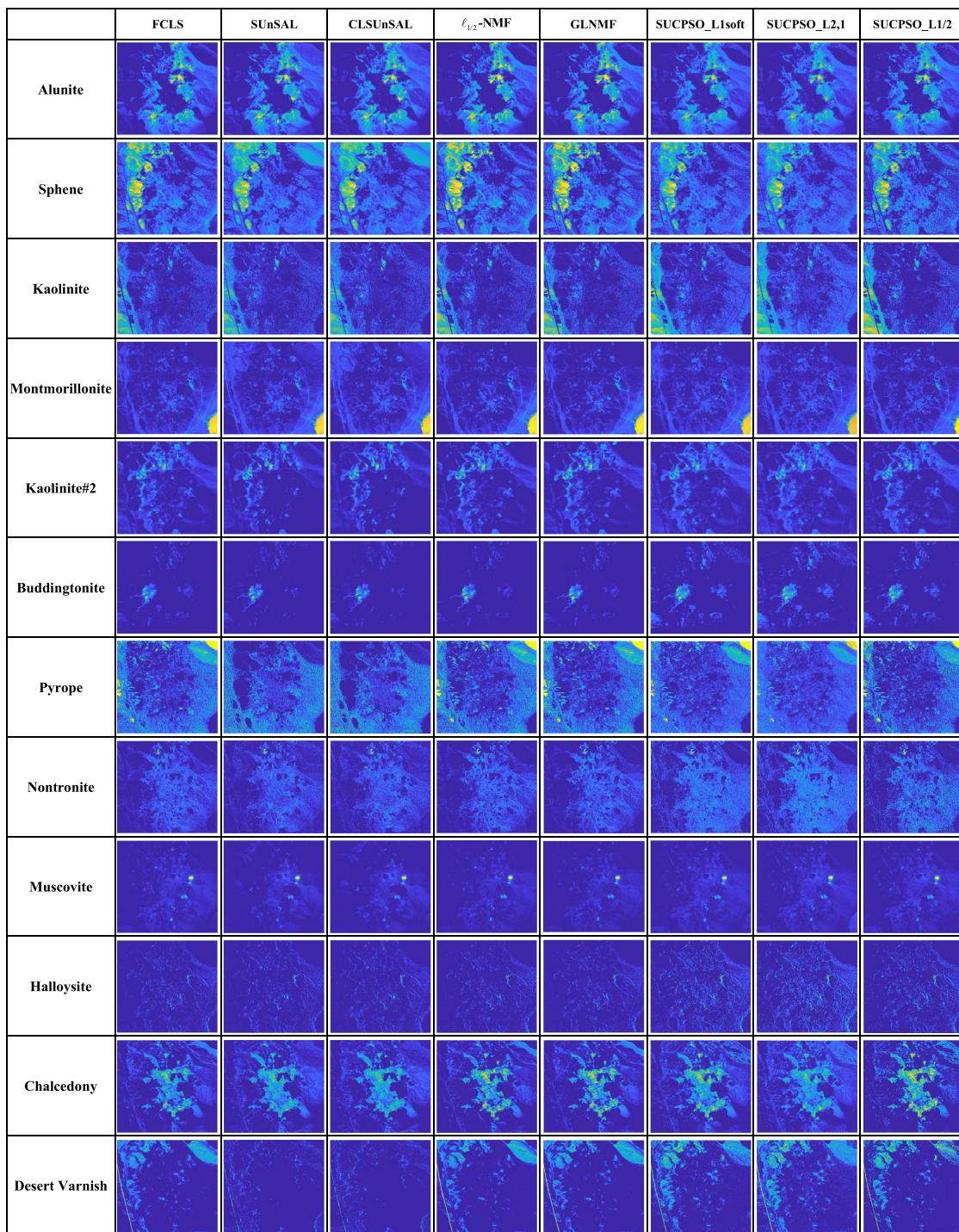


Fig. 11. Abundance maps of twelve endmembers in the observed region of Cuprite data.

