This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TEVC.2016.2598858, IEEE Transactions on Evolutionary Computation

A Multi-objective Cooperative Coevolutionary Algorithm for Hyperspectral Sparse Unmixing

Maoguo Gong, *Senior Member, IEEE,* Hao Li, Enhu Luo, Jing Liu, *Senior Member, IEEE,* and Jia Liu

***Abstract*—Sparse unmixing of hyperspectral data is an important technique aiming at estimating the fractional abundances of the endmembers. Traditional sparse unmixing is faced with the l_0-norm problem which is an NP-hard problem. Sparse unmixing is inherently a multi-objective optimization problem. Most of the recent works combine cost functions into single one to construct an aggregate objective function, which involves weighted parameters that are sensitive to different data sets and difficult to tune. In this paper, a novel multi-objective cooperative coevolutionary algorithm is proposed to optimize the reconstruction term, the sparsity term and the total variation regularization term simultaneously. A problem-dependent cooperative coevolutionary strategy is designed because sparse unmixing encounters a large scale optimization problem. The proposed approach optimizes the non-convex l_0-norm problem directly and can find a better compromise between two or more competing cost function terms automatically. Experimental results on simulated and real hyperspectral data sets demonstrate the effectiveness of the proposed method.**

***Index Terms*—Hyperspectral unmixing, multi-objective optimization, sparse unmixing, cooperative coevolution, evolutionary algorithm.**

I. Introduction

With the development of the earth observation technology, hyperspectral remote sensing imagery has been widely used in monitoring of environmental, mineral exploration urban processes, etc. A common problem of hyperspectral data is the existence of mixed pixels due to the sensor's insufficient spatial resolution and the spatial complexity [1]. Spectral unmixing aims at separating the mixed pixels into a collection of constituent spectra or endmembers, and estimating their fractional abundances [2].

In some situations, it is a challenge to identify the endmember signatures because of the limited spatial resolution of the sensors and mixtures occurring in different scales. However, the unmixing problem has been approached in semi-supervised fashion with the assumption that the observed signature is in the form of linear combinations of numbers of pure spectral signatures known in advance [1]. In these algorithms, it is not necessary to estimate the endmembers. The number of endmembers existing in a mixed pixel is small compared with the size of the spectral library [2], which means that the fractional abundance is sparse, and then unmixing can be converted to an optimization problem calling for efficient linear sparse regression techniques. To deal with these issues, the unmixing problem can be transformed into a sparse unmixing model that has been widely used in recent years [3]–[6]. The sparse unmixing via variable splitting and augmented Lagrangian (SUnSAL) [3] based on the alternating direction method of multipliers has been proposed to cope with the sparse unmixing problem. It converts the constrained problem into an unconstraint problem. However, SUnSAL only uses spectral information and does not concern relationship between each pixel vector and its neighbors. Sparse unmixing via variable splitting augmented Lagrangian and total variation (SUnSAL-TV) [4] was then developed to exploit the spatial contextual information present in images. Furthermore, non-local sparse unmixing methods [5], [6] were proposed to integrate the non-local spatial information.

Sparse unmixing of hyperspectral data has been attracting attention to deal with the difficulties of unmixing in the presence of complex data and non-convex models. There still exists some problems to deal with sparse unmixing. First, sparse unmixing is faced with the l_0-norm problem which is an NP-hard problem [7]. Traditional sparse unmixing methods relax the non-convex l_0-norm to obtain an approximate solution. Second, these methods have weight parameters in the objective criteria, and those parameters are sensitive to different data sets and have a great impact on the unmixing accuracy [5]. Third, only a single solution with a certain parameter can be obtained in a single run. Little insights into the sparse unmixing problems can be gained for users to make a decision.

It is interesting that all the aforementioned deficiencies can be naturally alleviated by employing multi-objective evolutionary algorithms (MOEAs) [8] for sparse unmixing. MOEAs optimize the conflicting objectives simultaneously and can obtain a set of nondominated solutions approximating the true Pareto-optimal front. Recently, researchers have proved that MOEAs can overcome the difficulty that limits the greedy algorithm in some NP-hard problems, such as minimum set cover problem [9], minimum cost coverage problem [10], ensemble pruning [11] and subset section [12]. In [9], Yu *et al.* used a multi-objective reformulation with an isolation function to efficiently achieve an H_n-approximation ratio for the unbounded set cover problem. Then Qian *et al.* found that multi-objective optimization was more efficient than the penalty function method for obtaining the optimal and approximate solutions on a minimum cost coverage instance [10]. In [11], the error and the size in ensemble pruning were optimized by MOEAs. The theoretical results were derived to reveal the advantage of

The authors are with Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, International Research Center for Intelligent Perception and Computation, Xidian University, No.2 South TaiBai Road, Xi'an 710071, China. E-mail: gong@ieee.org.

This work was supported by the National Natural Science Foundation of China (Grant nos. 61273317 and 61422209), the National Top Youth Talents Program of China, the Specialized Research Fund for the Doctoral Program of Higher Education (Grant no. 20130203110011) and the Fundamental Research Fund for the Central Universities (Grant no. K5051202053).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TEVC.2016.2598858, IEEE Transactions on Evolutionary Computation

MOEAs over the single-objective heuristic optimization methods. For sparse regression, the authors proved that MOEAs were able to achieve the best-so-far theoretically guaranteed approximation performance [12]. In these works, the authors converted a constrained optimization task as a multi-objective problem and optimized the reformulated problem by MOEAs [10]. The performance of MOEAs on NP-hard problems has been theoretically investigated [9]–[12], which encourages us to solve the sparse unmixing problems using MOEAs. It is natural to take the reconstruction error, the sparsity and the total variation regularization term in sparse unmixing as the objectives of multi-objective optimization. Then the objectives are simultaneously optimized by MOEAs to find the best trade-off solutions. The optimal solution is naturally achieved at the knee position of the Pareto front by employing some off-the-shelf tools in multi-objective optimization, such as the angle-based method [13]. We can find a better compromise between two or more competing cost function terms automatically without tuning weight parameters.

In this paper, a multi-objective sparse unmixing (MOSU) model is proposed. We consider the reconstruction term, the sparsity term and the total variation regularization term as the conflicting objectives. Then a novel multi-objective cooperative coevolutionary algorithm is proposed to optimize these objectives simultaneously. In the proposed algorithm, a random group strategy based on sparsity is introduced and the non-uniform mutation operator is modified to obtain more sparse solutions. Then a cooperation strategy is used among the species by using the information of knee solution. The proposed method optimizes the non-convex l_0-norm directly and can obtain nondominated solutions in a single run. In the traditional sparse unmixing methods, it is a difficult task for users to make a decision (choose appropriate parameters), who encounter the complex data and non-convex model and run the algorithm many times to make decisions (modify the parameters). The proposed technique directly optimizes the non-convex model to generate a set of non-dominated solutions in a single run. Then it is easy for users to make a decision based on these solutions.

The rest of this paper is structured as follows. Section II describes the sparse unmixing theory and our motivation for using MOEAs. Section III describes the proposed MOSU in detail. Experimental studies are shown in Section IV. Section V draws the conclusion of our work in this paper.

II. Background and Motivation

We will use the following notations throughout the paper:

- $\mathbf{A} \in \mathbf{R}^{L\times m}$ denotes the available spectral library. L is the number of spectral bands, and m denotes the number of endmembers in library $\mathbf{A}$.
- $\mathbf{Y} \in \mathbf{R}^{L\times n}$ denotes the observed data matrix. $\mathbf{X} \in \mathbf{R}^{m\times n}$ is the fractional abundances matrix. n is the number of the pixels in the total observed image.
- $\mathbf{y} \in \mathbf{R}^{L\times 1}$ is the measured spectrum of the pixel. $\mathbf{x} \in \mathbf{R}^{m\times 1}$ is the fractional abundance vector.
- $\mathbf{M} \in \mathbf{R}^{L\times q}$ is a endmember matrix containing q endmembers.
- $\boldsymbol{\alpha} \in \mathbf{R}^{q\times 1}$ is a vector containing the fractional abundances of the endmembers in the pixel.
- $\mathbf{n} \in \mathbf{R}^{L\times 1}$ is a vector collecting the errors affecting the measurements at each spectral band.
- $\iota_{\mathbf{S}}$ denotes the function of the set $\mathbf{S}$ (e.g., $\iota_{\mathbf{S}}(\mathbf{x}) = 0$ if $\mathbf{x} \in \mathbf{S}$, and $\iota_{\mathbf{S}}(\mathbf{x}) = \infty$ if $\mathbf{x} \notin \mathbf{S}$).

A. Linear Spectral Mixture Model

According to mixing scales at each mixed pixel, the mixture is linear or nonlinear [2]. The nonlinear model considers the mixing scale as intimate mixtures [14], while the linear model holds that the mixing scale is macroscopic and considers the observed spectrum of any mixed pixels as a linear combination of the endmembers signatures weighted by their corresponding abundances [15]. For the spectral signature of each pixel, the linear model can be written as follows:

$$y_i = \sum_{j=1}^{q} m_{ij}\alpha_j + n_i, \tag{1}$$

where y_i denotes the measured reflectance of a pixel at spectral band i, m_{ij} is the reflectance of the jth endmember at spectral band i, α_j stands for the fractional abundance of the jth endmember, and n_i represents an errors term for the spectral band i (e.g., noise and modeling errors). If the data has L spectral bands, Eq. (1) can be rewritten as follows:

$$\mathbf{y} = \mathbf{M}\boldsymbol{\alpha} + \mathbf{n}. \tag{2}$$

Two constraints are usually applied to the fractional abundances of the endmembers [16], the abundance non-negativity constraint (or ANC) and the abundance sum-to-one constraint (or ASC), which are shown as follows:

$$\begin{cases} ANC: & \alpha_j \geq 0, j = 1, 2, \ldots, q \\ ASC: & \sum_{j=1}^{q} \alpha_j = 1 \end{cases}. \tag{3}$$

B. Sparse Unmixing

The linear sparse unmixing aims at finding a linear combination of a few endmembers in a large spectral library for each observed mixed pixel [1]. Then the sparse unmixing models can be formulated as

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}. \tag{4}$$

Due to the fact that only a few endmembers contained in $\mathbf{A}$ contribute to an observed mixed pixel $\mathbf{y}$, $\mathbf{x}$ is sparse. The sparse unmixing problem for each mixed pixel can be formulated as

$$\begin{aligned} &\min_{\mathbf{x}} \quad \|\mathbf{x}\|_0 \\ &\text{s.t.} \quad \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \delta, \mathbf{x} \geq \mathbf{0}, \mathbf{1}^{\mathbf{T}}\mathbf{x} = 1, \end{aligned} \tag{5}$$

where $\|\mathbf{x}\|_0$ denotes the number of non-zero element of $\mathbf{x}$, and $\delta \geq 0$ is the errors tolerance because of noise and modeling errors.

