Endmember Selection of Hyperspectral Images based on Evolutionary Multitask

Yizhe Zhao*, Hao Li*, Yue Wu†, Shanfeng Wang‡ and Maoguo Gong*
*School of Electronic Engineering
Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education
Xidian University, Xi'an, Shaanxi Province 710071, China
Email: zhaoyizhe@stu.xidian.edu.cn; haoli@xidian.edu.cn; gong@ieee.com
†School of Computer Science and Technology, Xidian University, Xi'an, Shaanxi Province 710071, China
Email: ywu@xidian.edu.cn
‡School of Cyber Engineering, Xidian University, Xi'an, Shaanxi Province 710071, China
Email: sfwang@xidian.edu.cn

***Abstract*—Endmember selection of hyperspectral images is a practical yet difficult task due to the high spectral resolution and low spatial resolution of the hyperspectral cameras. The paradigm of multitask optimization has been investigated over two decades, which aim to handle multiple tasks simultaneously. To address these issues, we propose a novel multitasking framework based on multiobjective optimization evolutionary algorithm based on decomposition (MOEA/D). Specifically, we use a single population to simultaneously perform multiple subset selection tasks and apply it to a specific scene-the endmember selection of hyperspectral images. It is natural to consider that pixels in a homogeneous region of hyperspectral image as a task. Then, a within-task and between-task genetic transfer operator is constructed to reinforce the exchange of genetic material belonging to the same or different tasks for better and quicker search of the decision space. After that, this algorithm obtains a set of nondominated solutions for better decision of the active endmembers. Experiments on hyperspectral datasets show the effectiveness of our method in finding the real active endmembers.**

***Index Terms*—hyperspectral image, endmember selection, multitask optimization, evolutionary algorithm**

I. Introduction and Related Work

With the development of remote sensing technology, hyperspectral cameras (HSCs) with higher spectral resolution than multispectral cameras, can measure electromagnetic energy scattered in their instantaneous field view in hundreds or thousands of spectral channels [1]. Compared with RGB images, hyperspectral images contain more information, which greatly improve the ability of material identification. By using this feature, researchers have applied it to many fields, such as agricultural production, geological analysis, environmental monitoring, urban planning and military applications [2], [3]. However, in hyperspectral imagery, the pixels may contain more than one types of ground object signatures due to the low spatial resolution of the HSCs, which called the mixed pixels [4]. The ground object signatures mentioned above are also known as endmembers, which normally correspond to familiar macroscopic objects in the real scenario, such as water, soil, metal, vegetation, etc. [1], [4]. Because of the existence of mixed pixels, the accuracy of hyperspectral image processing is severely limited. Under the circumstances, how to identify the real active endmembers in mixed pixels is a key preprocessing technique for hyperspectral image analysis.

Extracting the pure spectral signatures (i.e. endmembers) in the scene is the first step of hyperspectral unmixing, and the second step is to estimate the corresponding proportions (called abundances) of these endmembers [4]. Note that we only focus on the endmember selection process since it is a prerequisite for the subsequent steps. There are two main models of hyperspectral unmixing: linear mixture model (LMM) [4], [5] and nonlinear mixture model (NLMM) [6]. Different from the NLMM, the LMM has more extensive applications to date because of it advantages of simplicity, flexibility, high efficiency and clear physical meaning.

In the LMM model, each mixed pixel of hyperspectral image is considered as a linear combination of endmembers weighted by its corresponding abundance fractions [4]. And under this theory, a group of unmixing approaches based on geometry [7], [8], statistics [9], and nonnegative matrix factorization (NMF) [10]–[13] has been proposed. These algorithms only require a small amount of prior knowledge of the hyperspectral image, nevertheless, some of these methods could extract virtual endmembers with no physical meaning [11] and others [7], [8] assume the existence of pure pixel, which is hard to guarantee in hyperspectral images. In addition, these algorithms perform poorly when hyperspectral image data in highly mixed scenario or have much noise [14].

In reality, the complexity and diversity of the ground substances increase the difficulty of endmember selection, that is to say, the endmembers extracted from the observed hyperspectral data are not accurate enough. In order to overcome inaccurate endmember extraction during the unmixing processing, some methods based on sparse representation [14], [15], have attracted more and more researchers' attention. The spectral library is used in sparse unmixing technology to constitute endmember matrix, which is a collection of reflectance spectral data of various kinds of ground objects measured by high spectral spectrometer under certain conditions, just like a "dictionary". To be precise, mixed pixels in the hyperspectral

978-1-7281-6929-3/20/$31.00 ©2020 IEEE

image have been modeled with a few spectra in the library, as illustrated in Fig. 1. In the actual spectral mixing phenomenon of hyperspectral images, the number of active endmembers existing in a pixel or image is far less than the number of the spectral signatures in the library, so this method often leads to sparse solution [14].