1) *Moffett Field Data*: The first image was obtained in 1997 by the Airborne Visible Infrared Imaging Spectrometer (AVIRIS) over Moffett Field at the South end of San Francisco, California. Its spatial resolution is 20 m, and each pixel contains 224 bands with a wavelength range from 0.4 to 2.5 μm . As shown in Fig. 8(a), the dataset for the experiments is a selected 50×50 -pixel subregion of the original image. After removing the water vapor absorption bands and bands corrupted heavily

by noises, we finally retained 185 bands (5–104, 116–149, 171–221) for unmixing. The subregion mainly contains vegetation, water, and soil. In the experiments, reference endmembers used for accuracy evaluation were manually selected from the data. The threshold value and λ are respectively 0.1 and 0.55 in SUCPSO_L2,1 and SUCPSO_L1/2. In SUCPSO_L1soft, the threshold value is set as 0.001. Other parameters remain the same as the simulated data experiments.

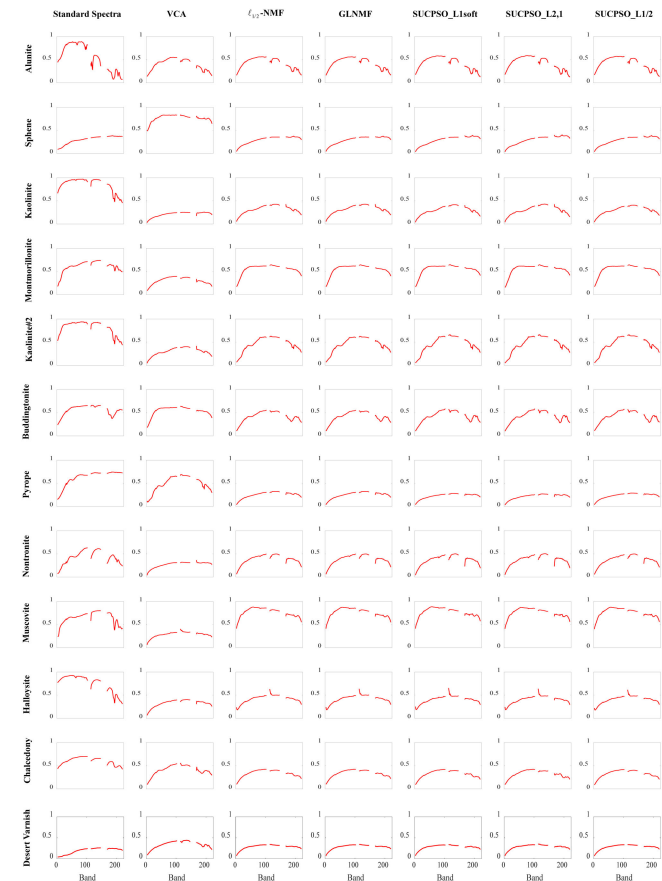


Fig. 12. Spectral curves of extracted endmembers by the compared algorithms for Cuprite data.

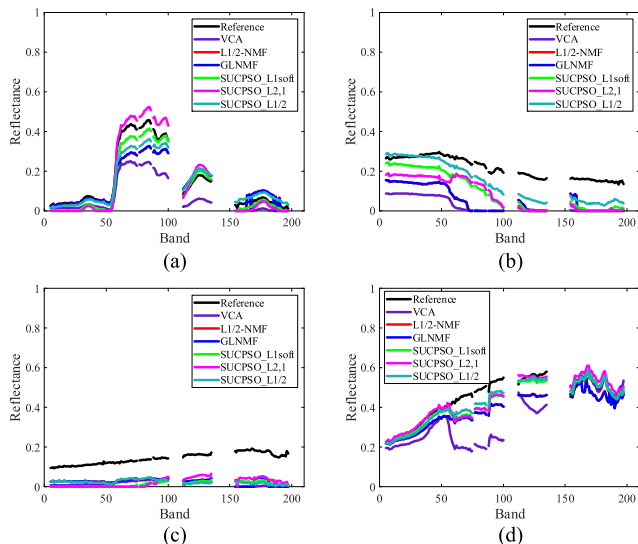


Fig. 13. Spectral curves of endmembers extracted by the compared algorithms for Urban data. (a) Vegetation. (b) Roof. (c) Asphalt. (d) Dirt.