Eq. (5) is an NP-hard problem, which is difficult to solve in a straightforward way. Candes and Tao proved that under

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TEVC.2016.2598858, IEEE Transactions on Evolutionary Computation

certain condition of the restricted isometric property (RIP) [7], [17], the l_0-norm can be replaced by the l_1-norm. Then the problem shown in Eq. (5) can be considered as

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|\mathbf{x}\|_1 \\ \text{s.t.} \quad & \|\mathbf{y}-\mathbf{A}\mathbf{x}\|_2 \leq \delta, \mathbf{x} \geq \mathbf{0}, \mathbf{1}^{\mathbf{T}}\mathbf{x} = 1, \end{aligned} \tag{6}$$

where $\|\mathbf{x}\|_1$ is the l_1-norms of $\mathbf{x}$. SUnSAL [3] converts the constraint problem into an unconstraint problem as follows :

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{A}\mathbf{x}-\mathbf{y}\|_2^2 + \lambda\|\mathbf{x}\|_1 + \iota_{\{1\}}(\mathbf{1}^{\mathbf{T}}\mathbf{x}) + \iota_{\mathbf{R}_+^m}(\mathbf{x}), \tag{7}$$

where $\lambda \geq 0$ is a parameter controlling the relative weight of the sparsity of the solution, $\iota_{\{\mathbf{1}\}}(\mathbf{1}^{\mathbf{T}}\mathbf{x})$ and $\iota_{\mathbf{R}_+^m}(\mathbf{x})$ stand for the ANC and ASC, respectively. Nevertheless, SUnSAL uses only spectral information and does not consider the relationship between each pixel vector and its neighbors. SUnSAL-TV is then put forward to exploit the spatial contextual information present in images [4]. SUnSAL-TV utilizes the spatial information in the sparse unmixing formulation by means of the TV regularizer. The model is shown as follows:

$$\min_{\mathbf{X}} \frac{1}{2}\|\mathbf{A}\mathbf{X}-\mathbf{Y}\|_2^2 + \lambda\|\mathbf{X}\|_{1,1} + \iota_{\mathbf{R}_+}(\mathbf{X}) + \lambda_{TV}TV(\mathbf{X}), \tag{8}$$

where

$$TV(\mathbf{X}) \equiv \sum_{\{i,j\}\in\varepsilon} \|\mathbf{x}_i - \mathbf{x}_j\|_1, \tag{9}$$

and $\mathbf{x}_j$ represents the neighboring pixels of $\mathbf{x}_i$ in $\mathbf{X}$, and ε stands for the set of the horizontal and vertical neighbors in $\mathbf{X}$.

C. Motivation of Using MOEAs for Sparse Unmxing

Without loss of generality, the multi-objective optimization problems (MOPs) in this paper can be set for minimization. An MOP can be stated as the following:

$$\begin{aligned} \min_{\mathbf{x}} F(\mathbf{x}) &= (f_1(\mathbf{x}), \cdots, f_m(\mathbf{x}))^T \\ \text{s.t.}\, \mathbf{x} &\in \Omega, \end{aligned} \tag{10}$$

where Ω is the feasible space, $\mathbf{x}$ is a solution to the MOP, R^m is the objective space, and the $F : \Omega \to R^m$ consists of m real-valued objective functions. In most instances, the objectives in an MOP are contradictory to each other, which means no point in feasible space can minimize all the objectives simultaneously. Hence, multi-objective optimization [8], [18] are designed to find the best trade-off relationship among them simultaneously.

For minimization, a solution $\mathbf{x}_u$ is said to dominate another solution $\mathbf{x}_v$, if and only if

$$\begin{aligned} &\forall i = 1, 2, \cdots, m \ f_i(\mathbf{x}_u) \leq f_i(\mathbf{x}_v) \\ \wedge &\exists j = 1, 2, \cdots, m \ f_i(\mathbf{x}_u) < f_i(\mathbf{x}_v). \end{aligned} \tag{11}$$

And a point $\mathbf{x}^*$ in Ω is called a Pareto optimal solution to (1) on condition that there is no such point $\mathbf{x}$ in Ω that makes $F(\mathbf{x})$ dominate $F(\mathbf{x}^*)$. Then $F(\mathbf{x}^*)$ is termed as Pareto optimal vector. The objectives in a Pareto optimal vector have such relationship: a decrease in one objective causes an increase in the others. All the Pareto optimal points constitute a set called Pareto optimal set [19], and their corresponding Pareto optimal objective vectors are called the Pareto optimal front (PF) [19]. For multi-objective optimization, it has been recognized that evolutionary algorithms (EAs) are well suited because EAs can deal with a set of possible solutions simultaneously [8], [20]. Since [21], various EAs to deal with MOPs have been proposed such as [22]–[31] and these EAs are termed as multi-objective evolutionary algorithms (MOEAs). MOEAs seek to obtain a set of Pareto optimal solutions for approximating the true PF in a single run.

The traditional sparse unmixing methods replace the l_0-norm by l_1-norm under a certain conditions of RIP. Candes and Tao [7] have proved that it is possible to recover $\mathbf{f}$ exactly from $\mathbf{y}$ if the vector of errors $\mathbf{e}$ is sparse enough, where, $\mathbf{f} \in \mathbf{R}^n$ is a vector to be recovered, $\mathbf{y} = \mathbf{A}\mathbf{f} + \mathbf{e}$ is a corrupted measurement, $\mathbf{A}$ is an $m \times n$ matrix, and $\mathbf{e}$ is an arbitrary and unknown vector of errors. However, in the sparse unmixing formulation $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}$, the errors term $\mathbf{n}$ is not as sparse as $\mathbf{e}$. Therefore the condition may not usually be satisfied during sparse unmixing. Evolutionary optimization has been widely used to solve non-convex optimization problems [8], [32]–[35]. In this paper, we use the evolutionary optimization to solve the non-convex problem directly, and do not need to consider the condition of RIP. Furthermore, many traditional sparse unmixing methods have parameters in the objective criteria and some of them are sensitive to different data sets. Such parameters have an important influence on the unmixing accuracy, which are difficult to tune. The regularization parameters have a great influence on the relative contribution of the data fidelity, sparse constraint and spatial relationships. To solve the difficulty of the determination of the parameters, we model the sparse unmixing problem as an MOP instead of scalar objective functions.

In this paper, a multi-objective cooperative coevolutionary algorithm is proposed for sparse unmixing. Due to the fact that the spectral library in sparse unmixing usually contains hundreds spectral samples, it will suffer high-dimensional optimization problems when we use the evolutionary optimization for sparse unmixing. In order to deal with the high-dimensional problems, cooperative coevolutionary scheme [36] is used in the proposed methods because it has been proved effective to large scale global optimization problems [37]–[43].

III. Methodology

In this section, the proposed multi-objective cooperative coevolutionary algorithm for hyperspectral sparse unmixing is described in detail. First, the multi-objective sparse unmixing model is introduced, and then the proposed multi-objective cooperative coevolutionary algorithm is described. In the proposed algorithm, a novel grouping strategy based on sparsity is designed. Then we modify the mutation operator by introducing a threshold to get more sparse solutions. A cooperation strategy is used among the species by using the information of knee solution. Finally, the analysis of the proposed multi-objective sparse unmixing model is given.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TEVC.2016.2598858, IEEE Transactions on Evolutionary Computation

A. Multi-objective Sparse Unmixing Model

Sparse unmixing searches for a best combination of a number of pure spectral signatures in a large spectral library for each observed pixel. According to the sparse unmixing model (5), the sparse unmixing can be considered as reconstructing the observed pixel with less error by using least spectral signatures from a large spectral library. Traditional sparse unmixing algorithms transform the competing cost function terms into one single term by constructing an aggregate objective function. The aggregate objective function involves weighted parameters which are sensitive to different data sets, difficult to tune and significantly influence the unmixing accuracy [5]. In order to address these drawbacks, we model the sparse unmixing problem as a multi-objective optimization problem and simultaneously optimize these conflicting objectives. We are the first to model the sparse unmixing problem as a multi-objective optimization problem. In general, two objectives, the reconstruction error $\mathbf{RSE} = \|\mathbf{Ax} - \mathbf{y}\|_2^2$ and the sparsity $\mathbf{SPA} = \|\mathbf{x}\|_0$ (the numbers of non-zero element of $\mathbf{x}$), are considered in the sparse unmixing. For each pixel, the sparse unmixing model can be formulated as the following bi-objective optimization model (Bi-MOSU):

$$\begin{aligned} \min_{\mathbf{x}} \quad & (\mathbf{RSE}, \mathbf{SPA}) \\ \text{s.t.} \quad & \mathbf{x} \geq \mathbf{0}, \mathbf{1^T x} = 1, \end{aligned} \tag{12}$$

where

$$\mathbf{RSE} = \|\mathbf{Ax} - \mathbf{y}\|_2^2, \mathbf{SPA} = \|\mathbf{x}\|_0. \tag{13}$$

However, these two objectives only use spectral information without considering relationship between each pixel vector and its neighbors. Total variation regularization accounts for the spatial homogeneity, which means that two neighboring pixels might have similar fractional abundances of the same endmember [4]. Therefore we consider the total variation regularizer as the third objective (**SNI**) to incorporate the spatial neighbor information. Then the tri-objective optimization model (Tri-MOSU) for sparse unmixing can be formulated as follows:

$$\begin{aligned} \min_{\mathbf{x}} \quad & (\mathbf{RSE}, \mathbf{SPA}, \mathbf{SNI}) \\ \text{s.t.} \quad & \mathbf{x} \geq \mathbf{0}, \mathbf{1^T x} = 1, \end{aligned} \tag{14}$$

where

$$\mathbf{SNI} = \sum_{j \in \varepsilon} \|\mathbf{x} - \mathbf{x}_j\|_1. \tag{15}$$

where $\mathbf{x}_j$ represents the neighboring pixels of $\mathbf{x}$, and ε stands for the set of the horizontal and vertical neighbors of $\mathbf{x}$ [4].

The schematic diagram of the proposed multi-objective sparse unmixing model is summarized in Fig. 1. The standard spectral library and observed image are considered as the inputs for the multi-objective sparse unmixing model. We choose the Eq. (12) or Eq. (14) as the objective functions, and then MOEAs are used to optimize them. Once the termination criterion is satisfied, the knee solution is selected from the PF for the sparse unmixing problem. All the outputs compose the abundance map.

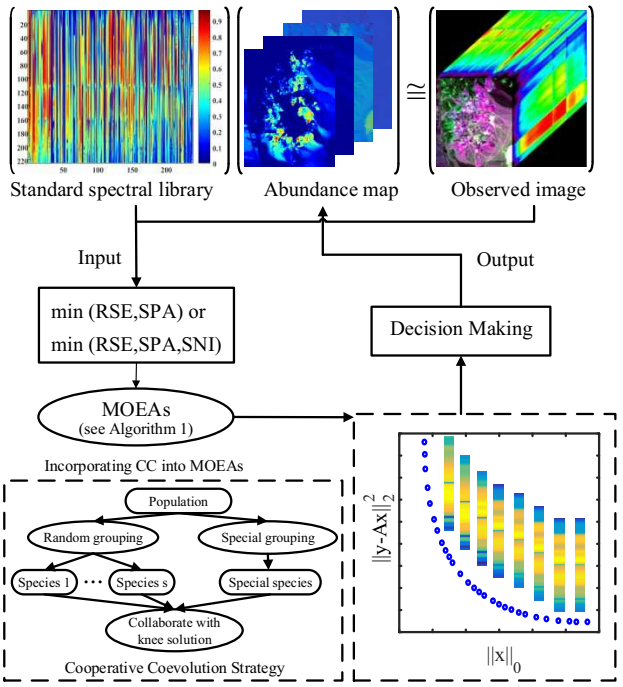


Fig. 1. The schematic diagram of the multi-objective sparse unmixing model.

