Evolutionary Multitasking Sparse Reconstruction: Framework and Case Study

Hao Li, Yew-Soon Ong, *Fellow, IEEE*, Maoguo Gong, *Senior Member, IEEE*, and Zhenkun Wang

***Abstract*—Real-world applications typically have multiple sparse reconstruction tasks to be optimized. In order to exploit the similar sparsity pattern between different tasks, this paper establishes an evolutionary multitasking framework to simultaneously optimize multiple sparse reconstruction tasks using a single population. In the proposed method, the evolutionary algorithm aims to search the locations of nonzero components or rows instead of searching sparse vector or matrix directly. Then the within-task and between-task genetic transfer operators are employed to reinforce the exchange of genetic material belonging to the same or different tasks. The proposed method can solve multiple measurement vector problems efficiently because the length of decision vector is independent of the number of measurement vectors. Finally, a case study on hyperspectral image unmixing is investigated in an evolutionary multitasking setting. It is natural to consider a sparse unmixing problem in a homogeneous region as a task. Experiments on signal reconstruction and hyperspectral image unmixing demonstrate the effectiveness of the proposed multitasking framework for sparse reconstruction.**

***Index Terms*—Evolutionary algorithm, hyperspectral unmixing, multitasking optimization, sparse reconstruction.**

I. Introduction

SPARSE reconstruction aims to find sparse solutions to large under-determined linear systems of equations. It has been widely used in many applications, such as signal processing, pattern recognition, computer vision, and so on [1]–[3]. Let $\boldsymbol{x} \in \mathbb{R}^N$ be a sparse vector that contains many zero elements, its reconstruction normally takes the following linear equation:

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} \tag{1}$$

Manuscript received May 17, 2018; revised September 18, 2018 and November 9, 2018; accepted November 14, 2018. Date of publication November 19, 2018; date of current version October 1, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant 61772393, in part by the National Key Research and Development Program of China under Grant 2017YFB0802200, in part by the Key Research and Development Program of Shaanxi Province under Grant 2018ZDXM-GY-045, in part by the Data Science and Artificial Intelligence Center, in part by the School of Computer Science and Engineering at Nanyang Technological University, and in part by the National Research Foundation Singapore through its AI Singapore Programme under Award AISG-RP-2018-00. *(Corresponding author: Maoguo Gong.)*

H. Li and M. Gong are with the School of Electronic Engineering, Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, Xidian University, Xi'an 710071, China (e-mail: omegalihao@gmail.com; gong@ieee.org).

Y.-S. Ong and Z. Wang are with the School of Computer Science and Engineering, Nanyang Technological University, Singapore 639798 (e-mail: asysong@ntu.edu.sg; wangzhenkun90@gmail.com).

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

Digital Object Identifier 10.1109/TEVC.2018.2881955

where $\boldsymbol{A} \in \mathbb{R}^{M\times N}$ with $M \leq N$ is known as the sensing matrix or over-complete dictionary and $\boldsymbol{y} \in \mathbb{R}^M$ is called the measurement vector. If noise is considered, sparse reconstruction is to recover the sparse vector from under-determined linear systems of equations

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{n} \tag{2}$$

where $\boldsymbol{n} \in \mathbb{R}^M$ is the additive independent identically distributed noise. In the noise-free or noisy cases, sparse reconstruction can be formulated as the following constrained optimization problems:

$$\min_{\boldsymbol{x}} \ \|\boldsymbol{x}\|_0, \quad \text{s.t.} \quad \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} \tag{3}$$

$$\min_{\boldsymbol{x}} \ \|\boldsymbol{x}\|_0, \quad \text{s.t.} \quad \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 \leq \sigma \tag{4}$$

where $\|\cdot\|_0$ is the l_0 norm that counts the number of nonzero values of a vector. With the penalty method, the above two problems can be converted into the following form:

$$\min_{\boldsymbol{x}} \ \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \lambda\|\boldsymbol{x}\|_0 \tag{5}$$

where λ is a positive regularization parameter.