Table IX compares the accuracy of six algorithms' estimated endmembers of vegetation, water, and soil. In this table, our method considering different sparse constraints performs better than VCA and the other two unsupervised sparse unmixing

algorithms. SUCPSO_L1soft gives the best results in the extraction of vegetation and soil. The water endmember identified by SUCPSO_L1/2 has a minor SAD. In Fig. 9, the spectral curves of every algorithm's endmembers and reference endmembers are visually compared. It can be observed that our method's endmember curves are much similar to the reference endmember curves, proving that accurate endmembers can be determined during the unsupervised sparse unmixing process of the proposed method.

In terms of the REs listed in Table X, SUCPSO_L1soft provides the smallest value, which illustrates the effectiveness of PSO unmixing framework indirectly. Table XI shows the execution time of eight algorithms for unmixing three real hyperspectral images. The proposed method with different sparse constraints costs more time than the other sparse unmixing algorithms. It implies that the improvement of computational efficiency may be a necessary task for swarm intelligence algorithms like PSO. Fig. 10 displays the estimated abundance maps. Vegetation and soil are clearly recognized by almost all the algorithms. However, large errors seem to occur in the abundance maps of SUnSAL and CLSUnSAL. Abundance maps of the proposed method with three sparse regularizers are considered to be in line with the published results. Moreover, the sparseness of ground covers' abundances can be revealed by SUCPSO_L1/2.

2) *Cuprite Data*: The second image was collected by the AVIRIS over a copper-iron mine named Cuprite in Nevada on June 19, 1997. Its spectral resolution is about 10 nm, and its spatial resolution is 20 m. In the wavelength range from 0.4 to 2.5 μm , 224 bands were continuously sampled. Fig. 8(b) displays a selected subimage containing 200×200 pixels of the whole dataset. In this area, various minerals such as alunite and kaolinite widely exist. In this experiment, spectral bands that are severely affected by noises were removed before unmixing, and 188 bands (3–103, 114–147, 168–220) were used for unmixing. According to the previous unmixing results in the literature, the number of endmembers was set as 12 by ignoring some rare materials. The standard mineral spectra in the USGS spectral library are used as reference endmembers for evaluating the extracted endmembers. The threshold value and λ of the proposed method are 0.001 and 0.01. Other parameters remain the same as the simulated data experiments.

Similar to the first experiment in this section, Table XII shows the SADs of each algorithm, which measures the similarity between the estimated endmembers and the reference endmembers. It is clear that SUCPSO_L1soft, SUCPSO_L2,1, and SUCPSO_L1/2 have obtained the best results of two kinds of materials, respectively. Although GLNMF has accurately estimated four endmembers, its average SAD is slightly worse than our method. Since many minerals presented in this area are considered to be highly mixed, VCA cannot provide accurate endmembers as the other algorithms in most situations though buddingtonite and pyrope have been accurately extracted.

In Fig. 11, abundance maps of twelve materials are displayed in detail. Abundance maps corresponding to pyrope and desert varnish obtained by SUnSAL and CLSUnSAL seem to be wrongly estimated. In comparison, the unmixing results of the proposed method can clearly reflect the actual distribution of