B. Proposed Multi-objective Cooperative Coevolutionary Algorithm for Sparse Unmixing

In our objective functions, Eq. (12) and Eq. (14), the available spectral library $\mathbf{A} \in \mathbf{R}^{L \times m}$ usually contains hundreds of spectral samples, which means that m is very large, e.g., 2000. $\mathbf{x} \in \mathbf{R}^m$ is the fractional abundance vector of a mixed pixel $\mathbf{y}$, which means $\mathbf{x}$ will be a 2000-dimensional decision vector to be estimated. When we code $\mathbf{x}$ and search $\mathbf{x}$ directly using the evolutionary algorithms, the search space of the decision vectors will be very large and the computation will be excessive. Suffered from high-dimensional optimization problem, the performance of EAs deteriorates with the increase of the dimensionality of the search space [44]. EAs can find good optimal solutions in low-dimension problems but may not perform well on high-dimensional problems. To deal with the high-dimensional optimization problems, cooperative coevolution was proposed and proved effective to large scale global optimization problems [36], [40], [45]. Inspired by these works, a multi-objective cooperative coevolutionary algorithm is proposed to deal with the high-dimensional optimization problems in sparse unmixing, which is shown in **Algorithm 1**.

For each mixed pixel, we use **Algorithm 1** to estimate the corresponding fractional abundance vector $\mathbf{x}$. First, we generate the initial population P^t randomly and each individual in P^t represents a complete solution of the mixed pixel. Then we obtain the PF by using nondominated sorting [22] and select the knee point on the PF as the representative individual P_{best} of P^t. Then the decision vector of m dimension is divided into n subcomponents of s dimension randomly, where $m = n \times s$. Meanwhile, as shown in Fig. 2, n subpopulation (species) are created and each one has pop individuals. Then the method

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TEVC.2016.2598858, IEEE Transactions on Evolutionary Computation

executes the cycling for a predefined generations.

Algorithm 1 Algorithm of multi-objective sparse unmixing.

Input: $Maxt$: the maximum number of generations in main loop, $Maxgen$: the maximum number of generations in the process of cooperation. pop: the number of individuals in the population.

Output: The fractional abundance vector.

Step 1) **Initialization:** Set $t = 0$ and generate the initial population $P^t = \{\mathbf{x_1}, \mathbf{x_2}, \cdots, \mathbf{x_{pop}}\}$.
Step 2) **Cycling**
  Step 2.1) Calculate the fitness of the population P^t and obtain the PF.
  Step 2.2) Obtain the knee point on the PF as the best individual (P_{best}) in the population P^t.
  Step 2.3) Split the population P^t into n species, and obtain the non-zeros subpopulation (species), $sp = \{sp_1, sp_2, \cdots, sp_n, sp_{non}\}$
  Step 2.4) The process of cooperation.
for $i = 1, 2, \cdots, n$, **do**
    Step 2.4.1) Set $gen = 0$.
    Step 2.4.2) *Mutation and Crossover*: Generate a new species sp_{mc} with mutation and crossover operators.
    Step 2.4.3) *Collaboration*: Combine sp_{mc} and sp_i, i.e. $sp_{com} = sp_{mc} \cup sp_i$, and then generate the population P^t_{com} by forming the collaboration with sp_{com} and P_{best}.
    Step 2.4.4) *Selection*: Calculate the fitness of the population P^t_{com} and obtain the non-dominated set; select the next generation P^{t+1}_{com} from P^t_{com} using elite-preserving operator [22].
    Step 2.4.5) Extract the evaluated species sp_i from P^{t+1}_{com}; If the $gen < Maxgen$, $gen = gen + 1$ and go to Step 2.4.2 for the next generation.
end for
Step 2.5) Update the population P^t.
Step 3) Cycling termination criterion: if $t < Maxc$, set $t = t + 1$ and go to Step 2 for the next cycle. Otherwise, stop the Cycling.
Step 4) Obtain the optimal PF of the population P^t, and find the knee point on the PF as the estimated fractional abundance vector.

1) Representation and Initialization: Each individual $\mathbf{x} = \{x_1, x_2, \ldots, x_m\}$ in P^t stands for a fractional abundance vector, where t is the current generation and m is the number of endmembers in the library. A good initialization mechanism may speed up convergence as well as promote population diversity. In this paper, in order to reduce the searching space as well as promote population diversity, the initial population P^0 is generated partly by SUnSAL or SUnSAL-TV and the rest is produced randomly.

2) Grouping strategy based on sparsity: In hyperspectral sparse unmixing, the fractional abundance vector of a pixel must fulfill the abundance sum-to-one constraint, which means that the decision vectors interact with each other and are interdependent. The early divide-and-conquer strategy failed to deal with nonseparable problems because of neglecting the interaction among different decision variables, which splits the high-dimensional decision variables into some low-dimensional decision vectors and evolves these decision vectors cooperatively with a predefined number of cycles. To solve the nonseparable problem, random groups scheme was introduced [37], [40] and theoretical analysis was presented to show why the scheme is effective for optimizing large nonseparable problems [37].

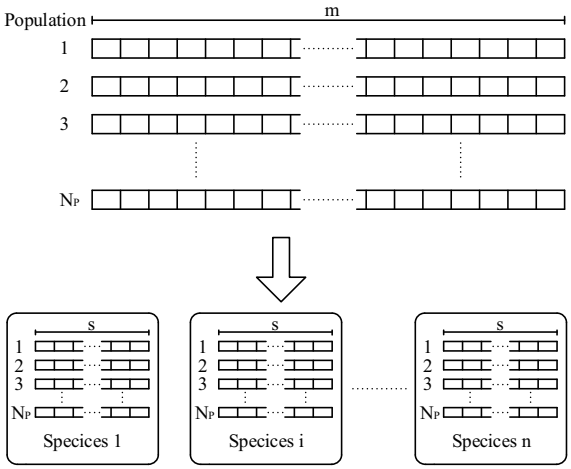


Fig. 2. Creation of the subpopulations (species). Here we assume that the population is with the size of $N_p \times m$, and it is split into n subpopulations (species) randomly. Each species is with the size of $N_p \times s$, where $m = n \times s$.

Due to the fact that only a few of spectral samples existing in a mixed pixel, which means that the fractional abundance is sparse. Each individual $\mathbf{x} = (x_1, x_2, \ldots, x_m)$ in the population is a sparse decision vector. Only a few of non-zero elements in $\mathbf{x}$, and they are the main elements that contribute to the observed mixed pixel $\mathbf{y}$. Therefore a grouping strategy based on sparsity is proposed for sparse unmixing. We select the non-zero elements of each individual in the population to form a species. For example, there are two individuals $\mathbf{a}$ and $\mathbf{b}$, only the $i-$th, $j-$th, $k-$th and $l-$th elements of $\mathbf{a}$ are non-zero, and only the $i-$th, $j-$th and $l-$th elements of b are non-zero. Then a and b can be formulated as $\mathbf{a} = (0, ..., a_i, ..., a_j, ..., a_k, ..., a_l, ..., 0)$ and $\mathbf{b} = (0, ..., b_i, ..., b_j, ..., b_l, ..., 0)$, respectively. When the mutation operator and crossover operator are implemented to combine $\mathbf{a}$ and $\mathbf{b}$ to form a new individual $\mathbf{ab}$, the $s-$th element of $\mathbf{ab}$ is more likely to be zeros if both the $s-$th element of $\mathbf{a}$ and the $s-$th element of $\mathbf{b}$ are zeros. The $s-$th element of $\mathbf{ab}$ have a great possibility being non-zero if the $s-$th element of $\mathbf{a}$ or $\mathbf{b}$ is non-zero. Compared with the zero elements, the non-zero elements have a greater possibility of contributing to an observed mixed pixel $\mathbf{y}$. Therefore it is reasonable to use only non-zeros elements of each individuals to compose a special species, named non-zeros species.

3) Genetic operators with sparseness constraints: In the proposed algorithm, in order to obtain a more sparse solution, we use an improved non-uniform mutation [46] operator based on sparsity. The modified non-uniform mutation operator is

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TEVC.2016.2598858, IEEE Transactions on Evolutionary Computation

shown as follows:

$$x_i'' = \begin{cases} 0, & \text{if } \mu > M \\ x_i', & \text{otherwise} \end{cases} \tag{16}$$

where x_i' is the non-uniform mutated result of x_i which is one element of a parent solution $\mathbf{x} = (x_1, \ldots, x_i, \ldots, x_n)$, x_i'' is the modified mutated result of x_i', $\mu \in [0, 1]$ is a random number, and M is a constant with value 0.9. This will increase the chance for the solution to be more sparse. A modified arithmetical crossover [47] operator is used in this paper, which is formulated as follows:

$$\mathbf{z} = \lambda \times \mathbf{x} + (1 - \lambda) \times \mathbf{y} \tag{17}$$

where $\mathbf{z} = (z_1, \ldots, z_i, \ldots, z_n)$ is a new solution, $\mathbf{x} = (x_1, \ldots, x_i, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_i, \ldots, y_n)$ are two parent solutions. Due to the abundance non-negativity constraint, for each element z_i of solution $\mathbf{z}$, we have the following constraint:

$$z_i = \begin{cases} 0, & \text{if } z_i < 0 \\ z_i, & \text{otherwise} \end{cases} \tag{18}$$

As for the abundance sum-to-one constraint, then we normalize solution $\mathbf{z}$ by its sum of all its elements:

$$\mathbf{z} = \mathbf{z} / \sum_{i=1}^{n} z_i \tag{19}$$

4) Cooperation with the knee solution: The main idea of C-C strategy is to divide the high-dimensional problem into some low-dimensional subcomponents (species) and evolve these species cooperative for a predefined cycling. In one cycling, one complete evolution of each species is done for a predefined number. In the process of the cooperation, a subcomponent from a species should cooperate with other subcomponents from other species so as to determine its fitness, which requires a lot of computation and great complexity. A number of cooperation strategies for multi-objective optimization have been put forward [48]–[50], and only a certain number of subcomponents in other species are used to determine the fitness of an individual.

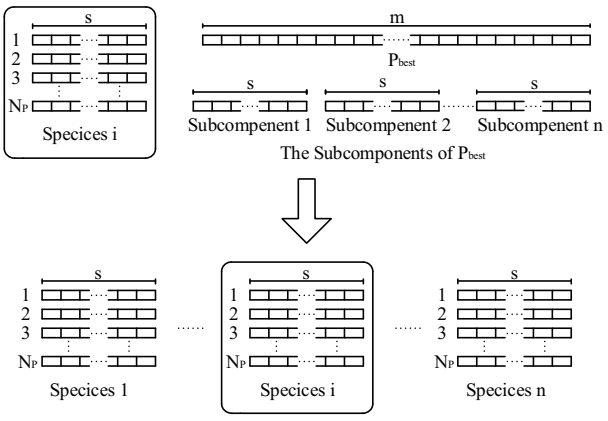


Fig. 3. Cooperation with the knee solution.

The knee solution is the most interesting solution on the PF and is considered as the best solution for sparse unmixing problem. The knee solution acts as a effective leader to guide each species to evolve toward a certain direction and speeds up the convergence of the proposed algorithms. For this reason, each subcomponent in a species are cooperated with the knee solution to form a complete solution. The framework of cooperation among the a species is shown in Fig. 3. Then the fitness of the complete solution is calculated and assigned back to the corresponding subcomponent.

C. Analysis of the Proposed Sparse Unmixing Method

In order to deal with high-dimensional optimization problem, we have proposed a multi-objective cooperative coevolution algorithm for sparse unmixing. The cooperative coevolution strategy includes two important parts, grouping scheme and cooperation strategy. A random grouping scheme is proposed to deal with nonseparable problems [37]. This scheme gives each element of the decision vector the same opportunity to evolve. However, in the sparse unmixing problem, the fractional abundance vector is sparse, which means only a few of non-zeros elements of the decision vector contribute to reconstruct the mixed pixel. Therefore it is reasonable to compose these non-zeros elements into a group (or species), and give these element more chance to evolve. In the cooperation scheme, we have proposed a cooperation strategy with the knee solution. The knee solution is the expected trade-off between two objective functions. We consider it as the best representative solution for the sparse unmixing problem. Like the cooperation in single objective optimization, we have the "best" individual and each species collaborates with the knee solution, which will avoid the large amount of calculation.