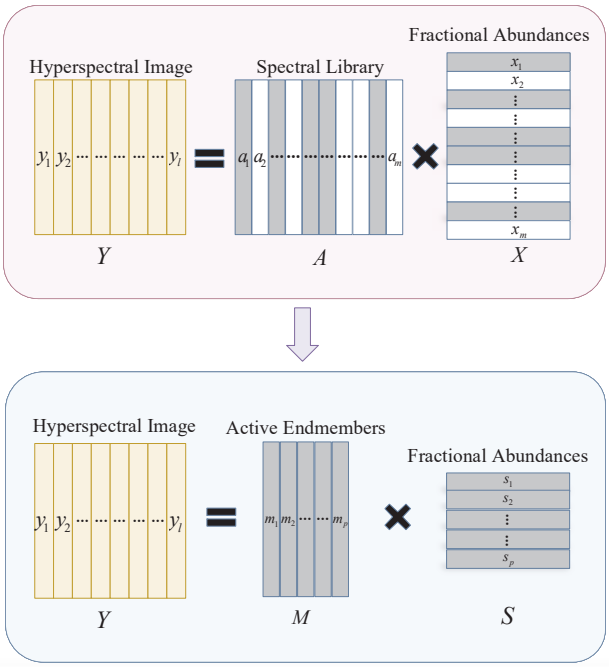


Fig. 1. Illustration on sparse regression-based endmember selection of hyperspectral data. $\boldsymbol{M}$ is formed by the active endmembers (in gray) which contributes to $\boldsymbol{Y}$, and $\boldsymbol{S}$ is formed by the corresponding rows of $\boldsymbol{M}$ in $\boldsymbol{X}$.

According to the above description, selecting active endmembers from spectral library is essentially a subset selection problem. Subset selection is to select a subset of size k from a total set of n variables for optimizing some criterion [16]. Subset selection can be considered as a bi-objective optimization problem that optimizes the given criteria and the size of the subset simultaneously [16], and these two objectives are generally contradictory to each other. In other words, the subset selection problem is essentially a multiobjective optimization problem (MOP). Therefore, there is no single solution in the feasible region that can simultaneously minimize all the objective functions. Instead, there is a set of Pareto optimal solutions that are tradeoffs between different objectives [17]. The Pareto set (PS) is a set of all Pareto optimal solutions and mapping it to objective space is the Pareto front (PF). A Pareto optimization for subset selection (POSS) algorithm has been proposed in [16]. The POSS solves the subset selection problem by assigning external variables to subsets gradually until the stop criterion is reached, but the convergence speed is slow, because of the single flipping strategy.

Multiobjective optimization evolutionary algorithms (MOEAs) aim to find a set of approximate solutions to the PF [17]–[20], therefore, it is natural that we use MOEA to solve the problem of hyperspectral endmember selection. In addition, MOEAs also have some common advantages, such as the low modeling complexity and the needless relaxation of the nonconvex regularizers [21]–[24]. In [24], a multiobjective optimization algorithm based on the NSGA-II has been proposed, but the nondominated sorting strategy used in it is time-consuming. In addition, the multiobjective optimization evolutionary algorithm based on decomposition (MOEA/D) proposed in [19] decomposes MOP into a sequence of scalar subproblems and optimizes them simultaneously, thus, it can generate evenly distributed objective vectors along the PF.

As mentioned earlier, evolutionary algorithms can solve optimization problem, notwithstanding, its are often designed to effectively solve one optimization problem at a time. In modern times, many scientists are devoted to the research of intelligent algorithms and systems that can efficiently perform multitasking, utilizing the basic commonality between different optimization tasks to simultaneously handle multiple optimization tasks [25]. These multitask algorithms are usually implemented through parameter sharing to improve the processing efficiency in the actual problem solving. Gupta et al. proposed multifactorial evolutionary algorithm (MFEA) in [26] in order to process multiple task simultaneouly and then extended to a multiobjective version in [27].

In this paper, we combine the MOEA/D with the idea of multitask, and propose a novel multitask endmember selection (MTES) framework. Multitask optimization refers to process multiple tasks simultaneously and utilizes the basic commonality between different optimization tasks so as to improve the processing efficiency in practical problem. The proposed framework uses a single population to simultaneously perform multiple subset selection tasks and applies it to a specific scene-the endmember selection of hyperspectral image. Therein, the within-task and between-task genetic transfer operator is constructed to reinforce the exchange of genetic material belonging to the same or different tasks for better and quicker search of the decision space. Experimental results on hyperspectral data reveal that MTES can process multiple regions of hyperspectral images simultaneously and has a faster convergence rate.

II. Evolutionary Multitasking Endmember Selection

Multitasking endmember selection aims to extracting the active endmembers of hyperspectral corresponding to the given task. We define $\boldsymbol{Y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_l]$ as a hyperspectral image that contains l pixels, where $\boldsymbol{y}_i$ denote the ith pixel in hyperspectral image with D spectral bands. $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_m]$ is a spectral library with m spectral signatures and each of them also has D spectral bands. Suppose $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_l]$ is the abundance fraction matrix. To ensure that the solution of the abundance has practical significance, the LMM introduces two constraints, named the abundance nonnegative constraint (ANC) [28] and

the abundance sum-to-one constraint (ASC) [28]. If noise is considered, $\boldsymbol{y}$ and $\boldsymbol{Y}$ can be expressed as

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{n} \quad s.t.\ \boldsymbol{x} \geq 0, \tag{1}$$

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{N} \quad s.t.\ \boldsymbol{X} \geq 0,\ 1_m^T \boldsymbol{X} = 1_l^T, \tag{2}$$

where $\boldsymbol{n} \in \mathbb{R}^D$ is the additive noise. $\boldsymbol{X} \geq 0$ is the ANC in the component-wise sense; $1_m^T \boldsymbol{X} = 1_l^T$ is the ASC for physical interpretability.