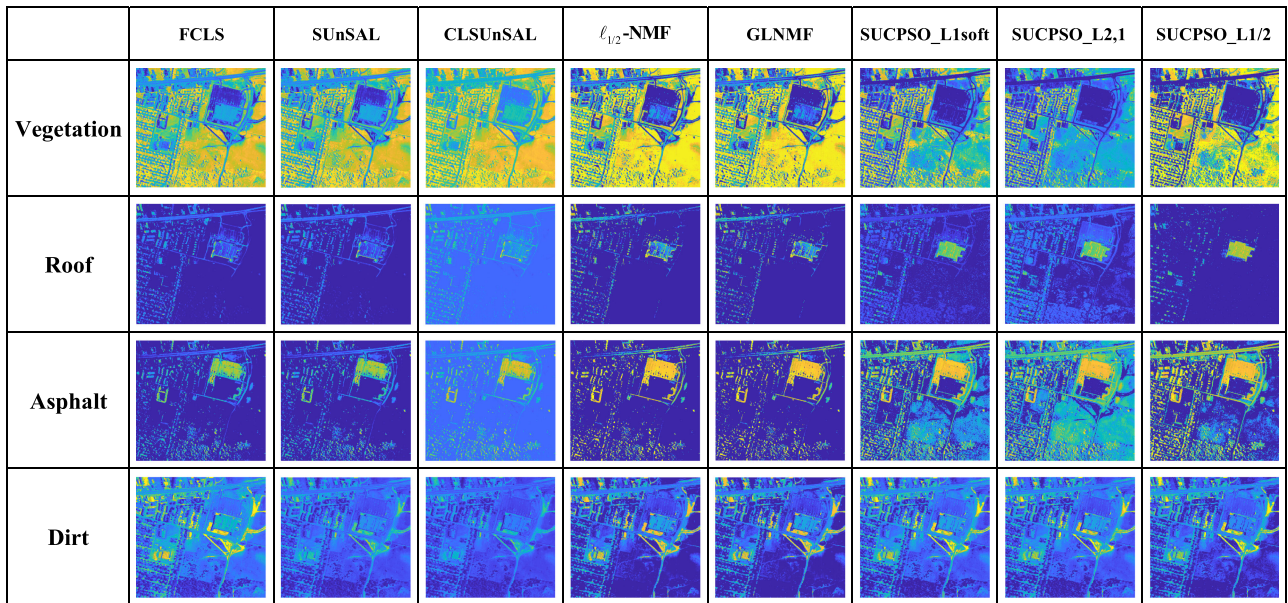


Fig. 14. Abundance maps of four endmembers in the Urban data.

12 materials. Fig. 12 compares the spectral curves of various materials extracted by six algorithms with standard spectra. The REs of each algorithm are listed in Table X. Table XI shows the time costs. One can see that the endmembers extracted by the proposed method are very close to their corresponding standard spectra. Moreover, in terms of RE, our approach also performs better than the other algorithms in all situations where different sparse constraints are considered.

3) *Urban Data*: The third image was collected by the Hyperspectral Digital Imagery Collection Experiment over an urban area at Copperas Cove, Texas, USA. Its spectral resolution is about 10 nm, and its spatial resolution is 2 m. In the wavelength range from 0.4 to 2.4 μm , 210 bands were continuously sampled. Fig. 8(c) displays the image containing 307×307 pixels of the whole dataset. Several types of natural round covers and artificial buildings can be clearly observed in the area. This hyperspectral image has been widely used for hyperspectral unmixing in the literature [36]–[38]. According to these existing works, the number of endmembers was set as four in this experiment. Specifically, vegetation, roof, asphalt, and dirt have been identified by the algorithms. It is known that the spatial distribution of these materials has strong sparsity, which makes this image suitable for the evaluation of sparse unmixing methods. Spectral bands that are severely affected by water absorption and atmospheric effects were removed before unmixing, and 162 bands (5–75, 77–86, 88–100, 112–135, 154–197) were used for unmixing. The threshold value α and λ of the proposed method are 0.1 and 0.4. Other parameters remain the same as the simulated data experiments.

According to the endmember extraction accuracies displayed in Table XIII, we can see that SUCPSO_L1/2 has obtained the best endmembers including roof, asphalt, and dirt. The SAD of vegetation endmember extracted by SUCPSO_L1soft is the smallest. $\ell_{2,1}$ -NMF and GLNMF have also recognized the dirt accurately. The RE and the time cost of each algorithm are

TABLE XIII
COMPARISON OF SADS OF ENDMEMBERS EXTRACTED BY THE ALGORITHMS FOR URBAN DATA

Endmember	VCA	$\ell_{1/2}$ -NMF	GLNMF	SUCPSO_L1soft	SUCPSO_L2,1	SUCPSO_L1/2
Vegetation	9.294	12.198	12.140	7.593	8.435	9.764
Roof	44.700	38.118	38.178	29.238	28.706	20.042
Asphalt	46.157	33.465	33.448	34.337	38.346	28.563
Dirt	17.358	5.311	5.292	5.143	6.426	4.112
<i>Average</i>	29.377	22.273	22.265	19.078	20.478	15.620

listed in Table X and Table XI, respectively. Spectral curves of the extracted endmembers are compared in Fig. 13 and Fig. 14 displays their corresponding abundance maps. Particularly, the roof abundance map of SUCPSO_L1/2 is closer to the ground truth than the other compared algorithms. From these unmixing results, it can be concluded that the proposed method can obtain accurate endmembers similar to the references and reflect the actual distribution of materials. The above experiments validate that our proposed method has great potential in sparse unmixing, and can serve as a general approach for the incorporation of different sparse constraints into unmixing.