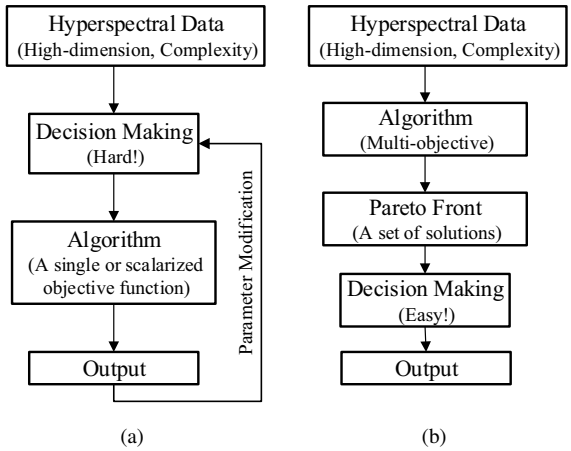


Fig. 4. Comparison between the traditional sparse unmixing model and the proposed multi-objective sparse unmixing model. (a) the traditional sparse unmixing model. (b) the proposed multi-objective sparse unmixing model.

Fig. 4 shows the comparison between the traditional sparse unmixing model and the proposed multi-objective sparse unmixing model. In the traditional sparse unmixing methods, the users must make a decision to choose the regularization parameters before algorithm execution. However, those parameters play an important role in controlling the unmixing accuracy, which are hard to tune because different data sets are acquired in different conditions. Furthermore, the traditional

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TEVC.2016.2598858, IEEE Transactions on Evolutionary Computation

sparse unmixing methods relax the l_0-morn with i_1-morn, while the l_1-morn might not always be equivalent to l_0-morn in the sparse unmixing problem. In the proposed method, the reconstruction term, the sparsity term, or the total variation regularization term are simultaneously optimized by the proposed multi-objective cooperative coevolutionary algorithm to gain more insights into the sparse unmixing task. We tackle the original sparse unmixing problem directly and a set of solutions can be obtained by the proposed method. Then the users can make a decision based on these meaningful solutions, which is a simple job. Because the proposed method is accomplished with evolutionary algorithms, MOSU costs more time in computer than the other spare unmixing algorithms. Although the proposed method costs more time in computer, the proposed technique reduces human resources costs because MOSU can generate a set of solutions and then the user can make a decision easily based on these solutions. Furthermore, as we all know, evolutionary algorithms are very easy to parallelize. For practical applications, the parallel processing will be considered in our future work for accelerating our algorithm.

IV. Experimental Study

In this section, we demonstrate the effectiveness of Bi-MOSU and Tri-MOSU for sparse unmixing. We first test the parameter M and the proposed cooperative coevolutionary strategy. Then four data sets are used to test the performance of the proposed algorithms. The abundance non-negativity constraint and abundance sum-to-one constraint are taken into account in the proposed algorithms. In the experiments, the population size is 100, the number of cycling is set to 100 in main loop, and the maximum generation $Maxgen$ is set to 20 for the evolution process of each species. The mutation rate and crossover rate are 0.1 and 0.8, respectively. All the experimental results are based on one run that is the nearest to the average of 20 independent runs. It does not make sense to show the average results of all the independent runs because the results are displayed in the form of images, which is different from the numerical optimization problems.

A. Experiments on the MOSU Model

This section describes several experiments to investigate the parameter M and the proposed cooperative coevolutionary strategy. A simulated data set [4] is used in the experiments, which is an image of size 75×75 and with 224 bands per pixel. In order to measure the diversity in the population, the coverage of two sets is calculated in the experiments [51]. Let $\mathbf{A}$, $\mathbf{B}$ be two approximate Pareto-optimal sets. The function I_C maps the ordered pair $(\mathbf{A}, \mathbf{B})$ to the interval [0,1]:

$$I_C(\mathbf{A},\mathbf{B}) = \frac{|\mathbf{b} \in \mathbf{B}; \exists \mathbf{a} \in \mathbf{A} : \mathbf{a} \succeq \mathbf{b}|}{|\mathbf{B}|} \qquad (20)$$

where $\succeq$ means dominate or equal (also called weakly dominate). The value $I_C(\mathbf{A},\mathbf{B}) = 1$ means that all decision vectors in $\mathbf{B}$ are weakly dominated by $\mathbf{A}$. $I_C(\mathbf{A},\mathbf{B}) = 0$ implies no decision vectors in $\mathbf{B}$ is weakly dominated by $\mathbf{A}$.

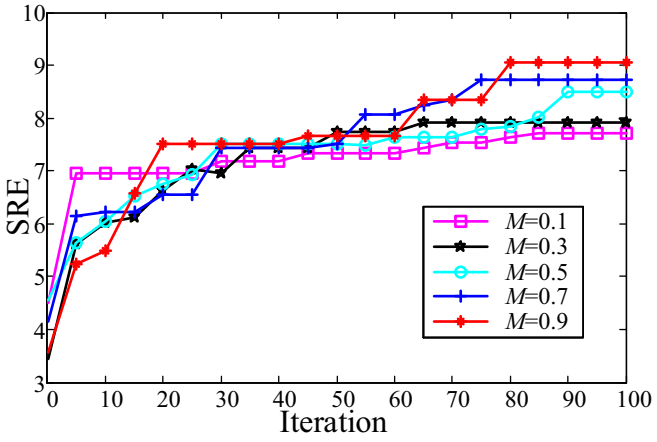


Fig. 5. The values of SRE during iterations with parameter M = 0.1, 0.3, 0.5, 0.7 and 0.9.

1) Test of the Parameter M: Fig. 5 graphs the values of SRE during iterations with parameter M = 0.1, 0.3, 0.5, 0.7 and 0.9. As shown in Fig. 5, the algorithm with $M = 0.1$ has a rapid convergence and cannot obtain better results than others when the iterations continue. The algorithm has the problem of premature convergence and obtains a bad result when the value of M is small. Therefore, in this paper, the value of M is set to 0.9 to get better results.

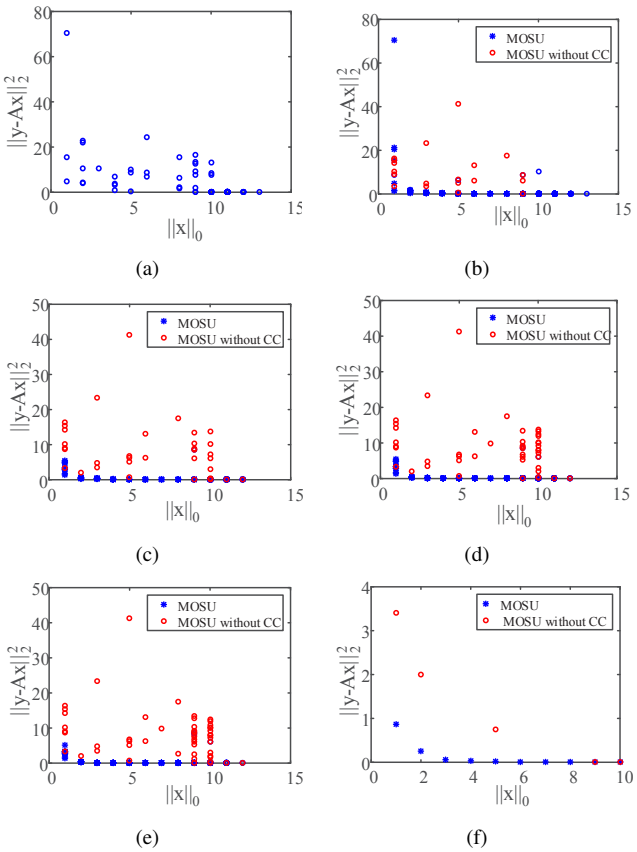
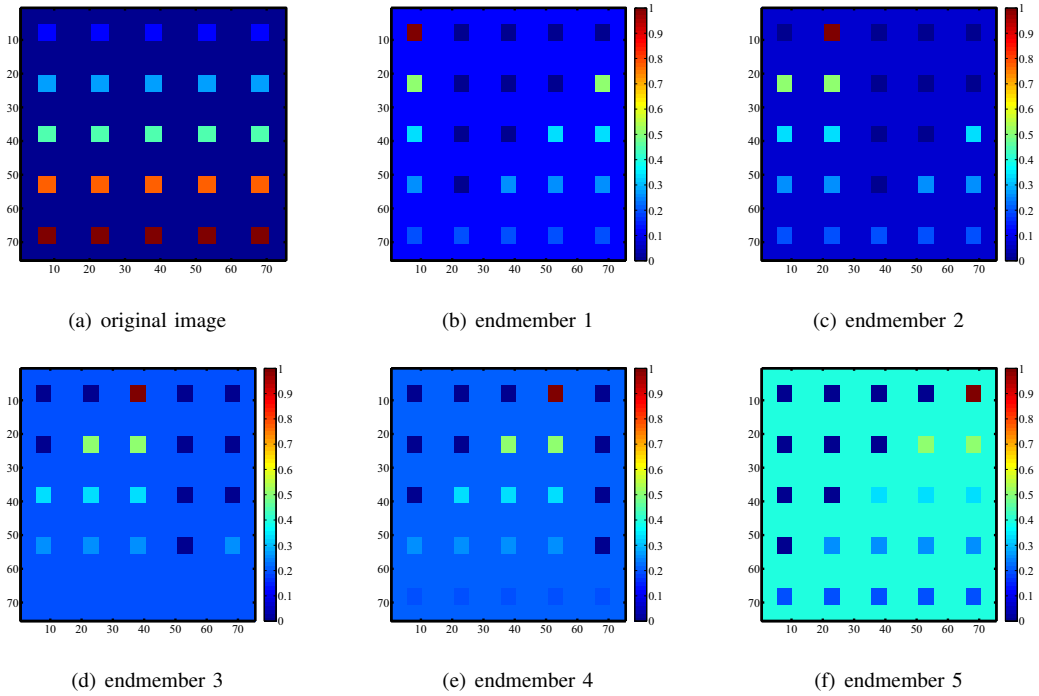


Fig. 6. The distribution of population with change in generation.

2) Test of the Cooperative Coevolutionary Strategy: In order to demonstrate the effectiveness of the proposed cooper-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TEVC.2016.2598858, IEEE Transactions on Evolutionary Computation



(a) original image (b) endmember 1 (c) endmember 2

(d) endmember 3 (e) endmember 4 (f) endmember 5

Fig. 8. True fractional abundances of data 1.