It is well known that the active endmembers might be different between pixels. In practice, the collaborative sparse unmixing model employed in [29], [30] assumes that all pixels in a hyperspectral image can share the same active set of endmembers. Compared to the pixels in the entire image, the pixels in the same homogeneous region are more likely to share the same active set of endmembers. Therefore, we partition the hyperspectral image into several homogeneous regions and find small set of endmembers (i.e., $\boldsymbol{M}_1$, $\boldsymbol{M}_2$, ...) to unmix all the pixels in the same region. Various clustering or classification methods for hyperspectral image have been proposed [31], so we can use them in this step. Assuming that we have K homogeneous regions, $\boldsymbol{Y}$ can be represented as

$$\boldsymbol{Y} = \bigcup_{k=1}^{K} \boldsymbol{Y}_k, \tag{3}$$

where $\boldsymbol{Y}_k$ is the pixels in the kth homogeneous region. As similar with (2), for each region, we have

$$\boldsymbol{Y}_k = \boldsymbol{A}\boldsymbol{X}_k + \boldsymbol{N}_k. \tag{4}$$

Because the estimated endmembers $\boldsymbol{M}$ are selected from the spectral library $\boldsymbol{A}$, a binary variable $\boldsymbol{I}$ formed by 0 and 1 is used to indicate the locations of the active endmembers in the spectral library. If the jth element of $\boldsymbol{I}$ is equal to 1, the signature in the jth column of the spectral library is selected as active, otherwise it is inactive. Thus, the active endmembers $\boldsymbol{M}_k$ is the nonzero rows of corresponding $\boldsymbol{A}$ in the kth homogeneous region and abundance matrix $\boldsymbol{S}_k$ is the subvector of $\boldsymbol{X}_k$ containing the rows specified in the $\boldsymbol{I}_k$.

In each task, the endmember selection problem is essentially an MOP, which can be described as

$$\begin{aligned} \min \quad & \boldsymbol{F}(\boldsymbol{x}) = (f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \cdots, f_n(\boldsymbol{x}))^\top \\ \text{s.t.} \quad & \boldsymbol{x} = (x_1, x_2, \cdots, x_m)^\top \in \boldsymbol{\Omega}, \end{aligned} \tag{5}$$

where $\boldsymbol{F}_i(\boldsymbol{x})$ is the objective functions with n real-valued of and $\boldsymbol{\Omega}$ is the feasible set. Assuming $\boldsymbol{x_1}, \boldsymbol{x_2} \in \boldsymbol{\Omega}$, $\boldsymbol{x_1}$ is said to be dominated by $\boldsymbol{x_2}$ (denoted as $\boldsymbol{x_1} \prec \boldsymbol{x_2}$) if and only if

$$\begin{aligned} &\forall i = 1, 2, \cdots, n \quad f_i(\boldsymbol{x_1}) \geq f_i(\boldsymbol{x_2}) \\ &\exists i = 1, 2, \cdots, n \quad f_i(\boldsymbol{x_1}) > f_i(\boldsymbol{x_2}). \end{aligned} \tag{6}$$

The solution $\boldsymbol{x}'$ is a Pareto optimal if and only if no other solution in $\boldsymbol{\Omega}$ dominates it. All the Pareto optimal solutions in $\boldsymbol{\Omega}$ constitute the Pareto set (PS) and the objective vectors of all the solutions in PS compose the Pareto front (PF) as

$$\mathrm{PF} = \{\boldsymbol{F}(\boldsymbol{x}) | \boldsymbol{x} \in \mathrm{PS}\} \tag{7}$$

In MOEA/D, tchebycheff approach [19] is use to decompose the MOP into a sequence of scalar optimization subproblems as

$$\begin{aligned} \min_x \ g^{te}(\boldsymbol{x}|\boldsymbol{\lambda}, \boldsymbol{z}^*) = \quad & \max_{1 \leq i \leq n} \{\lambda_i |f_i(\boldsymbol{x}) - z_i^*|\} \\ & \text{s.t. } \boldsymbol{x} \in \boldsymbol{\Omega}, \end{aligned} \tag{8}$$

where $\boldsymbol{z}^* = (z_1^*, \ldots, z_n^*)$ is the ideal or reference points, e.g., $z_i^* = 0$ or $z_i^* = \min\{f_i(x) | x \in \boldsymbol{\Omega}\}$. For each specific weight vector λ, there exists an optimal solution of (8) that solution is Pareto optimal with regard to (5). Therefore, different Pareto optimal solutions can be obtained by varying the weight vector, which plays an important role in maintaining the diversity of population.