V. CONCLUSION

In this article, we present an unsupervised sparse unmixing method using comprehensive-learning-based particle swarm optimization. Two swarms are first constructed to update endmembers and abundances alternately through sharing mutual best position information. Targeting for mitigating the premature convergence of basic PSO and facilitating more precise solution search, the swarms' fitness functions are formulated into fitness vectors based on dimension division. Moreover, under this improved PSO optimization framework, two comprehensive learning strategies are devised to further refine the possible element-level position search and improve local and global

abundances' sparsity. As a versatile sparse unmixing framework, the proposed method can be easily extended to accommodate different sparse constraints of abundances. Experimental results of both simulated data and real hyperspectral images indicate that the proposed method can generate abundances with high accuracy and reasonably expressed sparsity and outperform representative sparse unmixing methods.

However, the inherent mechanism of swarm intelligence techniques often induces the PSO-based method to be time-consuming. The effective balance of constraints' impact on unmixing is also a very challenging task. Therefore, it is of great significance to take a step to overcome these drawbacks in our future work.

REFERENCES

- [1] N. Keshava and J. F. Mustard, "Spectral unmixing," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 44–57, Jan. 2002.
- [2] D. Landgrebe, "Hyperspectral image data analysis," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 17–28, Jan. 2002.
- [3] D. Liao, Y. Qian, J. Zhou, and Y. Y. Tang, "A manifold alignment approach for hyperspectral image visualization with natural color," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3151–3162, Jun. 2016.
- [4] J. M. Bioucas-Dias *et al.*, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE J. Sel. Topics Appl. Earth Obs. Remote Sens.*, vol. 5, no. 2, pp. 354–379, Apr. 2012.
- [5] M. Berman, P. Conner, L. Whitbourn, D. Coward, B. Osborne, and M. Southan, "Classification of sound and stained wheat grains using visible and near infrared hyperspectral image analysis," *J. Near Infrared Spectrosc.*, vol. 15, no. 6, pp. 351–358, 2007.
- [6] A. Gowen, C. O'Donnell, P. Cullen, G. Downey, and J. Frias, "Hyperspectral imaging—an emerging process analytical tool for food quality and safety control," *Trends Food Sci. Technol.*, vol. 18, no. 12, pp. 590–598, 2007.
- [7] C. Gendrin, Y. Roggo, and C. Collet, "Pharmaceutical applications of vibrational chemical imaging and chemometrics: A review," *J. Pharmaceutical Biomed. Anal.*, vol. 48, no. 3, pp. 533–553, 2008.
- [8] H. Li, J. Liu, and H. Yu, "An automatic sparse pruning endmember extraction algorithm with a combined minimum volume and deviation constraint," *Remote Sens.*, vol. 10, no. 4, Mar. 2018, Art. no. 509.
- [9] A. Plaza, P. Martínez, R. Pérez, and J. Plaza, "A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, pp. 650–663, Mar. 2004.
- [10] B. Zhang, X. Sun, L. Gao, and L. Yang, "Endmember extraction of hyperspectral remote sensing images based on the ant colony optimization (ACO) algorithm," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 7, pp. 2635–2646, Jul. 2011.
- [11] Q. Du, N. Raksuntorn, N. H. Younan, and R. L. King, "End-member extraction for hyperspectral image analysis," *Appl. Opt.*, vol. 47, no. 28, pp. F77–F84, Jul. 2008.
- [12] D. Hong *et al.*, "Interpretable hyperspectral artificial intelligence: When nonconvex modeling meets hyperspectral remote sensing," *IEEE Geosci. Remote Sens. Mag.*, vol. 9, no. 2, pp. 52–87, Jun. 2021.
- [13] N. Dobigeon, J.-Y. Tourneret, C. Richard, J. C. M. Bermudez, S. McLaughlin, and A. O. Hero, "Nonlinear unmixing of hyperspectral images: Models and algorithms," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 82–94, Jan. 2014.
- [14] R. Heylen, M. Parente, and P. Gader, "A review of nonlinear hyperspectral unmixing methods," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 1844–1868, Jun. 2014.
- [15] A. Halimi, Y. Altmann, N. Dobigeon, and J.-Y. Tourneret, "Nonlinear unmixing of hyperspectral images using a generalized bilinear model," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4153–4162, Nov. 2011.