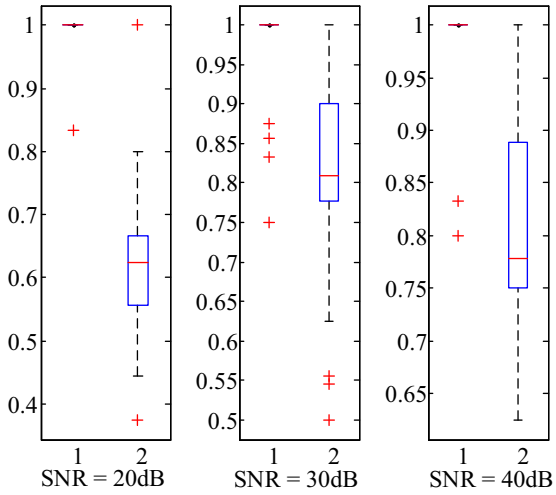


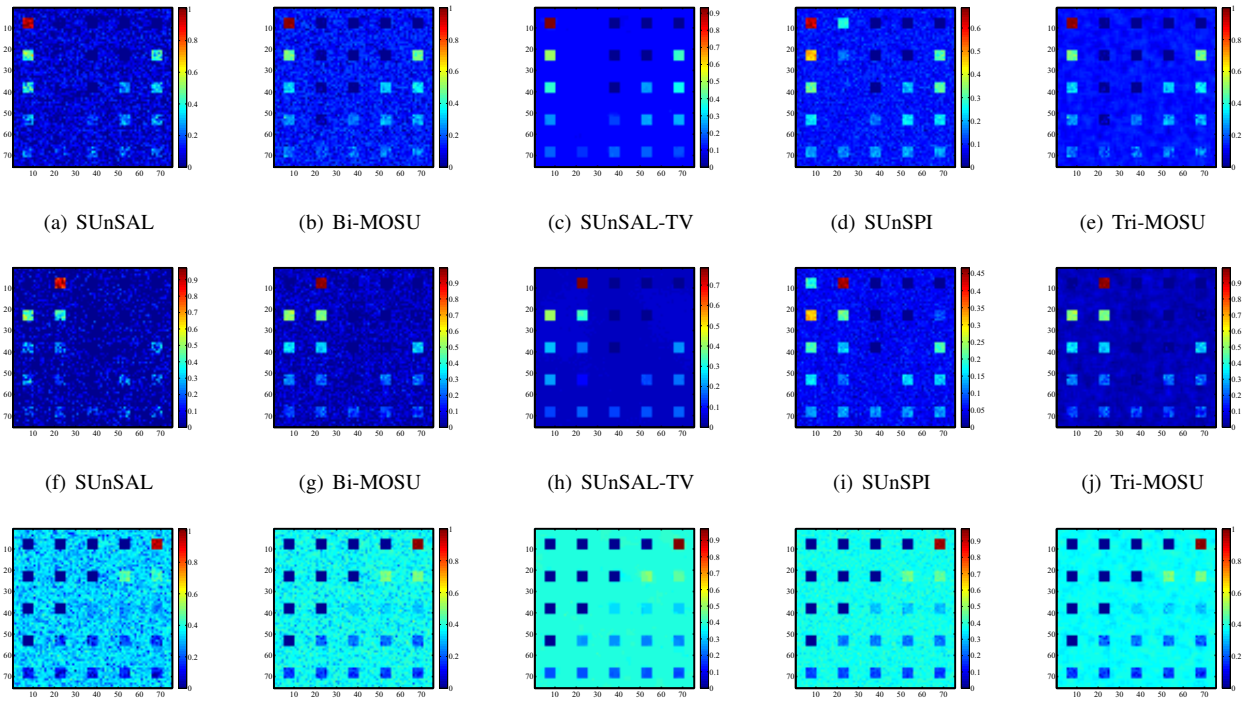
Fig. 7. Statistical values of the coverage of the two sets obtained by MOSU and MOSU without CC. $\mathbf{A}$ and $\mathbf{B}$ are obtained by MOSU and MOSU without CC, respectively. In each plot, the left box represents the distribution of $I_C(\mathbf{A},\mathbf{B})$ and the right box represents the distribution of $I_C(\mathbf{B},\mathbf{A})$.

ative coevolutionary strategy, we test the MOSU and MOSU without the cooperative coevolutionary strategy (MOSU without CC) in the sparse unmixing problems. MOSU without CC is used as the competing algorithm by removing the cooperative coevolutionary strategy from MOSU and keep the other strategies such as improved non-uniform mutation and arithmetical crossover operators. The population size is set to 100 for these two algorithms, and the number of species is set to 10 for MOSU. The number of iteration is 2000 for MOSU without CC and the number of cycling is set to 100 for MOSU, respectively. And the number of iteration per species is set to 20 for MOSU, so that MOSU and MOSU without CC have the same number of total iteration. The mixed pixel is selected from a simulated data set used in the experiment randomly. The changing trends of the population and the PFs of those two algorithms are shown in Fig. 6.

The initial population is shown in Fig. 6(a). Fig. 6(b)-(e) show the changes of population of the two algorithms and Fig. 6(f) shows the PFs obtained by the two methods. It is obvious that MOSU can speed up the convergence and has a better diversity of the population than MOSU without CC. The two PFs contain only a few solutions because each mixed pixel only contains several endmembers. The vector of decision variables is sparse and the objective function $\mathbf{SPA} = \|\mathbf{x}\|_0$ is discrete in the horizontal axis. Therefore the PFs of the two algorithms are different from other continuous problems. In the following experiments, we performed 20 independent runs on the simulated data set with SNR = 20, 30 and 40dB. Fig. 7 shows the statistical values of the coverage of the two sets obtained by MOSU and MOSU without CC. The two approximate Pareto-optimal sets $\mathbf{A}$ and $\mathbf{B}$ are obtained by MOSU and MOSU without CC, respectively. Then we calculate the coverage of two sets $I_C(\mathbf{A},\mathbf{B})$ and $I_C(\mathbf{B},\mathbf{A})$ by Eq. (20). Obviously, the values of $I_C(\mathbf{A},\mathbf{B})$ are higher than those of $I_C(\mathbf{B},\mathbf{A})$. Therefore, the results shown in Fig. 7 indicate that the proposed cooperative coevolutionary strategy is efficient.

From Fig. 6(f) and 7, the PF of MOSU contains more solutions and has a better population diversity than that of MOSU without CC, which is benefited from the proposed cooperative coevolutionary strategy. The random groups scheme treats all the elements equally and gives the same opportunity for each element to evolve. Some elements that really contribute to

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TEVC.2016.2598858, IEEE Transactions on Evolutionary Computation



(a) SUnSAL (b) Bi-MOSU (c) SUnSAL-TV (d) SUnSPI (e) Tri-MOSU

(f) SUnSAL (g) Bi-MOSU (h) SUnSAL-TV (i) SUnSPI (j) Tri-MOSU

(k) SUnSAL (l) Bi-MOSU (m) SUnSAL-TV (n) SUnSPI (o) Tri-MOSU

Fig. 9. Estimated abundance maps of endmember 1, 2, and 5 for data 1 obtained by different unmixing methods.

TABLE I
COMPARISONS OF THE PROPOSED ALGORITHMS AGAINST OTHER ALGORITHMS ON DATA 1.

SNR (dB)	SUnSAL	Bi-MOSU	SUnSAL-TV	SUnSPI	Tri-MOSU
20	4.6725	*5.8406*	8.1686	7.5108	**9.0943**
	($\lambda = 0.05$)	–	($\lambda = 5 \cdot 10^{-4}; \lambda_{TV} = 0.05$)	$\lambda_S = 0.01; \lambda_P = 0.005$	–
30	8.3773	*9.5078*	10.4845	11.3352	**11.9823**
	($\lambda = 0.01$)	–	($\lambda = 5 \cdot 10^{-4}; \lambda_{TV} = 0.01$)	$\lambda_S = 0.001; \lambda_P = 0.005$	–
40	10.6813	*11.9936*	20.8034	19.6017	**22.0538**
	($\lambda = 0.005$)	–	($\lambda = 5 \cdot 10^{-4}; \lambda_{TV} = 0.005$)	$\lambda_S = 0.005; \lambda_P = 0.005$	–

the observed mixed pixel $\mathbf{y}$ have the opportunity to become non-zero. The non-zeros grouping scheme can speed up the convergence of the proposed algorithms. The knee solution acts as a representative and guides each species to evolve toward a certain direction. Therefore MOSU can obtain a better PF and outperforms than MOSU without CC in the sparse unmixing problems.

B. Comparison of MOSU Against Other Methods

In this section, four data sets are used to test the performance of the proposed algorithms. In the following experiments, we compared the proposed algorithms, Bi-MOSU and Tri-MOSU, with other sparse unmixing methods: SUnSAL [3], SUnSAL-TV [4] and SUnSPI [52]. The signal-to-reconstruction error (SRE) is used to evaluate the accuracy assessment of the abundance estimations for all the experiments [1]. SRE gives more information about the relationship between the power of the error and the power of the signal, and it is defined as follows

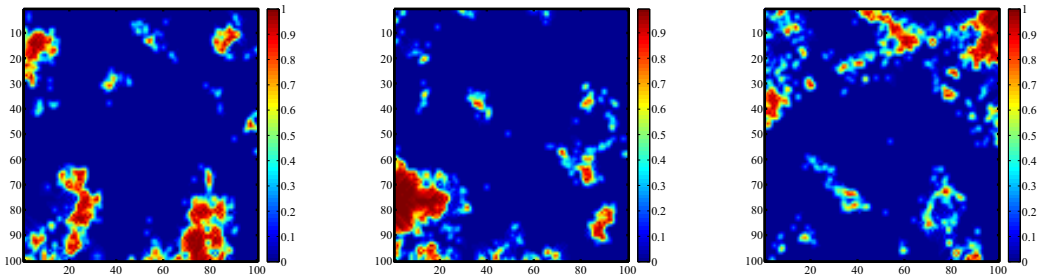
$$SRE(dB) \equiv 10 \log_{10}\left(\frac{E[\|\mathbf{x}\|_2^2]}{E[\|\mathbf{x} - \hat{\mathbf{x}}\|_2^2]}\right) \tag{21}$$

where $\hat{\mathbf{x}}$ stands for the estimated fractional abundance vector.

1) Experiments on Data 1: Data 1 is an image of size 75×75 and with 224 bands per pixel, and is provided by Bioucas-Dias [4]. As shown in Fig. 8(a), there are five endmembers in this image and both pure regions and mixed regions are distributed spatially in the form of distinct square regions. The true fractional abundance for each of the five endmembers is shown in Fig. 8(b) - Fig. 8(f), respectively. The background pixels are made up of a mixture of the same pixels with the fractional abundance values 0.1149, 0.0741, 0.2003, 0.2055, and 0.4051, respectively. The spectral library A_1 contains 240 spectral signatures that are selected from the USGS library [1]. The reflectance values of all the materials in the library are measured in 224 spectral bands and distributed uniformly in the interval of 0.4-2.5 μm.