Empirically, the real active endmembers are the least number of endmembers that can best represent the hyperspectral data. Therefore, the active endmembers can be estimated by

$$\min\ \boldsymbol{F}(\boldsymbol{x}) = \min\ (|\boldsymbol{I}|, \|\boldsymbol{Y} - \boldsymbol{M}\boldsymbol{S}\|_2). \tag{9}$$

To be precise, when the objective $\boldsymbol{I}$ is small, the endmembers contained in the matrix $\boldsymbol{M}$ are generally not enough to represent the data $\boldsymbol{Y}$, so the unmixing residuals will be large. When the number of endmembers increases sufficiently to represent the data $\boldsymbol{Y}$, the unmixing residuals will be small. Therefore, the two objectives in model (9) are suitable for estimating the active endmembers and the knee point in the PF should be selected to locate the active endmembers [32].

A multitasking optimization problem with K tasks can be mathematically expressed as

$$\begin{aligned} \{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_K\} = \quad & \arg\min \{\boldsymbol{F}_1(\boldsymbol{x}), \boldsymbol{F}_2(\boldsymbol{x}), \cdots, \boldsymbol{F}_K(\boldsymbol{x})\} \\ & \text{s.t.} \quad \boldsymbol{x}_i \in \boldsymbol{\Omega}_i, i = 1, 2, \cdots, K. \end{aligned} \tag{10}$$

In [33], Gupta et al. proposed a multifactorial optimization technique to solve the multitasking optimization problems, then extended it to deal with multiobjective optimization problems. In evolutionary multitasking optimization, every individual p_i in a population P is associated with a task. For the p_i in the kth task T_k, the factorial rank r_k^i, skill factor τ_i and scalar fitness ϕ_i are defined as follows:

- The factorial rank r_k^i is the index of p_i in the list of population members sorted in ascending order with respect to T_k.
- The skill factor τ_i indicates the associated task. $\tau_i = K$ means individual p_i belongs to the Kth task.
- The scalar fitness of p_i in a multitasking environment is given by $\phi_i = 1/r_{\tau_i}^i$.

Each individual is associated with the most effective task, on which the individual has the best factorial rank. If $\phi_1 > \phi_2$, the individual p_1 is considered to dominate p_2.

As mentioned in the preceding discussion, the pixels in hyperspectral image share similar sparsity pattern. We divide hyperspectral image into several homogeneous regions and make each region into a sub-task by selecting the endmembers. In order to take full advantage of the parallelism of population-based search, we are motivated to design evolutionary multitasking framework for endmember selection, as depicted in Fig. 2.

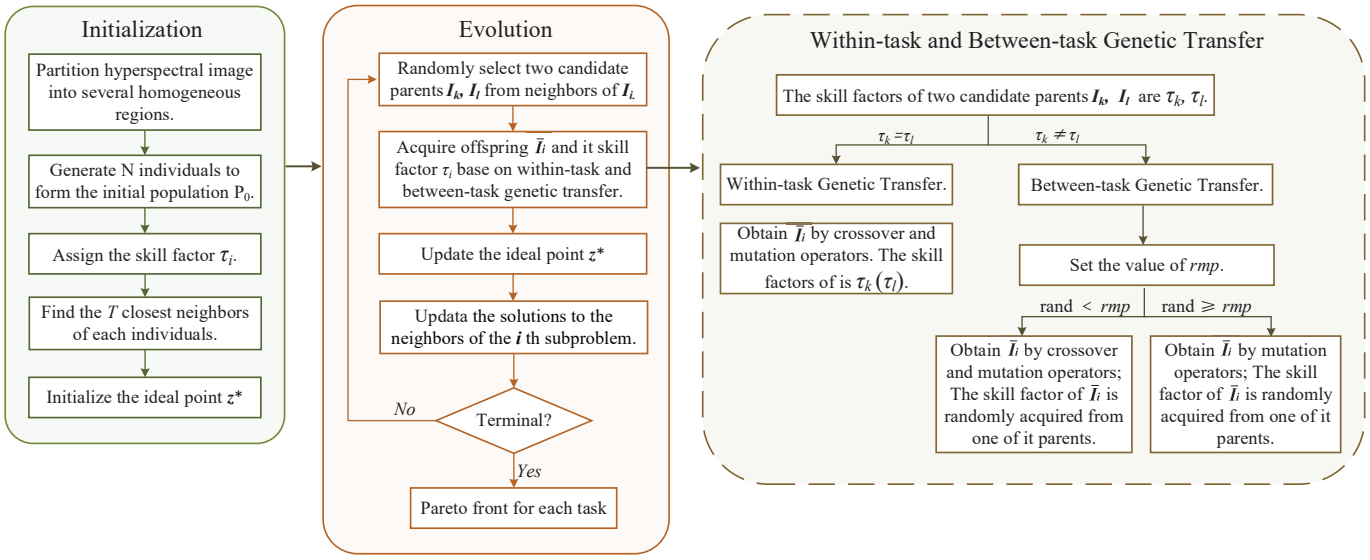


Fig. 2. Illustration on the proposed MTES framework. The procedures in green and orange regions represent initialization and population evolution, respectively. The procedure in gray region represents within-task and between-task genetic transfer.