- [16] Y. Altmann, A. Halimi, N. Dobigeon, and J.-Y. Tourneret, "Supervised nonlinear spectral unmixing using a postnonlinear mixing model for hyperspectral imagery," *IEEE Trans. Image Process.*, vol. 21, no. 6, pp. 3017–3025, Feb. 2012.
- [17] Z. Wang, L. Zhuang, L. Gao, A. Marinoni, B. Zhang, and M. K. Ng, "Hyperspectral nonlinear unmixing by using plug-and-play prior for abundance maps," *Remote Sens.*, vol. 12, no. 24, pp. 1–20, Dec. 2020.
- [18] L. Gao, Z. Wang, L. Zhuang, H. Yu, B. Zhang, and J. Chanussot, "Using low-rank representation of abundance maps and nonnegative tensor factorization for hyperspectral nonlinear unmixing," *IEEE Trans. Geosci. Remote Sens.*, to be published, doi: [10.1109/TGRS.2021.3065990](https://doi.org/10.1109/TGRS.2021.3065990).
- [19] C. Shi and L. Wang, "Linear spatial spectral mixture model," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3599–3611, Jun. 2016.
- [20] J. W. Boardman, F. A. Kruse, and R. O. Green, "Mapping target signatures via partial unmixing of AVIRIS data," in *Proc. Summaries 5th Annu. JPL Airborne Geosci. Workshop*, 1995, vol. 1, pp. 23–26.
- [21] M. E. Winter, "N-FINDR: An algorithm for fast autonomous spectral endmember determination in hyperspectral data," in *Proc. SPIE Conf. Imag. Spectrometry V*, 1999, vol. 3753, pp. 266–275.
- [22] J. M. P. Nascimento and J. M. B. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, Apr. 2005.
- [23] C. I. Chang, C. C. Wu, W. 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.
- [24] J. Gruninger, A. Ratkowski, and M. Hoke, "The sequential maximum angle convex cone (SMACC) endmember model," *Proc. SPIE*, vol. 5425, pp. 1–14, 2004.
- [25] T. Chan, W. Ma, A. Ambikapathi, and C. Chi, "A simplex volume maximization framework for hyperspectral endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4177–4193, Nov. 2011.
- [26] 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.
- [27] M. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2014–2039, Jun. 2011.
- [28] M. 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] M. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Collaborative sparse regression for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 341–354, Jan. 2014.
- [30] D. Wang, Z. Shi, and W. Tang, "Collaborative sparse unmixing of hyperspectral data using norm," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Beijing, China, 2016, pp. 6978–6981.
- [31] L. Sun, B. Jeon, Y. Zheng, and Y. Chen, "Hyperspectral unmixing based on sparsity and total variation," in *Proc. IEEE Int. Conf. Inf. Process.*, Phoenix, AZ, USA, 2016, pp. 4349–4353.
- [32] Z. Shi, T. Shi, M. Zhou, and X. Xu, "Collaborative sparse hyperspectral unmixing using norm," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5495–5508, Sep. 2018.
- [33] Y. Li and K. Li, "Hyperspectral unmixing based on weighted regularization," in *Proc. Int. Conf. Inf. Sci. Control Eng.*, Beijing, China, 2016, pp. 400–404.
- [34] P. V. Giampouras, K. E. Themelis, A. A. Rontogiannis, and K. D. Koutroumbas, "Simultaneously sparse and low-rank abundance matrix estimation for hyperspectral image unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4775–4789, Aug. 2016.
- [35] M. Rizkinia and M. Okuda, "Joint local abundance sparse unmixing for hyperspectral images," *Remote Sens.*, vol. 9, no. 12, Nov. 2017, Art. no. 1224.
- [36] 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.
- [37] L. Zhuang, C. Lin, M. A. T. Figueiredo, and J. M. Bioucas-Dias, "Regularization parameter selection in minimum volume hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 9858–9877, Dec. 2019.
- [38] Y. Qian, S. Jia, J. Zhou, and A. Robles-Kelly, "Hyperspectral unmixing via sparsity-constrained nonnegative matrix factorization," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4282–4297, Nov. 2011.
- [39] 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.
- [40] X. Li, X. Zhang, Y. Yuan, and Y. Dong, "Adaptive relationship preserving sparse NMF for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, to be published, doi: [10.1109/TGRS.2021.3066474](https://doi.org/10.1109/TGRS.2021.3066474).
- [41] L. Dong, Y. Yuan, and X. Lu, "Spectral-spatial joint sparse NMF for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 3, pp. 2391–2402, Mar. 2021.