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

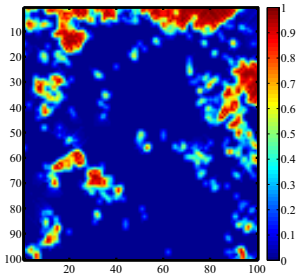
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TEVC.2016.2598858, IEEE Transactions on Evolutionary Computation



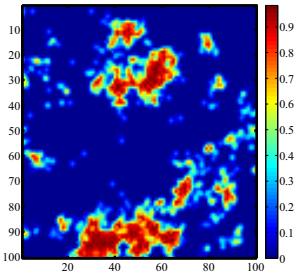
(a) endmember 1

(b) endmember 2

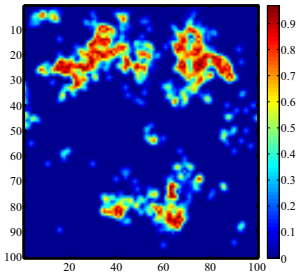
(c) endmember 3



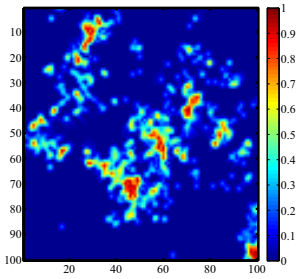
(d) endmember 4



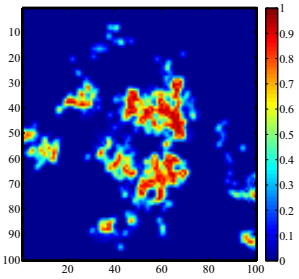
(e) endmember 5



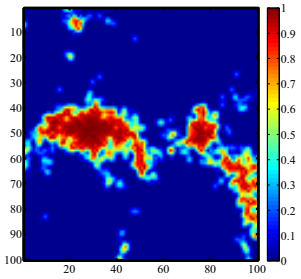
(f) endmember 6



(g) endmember 7



(h) endmember 8



(i) endmember 9

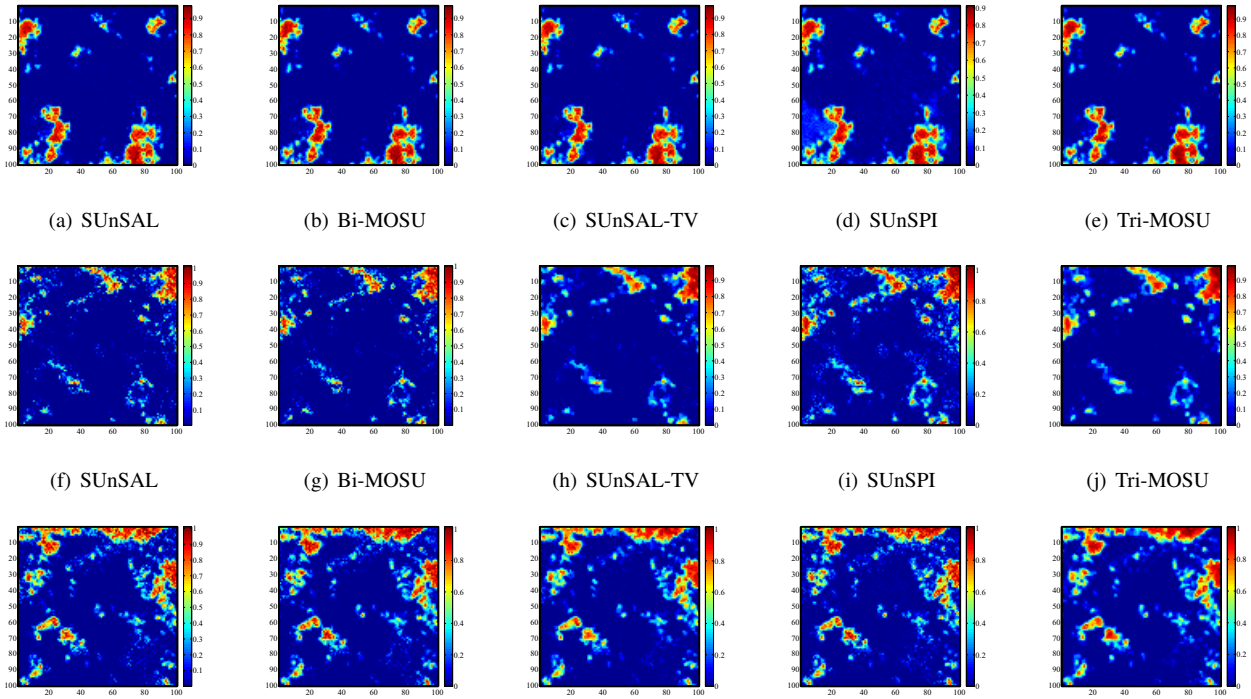
Fig. 10. True fractional abundances of data 2.

The unmixing results obtained by SUnSAL, Bi-MOSU, SUnSAL-TV, SUnSPI and Tri-MOSU are partly shown in Fig. 9, respectively. Table I shows the values of SRE obtained by different algorithms on the data 1 corrupted by different levels of white noise. Fig. 9 graphs the abundance maps estimated by different algorithms for three randomly selected endmembers on this data corrupted by 30dB white noise using spectral library A_1. Each row of images are the estimated abundance maps of one endmember obtained by different algorithms. From a visual analysis of Fig. 9, the estimated abundance maps obtained by SUnSAL and Bi-MOSU contain much noise points. It is hard to distinguish the endmember signatures from the mixed spectral signature without incorporating any spatial information between each pixel and its neighbors. The unmixing results obtained by SUnSAL-TV, SUnSPI and Tri-MOSU are close to the original image. The fractional abundance maps obtained by Tri-MOSU have less noise and are more smooth than those obtained by the other methods, but lose four patches in the abundance map of endmember 1. The fractional abundance maps of SUnSPI contain much more noise, but lose less patches in the abundance maps of endmember 1 and endmember 2. Although SUnSAL-TV can obtain smoother abundance maps than SUnSPI in data 1 corrupted by 30 dB white noise, the unmixing accuracy of SUnSAL-TV is lower than that of SUnSPI. The fractional abundance maps of Tri-MOSU almost lose no patches but are not smoother than those of SUnSAL-TV.

2) Experiments on Data 2: Data 2 contains 100×100 pixels and is provided by Iordache [4]. This data set exhibits a good spatial homogeneity as illustrated in Fig. 10, and the fractional abundances are piecewise smooth. And the true fractional abundance of nine endmembers signatures is shown in Fig. 10. The data set is also corrupted by different levels of correlated noise (SNR= 20, 30, 40 dB). The spectral library A_2 consists of 230 spectral signatures that are selected from the USGS library.

The unmixing results on data 2 are shown in Fig. 11. From Fig. 11, we can see that the algorithms which introduce spatial information obtain a better visual effect and contain fewer noise points. At first sight, the abundance maps of Bi-MOSU and SUnSAL may be the same, but it can be seen that the maps of Bi-MOSU contain less noise than that of SUnSAL in

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TEVC.2016.2598858, IEEE Transactions on Evolutionary Computation



(a) SUnSAL (b) Bi-MOSU (c) SUnSAL-TV (d) SUnSPI (e) Tri-MOSU

(f) SUnSAL (g) Bi-MOSU (h) SUnSAL-TV (i) SUnSPI (j) Tri-MOSU

(k) SUnSAL (l) Bi-MOSU (m) SUnSAL-TV (n) SUnSPI (o) Tri-MOSU

Fig. 11. Estimated abundance maps of endmember 1, 3, and 4 for data 2 obtained by different unmixing methods.

TABLE II
COMPARISONS OF THE PROPOSED ALGORITHMS AGAINST OTHER ALGORITHMS ON DATA 2.

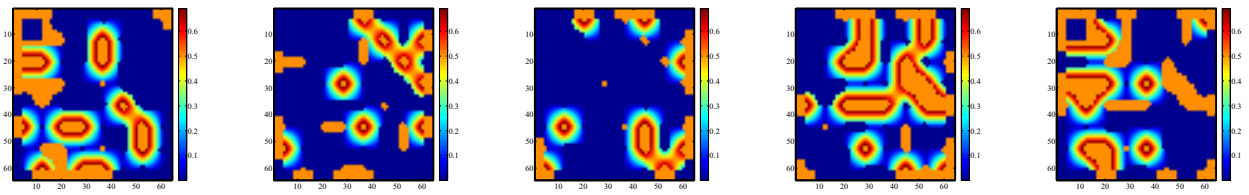
SNR (dB)	SUnSAL	Bi-MOSU	SUnSAL-TV	SUnSPI	Tri-MOSU
20	6.3007	*6.9007*	8.1628	8.3986	**8.8653**
	($\lambda = 0.01$)	–	($\lambda = 5 \cdot 10^{-4}; \lambda_{TV} = 0.005$)	($\lambda_S = 0.01; \lambda_P = 0.005$)	–
30	9.3115	*9.7115*	10.5989	9.9480	**11.1687**
	($\lambda = 0.005$)	–	($\lambda = 0.001; \lambda_{TV} = 5 \cdot 10^{-4}$)	($\lambda_S = 0.01; \lambda_P = 5 \cdot 10^{-4}$)	–
40	14.4291	*15.4291*	18.1507	16.5493	**20.0034**
	($\lambda = 5 \cdot 10^{-4}$)	–	($\lambda = 0; \lambda_{TV} = 5 \cdot 10^{-4}$)	($\lambda_S = 0.005; \lambda_P = 5 \cdot 10^{-4}$)	–

the rufous regions. The performances of SUnSAL-TV, SUnSPI and Tri-MOSU are nearly the same and their unmixing results are closer to the true abundance map. The abundance maps of SUnSAL-TV lose some small regions in the rufous regions, but are more clear than those of other methods. The maps of SUnSPI contain more noise than those of SUnSAL-TV and Tri-MOSU. The abundance maps of Tri-MOSU obtain smooth rufous regions but also have some small noise regions. The results of quantitative analysis for different algorithms are shown in Table II. From Table II, the proposed algorithms obtain satisfactory results. Both SUnSAL-TV and Tri-MOSU obtain a better spatial consistency due to the introduction of the TV regularization which exploits the spatial contextual information existing in the hyperspectral images. SUnSAL and SUnSAL-TV treat the sparse unmixing problem as a single optimization problem by imposing some parameters, which are sensitive to different data sets and have a great impact on the unmixing accuracy. In contrast, Bi-MOSU and Tri-MOSU optimize the l_0-norm problem directly and consider the sparse unmixing problem as multi-objective optimization problem.

3) Experiments on Data 3: Data 3, an image of size 64×64 pixels and with 224 bands per pixel, is provided by Prof. Z. Shi [52]. There is no pure pixel in this simulated hyperspectral data set. This data set is also homogeneous and the true fractional abundance for each of the five endmembers is shown in Fig. 12(a) - Fig. 12(e), respectively. Without loss of generality, the data is also corrupted by different levels of white noise (SNR= 20, 30 and 40 dB) in order to test the robustness of the proposed algorithms. The spectral library A_3 used in this subsection is also selected from the USGS digital spectral library.

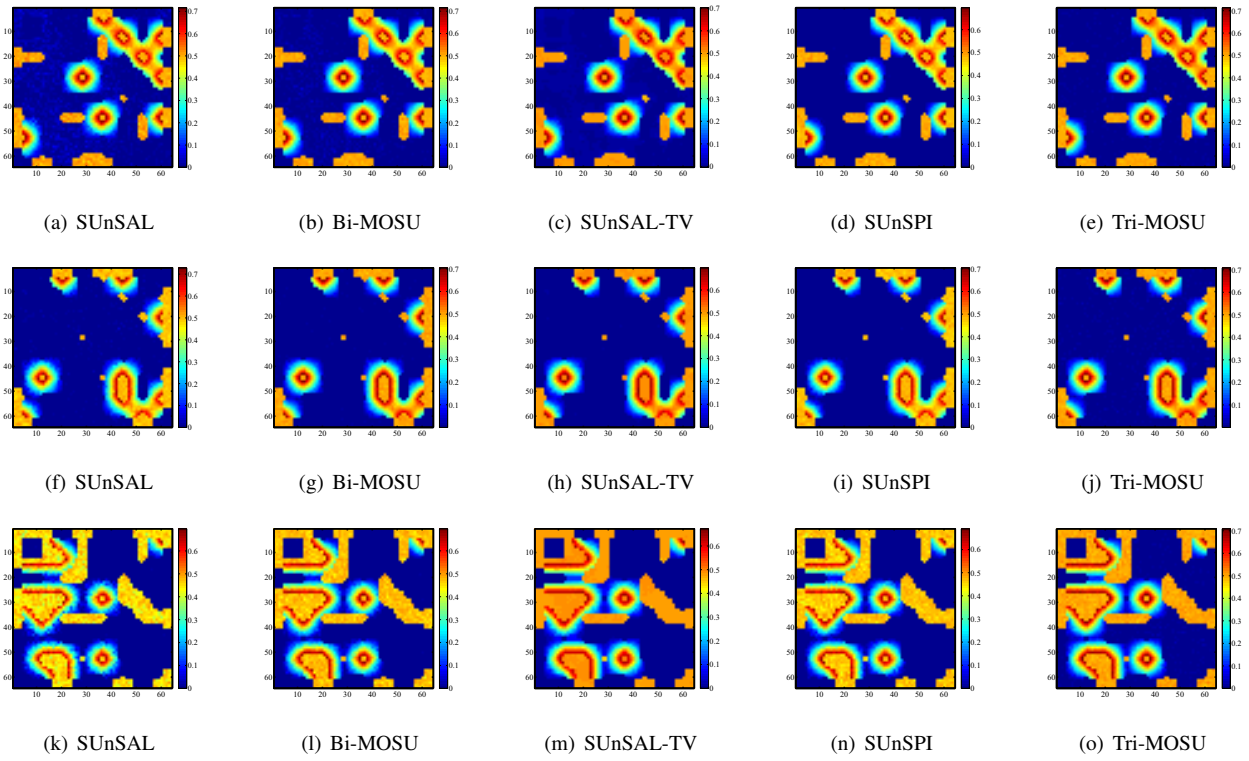
Fig. 13 shows the unmixing results of those algorithms on data 3. The first row of images are the estimated fractional abundance maps of endmember 2 obtained by different algorithms. The second and third row of images are the estimated fractional abundance maps of endmember 3 and of endmember 5, respectively. From Fig. 13, all the algorithms can obtain good edges. In the abundance map of endmember 5, it is obvious that the abundance map of Bi-MOSU contains less noise than that of SUnSAL. The maps of SUnSAL-TV almost

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TEVC.2016.2598858, IEEE Transactions on Evolutionary Computation



(a) endmember 1 (b) endmember 2 (c) endmember 3 (d) endmember 4 (e) endmember 5

Fig. 12. True fractional abundances of data 3.