In the initialization step (the green module in Fig. 2.), population is initialized by randomly generating N binary vectors (i.e., $\boldsymbol{I}_1$, $\boldsymbol{I}_2$, ..., $\boldsymbol{I}_N$) to form the initial population P_0. The length of each vector is equal to the column number of the spectral library. Individuals are encoded into a unified space, and we assign skill factor of each individual to indicate the associated task. After that, we generate N uniformly distributed vectors $\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \ldots, \boldsymbol{\omega}_N$ that uniformly distributed between (0, 1) and (1, 0). Size of neighbors T is set to 10 and the Euclidean distance is used to find the neighbor set $B(i)$ (i.e., $B(i) = \{i_1, i_2, \ldots, i_T\}$) for each weight vector $\boldsymbol{\omega}_i$. Lastly, the ideal point $\boldsymbol{z}^*$ is initialized with the objectives of the individual $\boldsymbol{I}_1$ in first task.

In the evolution phase (the orange module in Fig. 2.), we firstly acquire new individual $\bar{\boldsymbol{I}}_i$ and its skill factor τ_i. Then, the ideal point $\boldsymbol{z}^*$ updated by $\boldsymbol{z}^* = min\{\boldsymbol{z}^*, \boldsymbol{F}_{\tau_i}(\bar{I}_i)\}$ and the solutions to the neighbors of the ith subproblem should be updated by (6). To be precise, for each index $j \in B(i)$, let $\boldsymbol{I}_j = \bar{\boldsymbol{I}}_i$ if $g^{te}_{\tau_i}(\bar{\boldsymbol{I}}_i|\boldsymbol{\omega}^j) < g^{te}_{\tau_i}(\boldsymbol{I}_j|\boldsymbol{\omega}^j)$. The evolution process of all the subproblems is not terminated until the number of iterations (300 in this paper) meets the upper limit. Finally, we arrive at a set of nondominated solutions for each task. During the evolution of the population, the genetic materials can be transferred among the candidate parents belonging to the same or different tasks, as depicted in the gray module of Fig. 2. When the skill factors of the two parents are different, the random mating probability (rmp) is defined to determine whether the two parents belonging different tasks conduct genetic transfer. The skill factor of the offspring is obtained from one of its parents. When the value of rmp is close to 0, only culturally alike individuals are allowed to crossover. In contrast, when the value of rmp is close to 1, the between-task genetic transfer is able to enhance exploration of the entire search space. The transfer turns out to be beneficial if the genetic material from one task happens to be useful for the another task. We let $rmp = 0.8$ in this paper and use multipoint crossover operator and single-point mutation operator in [15].

III. Experimental Study

A. Data Sets and Evaluation Indicators

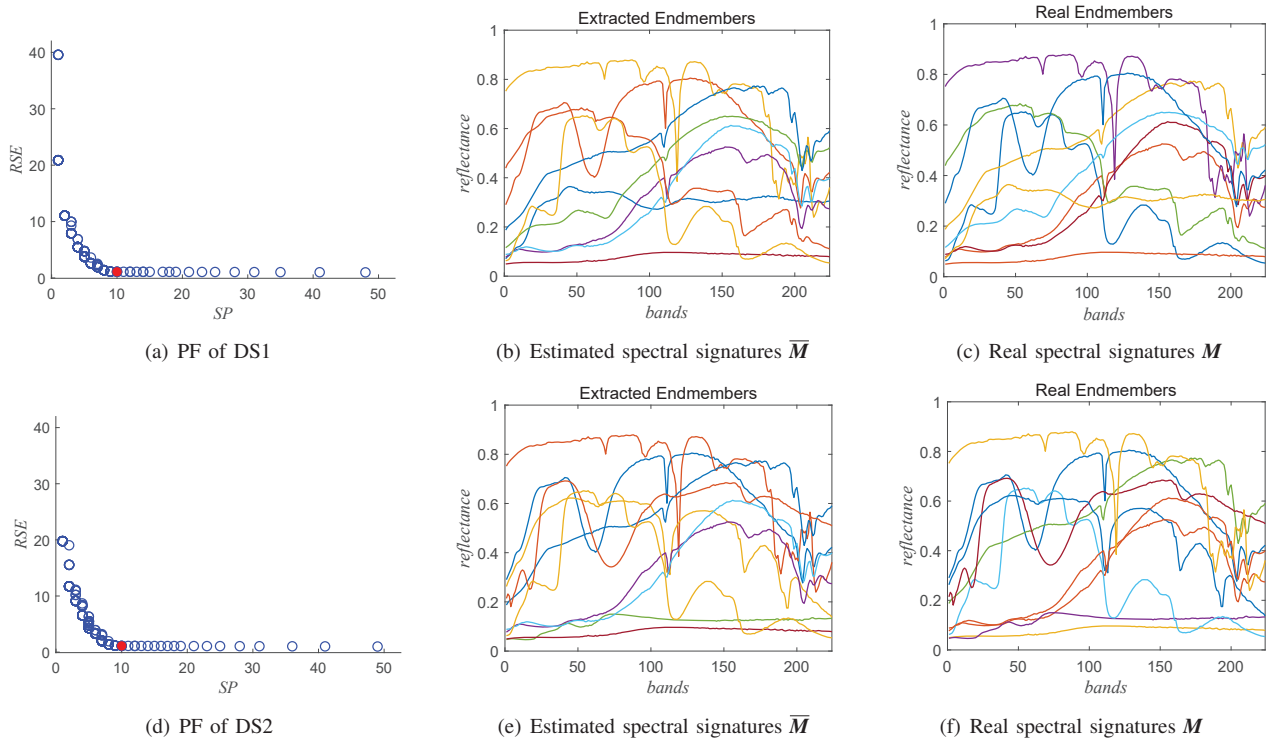
In this paper, we use famous USGS digital spectral library (denoted splib06[1]) to generate synthetic data based on the LMM. USGS library contains spectral signatures which reflectance values are measured in 224 spectral bands and uniformly distributed in the interval of 0.4-2.5 μm. In order to analyze the influence of the mutual coherence among the spectral signatures on the performance of the proposed method, we pruned the above spectral library through the spectral angle distance (SAD) to form two spectral libraries $\boldsymbol{A}1$ and $\boldsymbol{A}2$. To be precise, the SAD of any two signatures in $\boldsymbol{A}1$ is not smaller than 4.4° and the SAD of any two signatures in $\boldsymbol{A}2$ is not smaller than 10°. The synthetic data sets constructed by library $\boldsymbol{A}1$ and library $\boldsymbol{A}2$ based on the LMM were named as DS1 and DS2 respectively, and each of them contains 300 pixels.