- [42] J. Peng, Y. Zhou, W. Sun, Q. Du, and L. Xia, "Self-paced nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 2, pp. 1501–1515, Feb. 2021.
- [43] X. Feng, H. Li, J. Li, Q. Du, A. Plaza, and W. J. Emery, "Hyperspectral unmixing using sparsity-constrained deep nonnegative matrix factorization with total variation," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 6245–6257, Oct. 2018.
- [44] H. Fang, A. Li, H. Xu, and T. Wang, "Sparsity-constrained deep nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 15, no. 7, pp. 1105–1109, Jul. 2018.
- [45] X. Lu, L. Dong, and Y. Yuan, "Subspace clustering constrained sparse NMF for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 5, pp. 3007–3019, May 2020.
- [46] W. Wang, Y. Qian, and H. Liu, "Multiple clustering guided nonnegative matrix factorization for hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 5162–5179, Aug. 2020.
- [47] N. Guan, D. Tao, Z. Luo, and B. Yuan, "NeNMF: An optimal gradient method for nonnegative matrix factorization," *IEEE Signal Process.*, vol. 60, no. 6, pp. 2882–2898, Jun. 2012.
- [48] H. Laniéri, C. Theys, C. Richard, and C. Févotte, "Split gradient method for nonnegative matrix factorization," in *Proc. Eur. Signal Process. Conf.*, Aalborg, Denmark, 2010, pp. 1199–1203.
- [49] D. Hong, L. Gao, J. Yao, B. Zhang, A. Plaza, and J. Chanussot, "Graph convolutional networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 7, pp. 5966–5978, Jul. 2021.
- [50] L. Gao, Z. Han, D. Hong, B. Zhang, and J. Chanussot, "CyCU-Net: Cycle-consistency unmixing network by learning cascaded autoencoders," *IEEE Trans. Geosci. Remote Sens.*, to be published, doi: [10.1109/TGRS.2021.3064958](https://doi.org/10.1109/TGRS.2021.3064958).
- [51] D. Hong *et al.*, "Endmember-guided unmixing network (EGU-Net): A general deep learning framework for self-supervised hyperspectral unmixing," *IEEE Trans. Neural Netw. Learn. Syst.*, to be published, doi: [10.1109/TNNLS.2021.3082289](https://doi.org/10.1109/TNNLS.2021.3082289).
- [52] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Netw.*, 1995, vol. 4, pp. 1942–1948.
- [53] W. Chen *et al.*, "PSO-based endmembers extraction algorithm for hyperspectral imagery," *Comput. Eng. Appl.*, vol. 48, no. 8, pp. 189–193, 2012.
- [54] M. Mariana and X. Xu, "Particle swarm optimization algorithm for unmixing hyperspectral image," in *Proc. IEEE Int. Conf. Signal Process.*, 2010, pp. 897–901.
- [55] H. Dai, X. Wang, H. Hu, and Y. Wang, "Nonsmooth nonnegative matrix factorization algorithm based on particle swarm optimization," *Comput. Eng.*, vol. 39, no. 1, pp. 204–207, Jan. 2013.
- [56] B. Zhang, L. Zhuang, L. Gao, W. Luo, Q. Ran, and Q. Du, "PSO-EM: A hyperspectral unmixing algorithm based on normal compositional model," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 12, pp. 7782–7792, Dec. 2014.
- [57] L. Tong, B. Du, R. Liu, and L. Zhang, "An improved multiobjective discrete particle swarm optimization for hyperspectral endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 10, pp. 7872–7882, Oct. 2019.
- [58] B. Du, Q. Wei, and R. Liu, "An improved quantum-behaved particle swarm optimization for endmember extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 6003–6017, Aug. 2019.
- [59] M. Xu, B. Du, and Y. Fan, "Endmember extraction from highly mixed data using linear mixture model constrained particle swarm optimization," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 8, pp. 5502–5511, Aug. 2019.