(a) SUnSAL (b) Bi-MOSU (c) SUnSAL-TV (d) SUnSPI (e) Tri-MOSU

(f) SUnSAL (g) Bi-MOSU (h) SUnSAL-TV (i) SUnSPI (j) Tri-MOSU

(k) SUnSAL (l) Bi-MOSU (m) SUnSAL-TV (n) SUnSPI (o) Tri-MOSU

Fig. 13. Estimated abundance maps of endmember 2, 3, and 5 for data 3 obtained by different unmixing methods.

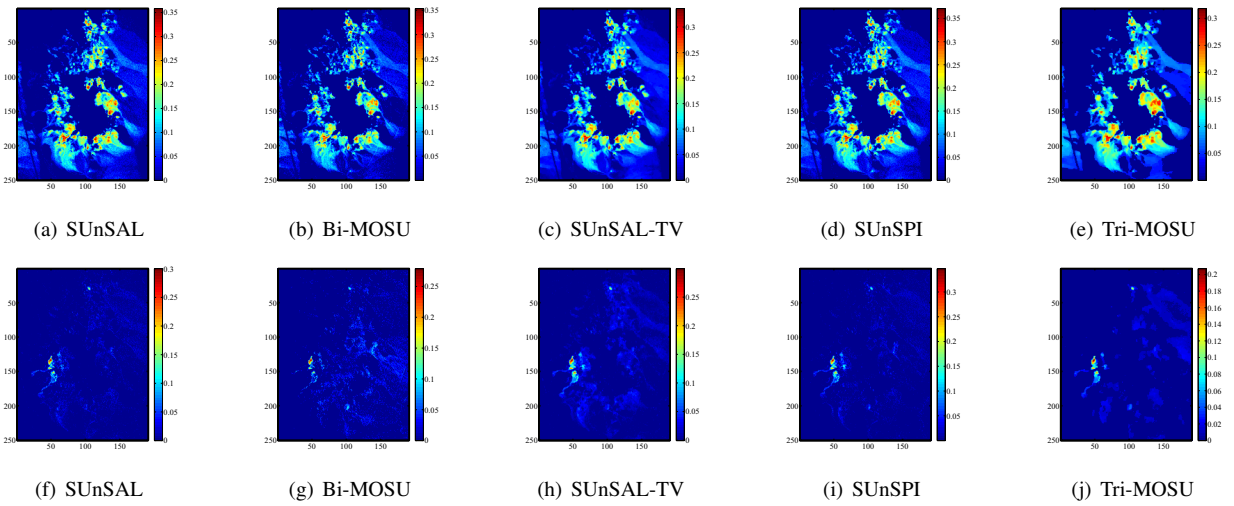
TABLE III
COMPARISONS OF THE PROPOSED ALGORITHMS AGAINST OTHER ALGORITHMS ON DATA 3.

SNR (dB)	SUnSAL	Bi-MOSU	SUnSAL-TV	SUnSPI	Tri-MOSU
20	7.4431	*8.3926*	12.1442	10.6149	**13.4202**
	($\lambda = 0.05$)	–	($\lambda = 5 \cdot 10^{-4}; \lambda_{TV} = 0.05$)	($\lambda_S = 5 \cdot 10^{-4}; \lambda_P = 0.005$)	–
30	15.2353	*17.8140*	22.3216	19.2026	**23.3075**
	($\lambda = 0.01$)	–	($\lambda = 5 \cdot 10^{-4}; \lambda_{TV} = 0.01$)	($\lambda_S = 5 \cdot 10^{-4}; \lambda_P = 0.001$)	–
40	23.1712	*23.8981*	30.1927	29.0447	**31.2155**
	($\lambda = 0.005$)	–	($\lambda = 5 \cdot 10^{-4}; \lambda_{TV} = 5 \cdot 10^{-4}$)	($\lambda_S = 5 \cdot 10^{-4}; \lambda_P = 5 \cdot 10^{-4}$)	–

contain no noise and also keep the good edges. The abundance maps of SUnSPI and Tri-MOSU still contain some noise more or less. At first sight, the color of the abundance maps of SUnSAL-TV are closer to the true abundance maps compared with SUnSPI and Tri-MOSU. On the contrary, the maximum scale of SUnSAL in the abundance maps of endmember 3 and endmember 5 is less than 0.7, and the maximum scale of SUnSPI and Tri-MOSU are lager than 0.7. The color of the maps of SUnSAL-TV is dark orange, but the value might be almost the same. The values of SRE shown in Table III show that Tri-MOSU achieves a higher accuracy than others.

4) Experiments on Data 4: The final data 4 used in the experiment is the AVIRIS Cuprite data set. This data set has been widely used to validate the performance of unmixing algorithms [4], [5], [53]. This data set comprises 224 spectral reflectance bands in the wavelength range 0.4-2.5 μm. Low SNR and water absorption bands (including bands 1-2, 105-115, 150-170, and 223-224) are removed. The materials maps obtained by he Tricorder software is shown in Fig. 14. We can use the mineral map to analyze the performance of dif-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TEVC.2016.2598858, IEEE Transactions on Evolutionary Computation



(a) SUnSAL (b) Bi-MOSU (c) SUnSAL-TV (d) SUnSPI (e) Tri-MOSU

(f) SUnSAL (g) Bi-MOSU (h) SUnSAL-TV (i) SUnSPI (j) Tri-MOSU

Fig. 15. Estimated abundance maps for the Cuprite data obtained by different unmixing methods.

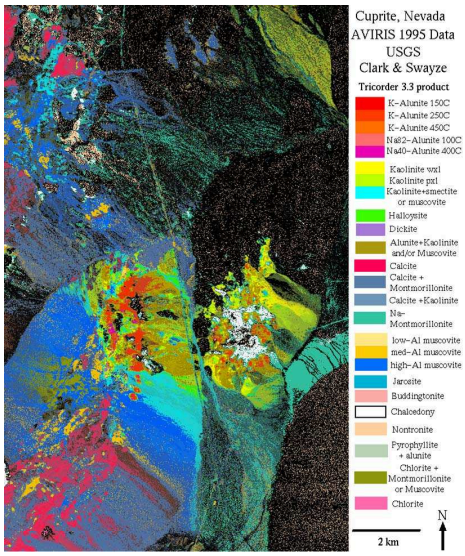


Fig. 14. Distribution of different minerals for Cuprite data obtained by Tricorder software.

ferent sparse unmixing algorithms. The subset of the AVIRIS Cuprite data set consists of 250×191 pixels and 188 spectral reflectance bands. We use the USGS digital spectral library as the spectral library (denoted by A). Calibration mismatches between the real image spectra and the spectra available in the library is undertaken due to the different acquisition conditions of the two data types. As a comparison, we use the minerals map (as shown in Fig. 14) produced by Tricorder software as a reference to analyze the performances of different algorithms.

A qualitative comparison among the fractional abundance maps of the two typical minerals: alunite and buddingtonite obtained by different algorithms to the subset of the AVIRIS Cuprite data set in this experiment are shown in Fig. 15. It is worth noting that the minerals map provided by Tricorder software and the abundance maps obtained by sparse unmixing algorithms are different. The minerals map provided by Tricorder software regards each pixel in the hyperspectral image as pure pixel and classify it as a certain mineral. However, unmixing regards a mixed pixel as a linear combination of the minerals with the fractional abundance. That is why the Tricorder map is different with the abundance maps generated by sparse unmixing algorithms. The highest abundance estimated by sparse unmixing algorithms is always consistent with the pixels classified as the corresponding materials in the Tricorder map. From Fig. 15, the abundance map of Bi-MOSU is similar to that of SUnSAL, but the brown regions that stand for large fractional abundance values distribute differently. The map of Bi-MOSU contains less noise points than that of SUnSAL according to the reference map produced by Tricorder software. The abundance maps of SUnSAL-TV, SUnSPI and Tri-MOSU are similar. But the difference between the brown regions with large fractional abundance values of those algorithms is closer to the maps provided by Tricorder software.

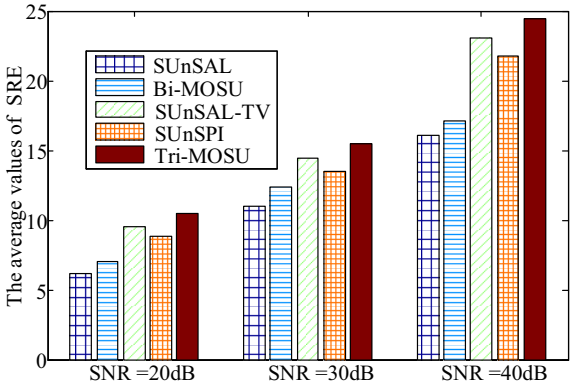


Fig. 16. The average values of SRE obtained by the five algorithms on the previous data sets with SNR = 20, 30 and 40dB.

5) Statistical Results: In order to further evaluate the performance of the proposed MOSU model, the statistical results obtained by the five algorithms are presented in this experiment. Fig. 16 shows the average values of SRE obtained by the five algorithms with SNR = 20, 30 and 40 dB.

Fig. 16 indicates that our approach can achieve the highest accuracies in the statistical sense when compared with the other algorithms.

V. Concluding Remarks

This paper has proposed a MOSU model to deal with the sparse unmixing for hyperspectral imagery. A multi-objective cooperative coevolutionary algorithm is used to solve this model. The reconstruction term, the sparsity term and the total variation regularization term are optimized by the proposed method simultaneously to gain more insights into the sparse unmixing problem. Specially, a problem-dependent cooperative coevolutionary strategy is designed to speed up the convergence and obtain high accuracy. Experimental results on simulated and real hyperspectral data sets has demonstrated the effectiveness of the proposed technique on sparse hyperspectral unmixing. For practical applications, the parallel processing will be considered in our future work for accelerating our algorithms. We will focus on designing some other regularizer for sparse unmixing to make full use of the spatial information. It is also feasible to enhance our cooperative coevolutionary strategy to improve the efficiency of the proposed method.

References

[1] M.-D. 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, 2011.

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

[3] J. M. Bioucas-Dias and M. A. Figueiredo, "Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing," in *Proc. 2nd Workshop Hyperspectr. Image Signal Process.: Evol. Remote Sens.*, vol. 1, 2010, pp. 1–4.