In order to compare the performance of our endmember selection algorithm, three evaluation indicators are considered in the experiments:

a) Average number (AN): This index can be used to measure the accurately estimated endmembers, which is defined as

$$\mathrm{AN} = \frac{\sum_{n=1}^{\mathrm{NT}} N_n}{\mathrm{NT}}, \tag{11}$$

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



(a) PF of DS1 (b) Estimated spectral signatures $\overline{\boldsymbol{M}}$ (c) Real spectral signatures $\boldsymbol{M}$

(d) PF of DS2 (e) Estimated spectral signatures $\overline{\boldsymbol{M}}$ (f) Real spectral signatures $\boldsymbol{M}$

Fig. 3. Illustration on the use of MTES. The top row is the results on the DS1 and the bottom row is the results on the DS2 (SNR = 40).

where NT is the total number of the estimation trials and N_n is the number of accurately estimated endmembers in the nth trial.

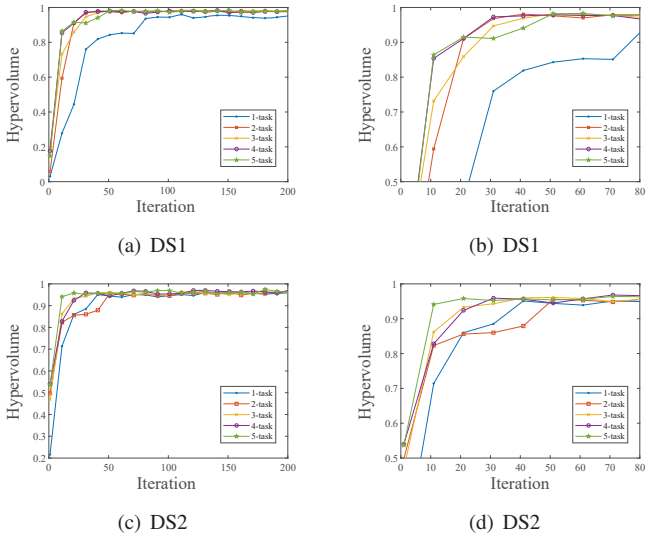
b) Correct estimation rate (CER): The CER can be used to measure the number of correct estimations of the whole trials. NC denotes the number of correct estimations of the whole endmembers; NT is the total number of the estimation trials. The CER can be calculated by

$$\text{CER} = \text{NC}/\text{NT}. \tag{12}$$

c) Hypervolume: The hypervolume indicator is used to measure the quality of the Pareto front, which can evaluate convergence and distribution of the solution set simultaneously. We prefer hypervolume indicator because the true Pareto front is not known beforehand for real-world problems. Hypervolume is calculated using a reference point 1% larger in every component than the corresponding nadir point [20].

B. Results and Discussions

To demonstrate the ability of the proposed MTES for finding the real endmembers in multitasking environment, we conducted two experiments on $\boldsymbol{A}1$ and $\boldsymbol{A}2$ respectively. Each experiment consisted of two tasks and the endmember numbers of each task was set to 10, moreover, the Gaussian white noise with SNR = 40 dB was imposed on each hyperspectral data. Note that Fig. 3 only shows the results of the first task of the two experiments, and the top row is the results on DS1 , the bottom row is the results on DS2. The PF of above two experiments show in Fig. 3(a) and (d) and the knee points (in red) are located to find the corresponding estimated endmembers. Extracted endmembers of DS1 and DS2 were shown in Fig. 3(b) and (e), respectively. It can be seen that the estimated endmembers presented in Fig. 3(b) and (e) are identical to the real active endmembers presented in Fig. 3(c) and (f). The only difference lies in that the column indices of the same endmember in estimated endmember set $\overline{\boldsymbol{M}}$ and real active endmember set $\boldsymbol{M}$ are inconsistent.