The sparse reconstruction problem is highly nonconvex and has been proven to be an NP-hard optimization problem [4]. Many schemes have been developed to deal with the sparse reconstruction problem. Among them, greedy algorithms and relaxation methods are two widely used techniques for sparse reconstruction. Greedy algorithms (e.g., the orthogonal matching pursuit and its variants [5]) select the column most correlated with the current residuals so as to add the corresponding entry into the set of nonzero entries in each iteration. Greedy algorithms have good performance when the nonzero entries are identified correctly. However, if noise exists, the performance of greedy algorithms may degenerate significantly since some zero components may be considered as the nonzero ones. Relaxation methods use the $l_p(0 < p \leq 1)$ norm minimization instead of the l_0 norm so that they are less sensitive to noise compared with greedy algorithms [6]. Nevertheless, the quality of the signal reconstructed with the relaxation method is easily affected by the selection of Lagrangian parameter [7], [8].

In this paper, the sparse reconstruction in case of multiple measurement vectors (MMVs) is also considered. Its model can be generalized from the basic single measurement vector (SMV) model [i.e., (2)] [9]–[11]. Given L structurally equivalent SMV models, the MMV model can be expressed as

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{N} \tag{6}$$

where $\boldsymbol{Y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_L]$ and $\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_L]$. Similar to the sparse constraint in the SMV model, the number of nonzero rows in $\boldsymbol{X}$ should be less than a threshold to ensure sparsity. The MMV model has been demonstrated to be more suitable for some applications [9], [10], [12]–[16]. For example, a theoretical result of hyperspectral unmixing shows that the probability of recovery failure decays exponentially in the number of measurement vectors [15]. Iordache *et al.* [15] proposed a collaborative hyperspectral unmixing method based on the MMV model. The sparsity is simultaneously imposed to all pixels in the data set. Then Iordache *et al.* [16] proposed a two-step algorithm to exploit the usual low dimensionality of the hyperspectral data sets. The algorithm identifies a subset of the library elements and then applies collaborative hyperspectral unmixing.

Recently, multiobjective evolutionary algorithms (MOEAs) have been employed to solve the sparse reconstruction problems [7], [8], [17]–[19]. In order to balance the measurement error and the sparsity constraint, Li *et al.* [17] modeled sparse reconstruction as a multiobjective optimization problem involving measurement error and a sparsity-inducing term. In [7], a local search strategy based on a soft thresholding method was embedded into the NSGA-II algorithm [20] to increase the speed of convergence. Luo *et al.* [18] proposed a multiobjective sparse spectral clustering method and designed a ratio cut-based method to select a tradeoff solution from the Pareto front (PF). Zhou *et al.* [8] proposed a two-phase evolutionary approach for sparse reconstruction. In the first phase, MOEA/D [21] is applied to generate a set of robust solutions by optimizing l_1 norm of the solutions. In the second phase, a forward-based selection method is proposed to further update the solution set to make the identification of nonzero entries more precise. In [19], a preference-based multiobjective approach was proposed to solve sparse optimization problems. The randomness of the sparsity level can be viewed as a mutation operator. Notwithstanding, most of these MOEA-based methods are proposed to deal with SMV problems. For MMV problems, the solution representation is a matrix, not a vector. It is a difficult task to design a suitable representation strategy and genetic operators for considering the structure of MMVs. Furthermore, it suffers from the curse of dimensionality with the increase of the length of sparse vector and the number of measurement vectors.

In many real-world applications, there often exist more than one sparse reconstruction tasks that have to be solved at the same time and notably these tasks typically share similar sparsity pattern [22]. Taking the cue, in this paper, we extend the evolutionary multitasking framework proposed in [23]–[25] for simultaneous optimization of multiple sparse reconstruction tasks by taking advantage of the similar sparsity pattern found across the tasks. Evolutionary multitasking optimization was first proposed by Gupta *et al.* [23] and then extended to a multiobjective version in [24]. In evolutionary multitasking optimization, each individual is assigned to a skill factor for indicating the cultural trait of the associated task. Then the individuals are encoded in a unified search space and the genetic operators are applied to produce offspring in this space. The offspring also inherits the parents' skill factors through the vertical cultural transmission. In the phase of evaluation, these individuals are decoded into the solution representation of the associated task.

In summary, a novel multitasking sparse reconstruction (MTSR) framework is proposed in this paper to solve both SMV and MMV problems in the evolutionary multitasking setting. In the proposed multitasking framework, each task is modeled as a multiobjective optimization problem. Instead of coding the sparse vector or matrix directly, evolutionary algorithm in the proposed method aims to search the locations of the nonzero components of the sparse vector in SMV problems or the nonzero rows of the sparse matrix in MMV problems. Then the within-task and between-task genetic transfer occurred in the candidate parents is used to generate the next population. Eventually, a set of nondominated solutions is attained for each task. A case study on hyperspectral sparse unmixing is also investigated to validate the effectiveness of the proposed method. In this case study, a sparse unmixing problem in a homogeneous region is considered as a task. The pixels in the homogeneous region are more likely to share the same active set of endmembers than that of the pixels in the whole image.