- [60] T. Hendtlass, "Particle swarm optimization and high dimensional problem spaces," in *Proc. IEEE Congr. Evol. Comput.*, May 2009, pp. 1988–1994.
- [61] B. Yang, W. Luo, and B. Wang, "Constrained nonnegative matrix factorization based on particle swarm optimization for hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 8, pp. 3693–3710, Aug. 2017.
- [62] F. van den Bergh and A. P. Engelbrecht, "A cooperative approach to particle swarm optimization," *IEEE Trans. Evol. Comput.*, vol. 8, no. 3, pp. 225–239, Jun. 2004.
- [63] J. J. Liang, A. K. Qin, P. N. Suganthan, and S. Baskar, "Comprehensive learning particle swarm optimizer for global optimization of multimodal functions," *IEEE Trans. Evol. Comput.*, vol. 10, no. 3, pp. 281–295, Jun. 2006.
- [64] W. Zhao, H. Ma, and N. Li, "A new non-negative matrix factorization algorithm with sparseness constraints," in *Proc. Int. Conf. Mach. Learn. Cybern.*, Guilin, China, 2011, pp. 1449–1452.
- [65] P. Ounsrimuang and S. Nootyaskool, "Roulette wheel selection to encourage discrete particle swarm optimization solving toll-keeper scheduling problem," in *Proc. Int. Conf. Sci. Technol.*, Pathum Thani, Thailand, 2015, pp. 390–393.
- [66] R. A. Borsoi, T. Imbiriba, J. C. M. Bermudez, and C. Richard, "A fast multiscale spatial regularization for sparse hyperspectral unmixing," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 4, pp. 598–602, Apr. 2019.
- [67] D. Hong, N. Yokoya, J. Chanussot, and X. X. Zhu, "An augmented linear mixing model to address spectral variability for hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 1923–1938, Apr. 2019.
- [68] T. Uezato, M. Fauvel, and N. Dobigeon, "Hyperspectral unmixing with spectral variability using adaptive bundles and double sparsity," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 6, pp. 3980–3992, Jun. 2019.
- [69] A. Ambikapathi, T. Chan, W. Ma, and C. Chi, "Chance-constrained robust minimum-volume enclosing simplex algorithm for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4194–4209, Nov. 2011.
- [70] R. Fernandez-Beltran, A. Plaza, J. Plaza, and F. Pla, "Hyperspectral unmixing based on dual-depth sparse probabilistic latent semantic analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 11, pp. 6344–6360, Nov. 2018.



Yapeng Miao received the B.S. degree in network engineering from the School of Computer Science and Technology, Hunan University of Technology, Hunan, China, in 2019. He is currently working toward the M.S. degree in computer science and technology from the School of Computer Science and Technology, Donghua University, Shanghai, China.

His research interests include hyperspectral unmixing, pattern recognition, and swarm computing intelligence.



Bin Yang (Member, IEEE) received the B.S. degree in geographic information system from Huazhong Agricultural University, Wuhan, China, in 2011, the M.S. degree in cartography and geographic information system from South China Normal University, Guangzhou, China, in 2015, and the Ph.D. degree in circuit and system from Fudan University, Shanghai, China, in 2019.

Since July 2019, he has been with the School of Computer Science and Technology, Donghua University, Shanghai, China, where he is currently a Lecturer. His main research interests include hyperspectral remote sensing image analysis, machine learning, pattern recognition, and computational intelligence.