[4] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Total variation spatial regularization for sparse hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4484–4502, 2012.

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

[6] R. Feng, Y. Zhong, and L. Zhang, "Adaptive non-local euclidean medians sparse unmixing for hyperspectral imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 97, pp. 9–24, 2014.

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

[8] K. Deb, *Multi-objective optimization using evolutionary algorithms*. New York: Wiley, 2001.

[9] Y. Yu, X. Yao, and Z.-H. Zhou, "On the approximation ability of evolutionary optimization with application to minimum set cover," *Artif. Intell.*, vol. 180, pp. 20–33, 2012.

[10] C. Qian, Y. Yu, and Z.-H. Zhou, "On constrained boolean pareto optimization," in *Proc. 24th Int. Joint Conf. Artif. Intell.*, 2015, pp. 389–395.

[11] C. Qian, Y. Yu, and Z.-H. Zhou, "Pareto ensemble pruning." in *Proc. 29th AAAI Conf. Artif. Intell.*, 2015, pp. 2935–2941.

[12] C. Qian, Y. Yu, and Z.-H. Zhou, "Subset selection by pareto optimization," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 1765–1773.

[13] J. Branke, K. Deb, H. Dierolf, and M. Osswald, "Finding knees in multi-objective optimization," in *Proc. Parallel Problem Solving from Nature-PPSN VIII*, 2004, pp. 722–731.

[14] C. C. Borel and S. A. Gerstl, "Nonlinear spectral mixing models for vegetative and soil surfaces," *Remote Sens. Environ.*, vol. 47, no. 3, pp. 403–416, 1994.

[15] R. B. Singer and T. B. McCord, "Mars-large scale mixing of bright and dark surface materials and implications for analysis of spectral reflectance," in *Proc. 10th Lunar Planetary Sci. Conf.*, vol. 10, 1979, pp. 1835–1848.

[16] 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, 2001.

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

[18] C. A. C. Coello, D. A. Van Veldhuizen, and G. B. Lamont, *Evolutionary algorithms for solving multi-objective problems*. Norwell, MA: Kluwer, 2002.

[19] K. Miettinen, *Nonlinear multiobjective optimization*. Norwell, MA: Kluwer, 1999.

[20] C. M. Fonseca and P. J. Fleming, "An overview of evolutionary algorithms in multiobjective optimization," *Evol. Comput.*, vol. 3, no. 1, pp. 1–16, 1995.

[21] J. D. Schaffer, "Multiple objective optimization with vector evaluated genetic algorithms." in *Proc. 1st Int. Conf. Genetic Algorithms*, 1985, pp. 93–100.

[22] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, vol. 6, no. 2, pp. 182–197, 2002.

[23] C. A. C. Coello, G. T. Pulido, and M. S. Lechuga, "Handling multiple objectives with particle swarm optimization," *IEEE Trans. Evol. Comput.*, vol. 8, no. 3, pp. 256–279, 2004.

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

[25] O. Soliman, L. T. Bui, and H. Abbass, "A memetic coevolutionary multi-objective differential evolution algorithm," *Multi-Objective Memetic Algorithms*, pp. 369–388, 2009.

[26] K. Deb and H. Jain, "An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, Part I: Solving problems with box constraints," *IEEE Trans. Evol. Comput.*, vol. 18, no. 4, pp. 577–601, 2014.

[27] Z. Zhu, J. Xiao, J. Li, F. Wang, and Q. Zhang, "Global path planning of wheeled robots using multi-objective memetic algorithms," *Integr. Comput.-Aided Eng.*, vol. 22, no. 4, pp. 387–404, 2015.

[28] L. Wang, Q. Zhang, A. Zhou, M. Gong, and L. Jiao, "Constrained subproblems in a decomposition-based multiobjective evolutionary algorithm," *IEEE Trans. Evol. Comput.*, vol. 20, no. 3, pp. 475–480, 2016.

[29] X. Qiu, J. X. Xu, K. C. Tan, and H. A. Abbass, "Adaptive cross-generation differential evolution operators for multiobjective optimization," *IEEE Trans. Evol. Comput.*, vol. 20, no. 2, pp. 232–244, 2016.

[30] X. Ma, F. Liu, Y. Qi, X. Wang, L. Li, L. Jiao, M. Yin, and M. Gong, "A multiobjective evolutionary algorithm based on decision variable analyses for multiobjective optimization problems with large-scale variables," *IEEE Trans. Evol. Comput.*, vol. 20, no. 2, pp. 275–298, 2016.

[31] Z. Zhu, J. Xiao, S. He, Z. Ji, and Y. Sun, "A multi-objective memetic algorithm based on locality-sensitive hashing for one-to-many-to-one dynamic pickup-and-delivery problem," *Inf. Sci.*, vol. 329, pp. 73–89, 2016.

[32] P. S. Oliveto, J. He, and X. Yao, "Analysis of the (1+1)-EA for finding approximate solutions to vertex cover problems," *IEEE Trans. Evol. Comput.*, vol. 13, no. 5, pp. 1006–1029, 2009.

[33] Y. Yu, X. Yao, and Z.-H. Zhou, "On the approximation ability of evolutionary optimization with application to minimum set cover," *Artif. Intell.*, vol. 180–181, pp. 20–33, 2012.

[34] X. Lai, Y. Zhou, J. He, and J. Zhang, "Performance analysis of evolutionary algorithms for the minimum label spanning tree problem," *IEEE Trans. Evol. Comput.*, vol. 18, no. 6, pp. 860–872, 2014.

[35] C. Qian, Y. Yu, and Z.-H. Zhou, "An analysis on recombination in multi-objective evolutionary optimization," *Artif. Intell.*, vol. 204, pp. 99–119, 2013.

[36] Y. Liu, X. Yao, Q. Zhao, and T. Higuchi, "Scaling up fast evolutionary programming with cooperative coevolution," in *Proc. IEEE Congr. Evol. Comput.*, vol. 2, 2001, pp. 1101–1108.

[37] Z. Yang, K. Tang, and X. Yao, "Large scale evolutionary optimization using cooperative coevolution," *Inf. Sci.*, vol. 178, no. 15, pp. 2985–2999, 2008.

[38] X. Li and X. Yao, "Cooperatively coevolving particle swarms for large scale optimization," *IEEE Trans. Evol. Comput.*, vol. 16, no. 2, pp. 210–224, 2012.

[39] Y. Mei, X. Li, and X. Yao, "Cooperative coevolution with route distance grouping for large-scale capacitated arc routing problems," *IEEE Trans. Evol. Comput.*, vol. 18, no. 3, pp. 435–449, 2014.

[40] M. N. Omidvar, X. Li, Y. Mei, and X. Yao, "Cooperative co-evolution with differential grouping for large scale optimization," *IEEE Trans. Evol. Comput.*, vol. 18, no. 3, pp. 378–393, 2014.

[41] S. Chand and R. Chandra, "Multi-objective cooperative coevolution of neural networks for time series prediction," in *Proc. Int. Joint Conf. Neural Netw.*, 2014, pp. 190–197.

[42] M. A. Rad and A. Hamzeh, "A coevolutionary approach to many objective optimization based on a novel ranking method," *Intelligent Data Analysis*, vol. 20, no. 1, pp. 129–151, 2016.

[43] S. He, Z. Zhu, G. Jia, and D. Tennant, "Cooperative co-evolutionary module identification with application to cancer disease module discovery," *IEEE Trans. Evol. Comput.*, 2016, doi: 10.1109/TEVC.2016.2530311.

[44] 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, 2004.

[45] M. A. Potter and K. A. De Jong, "A cooperative coevolutionary approach to function optimization," in *Proc. 3rd Conf. Parallel Problem Solving From Nature*, 1994, pp. 249–257.

[46] A. Neubauer, "A theoretical analysis of the non-uniform mutation operator for the modified genetic algorithm," in *Proc. IEEE Congr. Evol. Comput.*, 1997, pp. 93–96.

[47] Z. Michalewicz, *Genetic algorithms+ data structures= evolution programs*. Springer-Verlag, 1996.

[48] K. C. Tan, Y. Yang, and C. K. Goh, "A distributed cooperative coevolutionary algorithm for multiobjective optimization," *IEEE Trans. Evol. Comput.*, vol. 10, no. 5, pp. 527–549, 2006.

[49] C.-K. Goh and K. C. Tan, "A competitive-cooperative coevolutionary paradigm for dynamic multiobjective optimization," *IEEE Trans. Evol. Comput.*, vol. 13, no. 1, pp. 103–127, 2009.

[50] L. M. Antonio and C. Coello Coello, "Use of cooperative coevolution for solving large scale multiobjective optimization problems," in *Proc. IEEE Congr. Evol. Comput.*, 2013, pp. 2758–2765.

[51] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: A comparative case study and the strength pareto approach," *IEEE Trans. Evol. Comput.*, vol. 3, no. 4, pp. 257–271, 1999.

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

[53] J. M. 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, 2005.

Maoguo Gong (M'07-SM'14) received the B.S. degree in electronic engineering (first class honors) and the Ph.D. degree in electronic science and technology from Xidian University, Xi'an, China, in 2003 and 2009, respectively.

Since 2006, he has been a Teacher with Xidian University. In 2008 and 2010, he was promoted as an Associate Professor and as a Full Professor, respectively, both with exceptive admission. His research interests are in the area of computational intelligence with applications to optimization, learning, data mining and image understanding.

Dr. Gong received the prestigious National Program for the support of Top-Notch Young Professionals from the Central Organization Department of China, the Excellent Young Scientist Foundation from the National Natural Science Foundation of China, and the New Century Excellent Talent in University from the Ministry of Education of China. He is the Vice Chair of the IEEE Computational Intelligence Society Task Force on Memetic Computing, an Executive Committee Member of the Chinese Association for Artificial Intelligence, and a Senior Member of the Chinese Computer Federation. Please see his homepage (http://see.xidian.edu.cn/faculty/mggong) for more information.

Hao Li received the B.S. degree in electronic engineering from Xidian University, Xi'an, China, in 2013. He is currently pursuing the Ph.D. degree in pattern recognition and intelligent systems at the School of Electronic Engineering, Xidian University, Xi'an, China.

His research interests include computational intelligence and image understanding.

Enhu Luo received the B.S. degree in electronic engineering from Xidian University, Xi'an, China, in 2013. He is currently pursuing the M.S. degree in pattern recognition and intelligent systems at the School of Electronic Engineering, Xidian University, Xi'an, China.

His research interests include computational intelligence and image understanding.

Jing Liu (M'05-SM'15) received the B.S. degree in computer science and technology and the Ph.D. degree in circuits and systems from Xidian University in 2000 and 2004, respectively. In 2005, she joined Xidian University as a lecturer, and was promoted to a full professor in 2009. From Apr. 2007 to Apr. 2008, she worked at The University of Queensland, Australia as a postdoctoral research fellow, and from Jul. 2009 to Jul. 2011, she worked at The University of New South Wales at the Australian Defence Force Academy as a research associate.

Now, she is a full professor in the Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, Xidian University. Her research interests include evolutionary computation, complex networks, fuzzy cognitive maps, multiagent systems, and data mining. She is the associate editor of *IEEE Trans. Evolutionary Computation*. Please see her homepage (http://see.xidian.edu.cn/faculty/liujing) for more information.

Jia Liu received the B.S. degree in electronic engineering from Xidian University, Xi'an, China, in 2013. He is currently pursuing the Ph.D. degree in pattern recognition and intelligent systems at the School of Electronic Engineering, Xidian University, Xi'an, China.

His research interests include computational intelligence and image understanding.