(a) DS1 (b) DS1

(c) DS2 (d) DS2

Fig. 4. Comparing the evolution of the hypervolume indicator for DS1 and DS2 (SNR = 40). (b) and (d) are the first 80 iterations of (a) and (c).

TABLE I
PERFORMANCE OF THE MUSIC-CSR, TP-MOSU AND MTES UNDER THE VARYING NUMBER OF ENDMEMBERS ON DS1 AND DS2 (SNR = 40)

p	MUSIC-CSR				Tp-MoSU				MTES			
	AN		CER		AN		CER		AN		CER	
	DS1	DS2	DS1	DS2	DS1	DS2	DS1	DS2	DS1	DS2	DS1	DS2
4	3.6	**4**	0.9	**1**	**4**	**4**	**1**	**1**	**4**	**4**	**1**	**1**
6	5.5	5.8	0.7	0.8	**6**	**6**	**1**	**1**	**6**	**6**	**6**	**1**
8	5.2	7.2	0.3	0.6	**8**	**8**	**1**	**1**	**8**	**8**	**1**	**1**
10	6.3	8.7	0.1	0.3	9.8	**10**	0.8	**1**	**10**	**10**	**1**	**1**
20	8.5	14.6	0	0	15.3	19.8	0	0.9	**20**	**20**	**1**	**1**
30	11.5	19.7	0	0	18.6	27.3	0	0	**27.3**	**28.4**	**0.3**	**0.6**

TABLE II
EXPERIMENTAL RESULTS OF AN BY THE MTES UNDER THE VARYING NUMBER OF TASKS ON DS1 (SNR = 20, 30, 40)

TASK	p	AN		
		SNR=20	SNR=30	SNR=40
1	10	9.35	9.6	9.8
2	20	16.45	18.75	19.9
3	30	24.2	28.5	29.8
4	40	30.8	38.25	40
5	50	34.7	47.65	49.9

TABLE III
EXPERIMENTAL RESULTS OF CER BY THE MTES UNDER THE VARYING NUMBER OF TASKS ON DS1 (SNR = 20, 30, 40)

TASK	p	CER		
		SNR=20	SNR=30	SNR=40
1	10	0.6	0.7	0.8
2	20	0.1	0.43	0.95
3	30	0.07	0.55	0.93
4	40	0.03	0.61	1
5	50	0	0.59	0.98

Next, we did two experiments with different number of tasks to analyze the influence of the number of tasks on our algorithm. When TASK is equal to 1 (denoted as TASK1), it consists of one task, TASK2 consist of two tasks (each task had 10 endmembers), and so on. Note that there is no between-task genetic transfer in TASK1, because it only has one task. In addition, we also considered the robustness to the noise interference. We conducted 40 repeated trails on each experiment with SNR = 20,30,40, respectively, then obtained the mean values of AN and CER. The quantitative results of AN and CER are presented in Table II and Table III on the DS1 respectively. EA-based endmember selection approaches face the curse of dimensionality [33]. It is worth noting that with the increase of the number of total endmembers p, AN and CER decreased somewhat but not dramatically in Table II and Table III, indicating that our multitasking framework could effectively avoid the precision attenuation caused by too many endmembers.

The evolution of the hypervolume indicator for DS1 and DS2 is shown in Fig. 4. We conducted 5 experiments on DS1 and DS2, respectively, corresponding to five different numbers of tasks. In Fig. 4(a), five broken lines represent five experimental results of the hypervolume indicator on DS1, which line "1-task" represents the experiment with one task and line "5-task" represents the experiment with five tasks. It can be concluded from the convergence trends of the hypervolume in Fig. 4(b), (d), when the number of iterations is small, the experiments with between-task genetic transfer obtains higher hypervolume values than experiment not have between-task genetic transfer (e.g. line "1-task"). That is to say, when multiple relevant tasks are optimized simultaneously by the MTES with the between-task genetic transfer, the overall convergence characteristics can be significantly improved.

In the end, three methods are compared in this paper: 1) MUSIC-CSR [30]; 2) Tp-MoSU [15]; 3) MTES proposed in this work. The first two algorithms are single-task hyperspectral endmember selection methods, while ours is a multi-task algorithm (we used the bi-task in this experiment.). As is shown in Table I, the proposed MTES significantly outperforms other methods on both two datasets with the increase of active endmembers p.

IV. CONCLUSION

In this paper, we proposed a MTES framework for better handling of the hyperspectral endmember selection. The MTES considers that the pixels in a homogeneous region are considered as a task and then solves these tasks simultaneously. Experimental studies on hyperspectral data have demonstrated that the proposed algorithm was robust to the number of active endmembers because of the between-task genetic transfer. Quantitative results on MUSIC-CSR, Tp-MoSU demonstrate the effectiveness of the proposed framework. In the future, we hope to improve the proposed framework for more applications.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China under Grant 61772393 and Grant 61906146, and in part by the Key Research and Development Program of Shaanxi Province under Grant 2018ZDXM-GY-045.

REFERENCES

[1] J. M. Bioucas-Dias, A. Plaza, N. Dobigeon, M. Parente, Q. Du, P. Gader, and J. Chanussot, "Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 5, no. 2, pp. 354–379, 2012.

[2] D. Landgrebe, "Hyperspectral image data analysis," *IEEE Signal Processing Magazine*, vol. 19, no. 1, pp. 17–28, 2002.

[3] M. W. Mwaniki, M. S. Matthias, and G. Schellmann, "Application of remote sensing technologies to map the structural geology of central region of kenya," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 4, pp. 1855–1867, 2015.

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

[5] C. Shi and L. Wang, "Linear spatial spectral mixture model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 6, pp. 3599–3611, 2016.

[6] A. Marinoni, A. Plaza, and P. Gamba, "Harmonic mixture modeling for efficient nonlinear hyperspectral unmixing," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 9, pp. 4247–4256, 2016.

[7] J. W. Boardman, "Automating spectral unmixing of aviris data using convex geometry concepts," *in Proc. Annual JPL Airborne Geoscience Workshop*, vol. 1, pp. 11–14, 1993.

[8] J. M. Nascimento and J. M. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 4, pp. 898–910, 2005.

[9] 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 Transactions on Geoscience and Remote Sensing*, vol. 42, no. 10, pp. 2085–2095, 2004.

[10] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[11] L. Miao and H. Qi, "Endmember extraction from highly mixed data using minimum volume constrained nonnegative matrix factorization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 3, pp. 765–777, 2007.

[12] G. Zhou, S. Xie, Z. Yang, J.-M. Yang, and Z. He, "Minimum-volume-constrained nonnegative matrix factorization: Enhanced ability of learning parts," *IEEE Transactions on Neural Networks*, vol. 22, no. 10, pp. 1626–1637, 2011.

[13] J. Li, J. M. Bioucas-Dias, A. Plaza, and L. Liu, "Robust collaborative nonnegative matrix factorization for hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 10, pp. 6076–6090, 2016.

[14] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 6, pp. 2014–2039, 2011.

[15] X. Jiang, M. Gong, H. Li, M. Zhang, and J. Li, "A two-phase multiobjective sparse unmixing approach for hyperspectral data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 1, pp. 508–523, 2017.

[16] C. Qian, Y. Yu, and Z. H. Zhou, "Subset selection by pareto optimization," in *Advances in Neural Information Processing Systems*, 2015.

[17] K. Deb, *Multi-objective optimization using evolutionary algorithms*. John Wiley & Sons, 2001, vol. 16.

[18] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002.

[19] Q. Zhang and H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 6, pp. 712–731, 2007.

[20] N. Beume, B. Naujoks, and M. Emmerich, "SMS-EMOA: Multiobjective selection based on dominated hypervolume," *European Journal of Operational Research*, vol. 181, no. 3, pp. 1653–1669, 2007.

[21] E. Candes and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Problems*, vol. 23, no. 3, pp. 969–985, 2007.

[22] L. Li, X. Yao, R. Stolkin, M. Gong, and S. He, "An evolutionary multiobjective approach to sparse reconstruction," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 6, pp. 827–845, 2013.

[23] M. Gong, H. Li, E. Luo, J. Liu, and J. Liu, "A multiobjective cooperative coevolutionary algorithm for hyperspectral sparse unmixing," *IEEE Transactions on Evolutionary Computation*, vol. 21, no. 2, pp. 234–248, 2016.

[24] X. Xu and Z. Shi, "Multi-objective based spectral unmixing for hyperspectral images," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 124, pp. 54–69, 2017.

[25] C. M. Dzubak *et al.*, "Multitasking: The good, the bad, and the unknown," *The Journal of the Association for the Tutoring Profession*, vol. 1, no. 2, pp. 1–12, 2008.

[26] A. Gupta, Y. S. Ong, and L. Feng, "Multifactorial evolution: Toward evolutionary multitasking." *IEEE Transactions on Evolutionary Computation*, vol. 20, no. 3, pp. 343–357, 2016.

[27] A. Gupta, Y. S. Ong, L. Feng, and K. C. Tan, "Multiobjective multifactorial optimization in evolutionary multitasking," *IEEE Transactions on Cybernetics*, vol. 47, no. 7, pp. 1652–1665, 2017.

[28] C.-I. Chang and D. C. Heinz, "Constrained subpixel target detection for remotely sensed imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 3, pp. 1144–1159, 2000.

[29] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Collaborative sparse regression for hyperspectral unmixing," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 1, pp. 341–354, 2013.

[30] M.-D. Iordache, J. M. Bioucas-Dias, A. Plaza, and B. Somers, "Music-csr: Hyperspectral unmixing via multiple signal classification and collaborative sparse regression," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, no. 7, pp. 4364–4382, 2013.

[31] A. Paoli, F. Melgani, and E. Pasolli, "Clustering of hyperspectral images based on multiobjective particle swarm optimization," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 12, pp. 4175–4188, 2009.

[32] J. Branke, K. Deb, H. Dierolf, and M. Osswald, "Finding knees in multi-objective optimization," in *International Conference on Parallel Problem Solving from Nature*. Springer, 2004, pp. 722–731.

[33] Y.-S. Ong and A. Gupta, "Evolutionary multitasking: a computer science view of cognitive multitasking," *Cognitive Computation*, vol. 8, no. 2, pp. 125–142, 2016.