The contributions of this paper are threefold.

1) This paper solves the sparse reconstruction problems in an evolutionary multitasking setting to facilitate improved convergence characteristics. The overall convergence characteristics can be significantly improved by the between-task transfer in the multitasking framework.
2) While existing MOEA-based algorithms are designed for SMV problems, this paper is the first one to consider MMV problems.
3) A case study on hyperspectral sparse unmixing is investigated in the evolutionary multitasking setting.

This paper is organized as follows. Section II gives the background knowledge on multiobjective optimization and multitasking optimization. Our motivation of using evolutionary multitasking optimization is also discussed. In Section III, we describe the proposed MTSR framework in detail. Section IV presents a case study on hyperspectral image unmixing. The experimental results on signal reconstruction and hyperspectral unmixing are provided in Section V. Finally, the main conclusions are drawn in Section VI.

II. Background and Motivation

In this section, we introduce the basic concepts of multiobjective optimization and multitasking optimization. In particular, we distinguish the differences between the two paradigms. Moreover, we present the motivation of using multitasking optimization for solving sparse reconstruction problems.

A. Background

1) Evolutionary Multiobjective Optimization: Generally, a multiobjective optimization problem (MOP) with m decision variables and n objectives can be expressed as

$$\begin{aligned} \min \ & F(\boldsymbol{x}) = (f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \ldots, f_n(\boldsymbol{x}))^{\mathsf{T}} \\ \text{s.t.} \ & \boldsymbol{x} = (x_1, x_2, \ldots, x_m)^{\mathsf{T}} \in \Omega \end{aligned} \tag{7}$$

where Ω is the *feasible set*, and $\boldsymbol{x}$ is the *decision variable vector* in it. $F : \Omega \to \mathbb{R}^n$ consists of n real-valued objective functions. Normally, 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 [26]. The Pareto set is the set of all Pareto-optimal solutions and its mapping to objective space is the PF. MOEAs aim to find a set of approximate solutions to the PF [20], [21], [26], [27].

2) Evolutionary Multitasking Optimization: The paradigm of multitask learning has been investigated over two decades [28]. The idea is to employ relevant information available in different tasks by performing learning using a shared representation. Evolutionary multitasking optimization aims to reveal the multitasking potential of evolutionary algorithms [29], [30]. A multitasking optimization problem with K tasks can be mathematically expressed as

$$\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_K\} = \arg\min \ \{F_1(\boldsymbol{x}), F_2(\boldsymbol{x}), \ldots, F_K(\boldsymbol{x})\} \quad \text{s.t.} \quad \boldsymbol{x}_i \in \Omega_i, i = 1, 2, \ldots, K. \tag{8}$$

It is easy to distinguish the differences between multiobjective optimization and multitasking optimization. For multiobjective optimization, there exists a design space for a given task and all variables are contained in this single space. Evolutionary multitasking optimization has multiple heterogeneous design spaces and every individual is encoded into a unified space. Furthermore, each task in the multitasking setting can be one multiobjective optimization problem. In [24], a multifactorial optimization technique was proposed to solve the multitasking optimization problems. Then Gupta *et al.* [24] 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 ith individual p_i with the jth task T_j, the factorial rank r_j^i, skill factor τ_i and scalar fitness ϕ_i are defined as follows.

1) r_j^i*:* The factorial rank r_j^i is the index of p_i in the list of population members sorted in ascending order with respect to T_j.
2) τ_i*:* The skill factor τ_i indicates the associated task. If p_i is evaluated for all tasks, we can have $\tau_i = \arg\min\{r_j^i\}$, where $j \in \{1, 2, \ldots, K\}$.
3) ϕ_i*:* The scalar fitness of p_i in a multitasking environment is given by $\phi_i = 1/r_{\tau_i}^i$.

In the multiobjective case, nondominated sorting and crowding distance are employed to order population members [20], [24]. 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.

B. Motivation

As described in Section I, some MOEA-based sparse reconstruction methods have been proposed to deal with SMV problems. These schemes always consider the measurement error and sparsity-inducing terms as two objective functions and then search the sparse vector directly. Whereas, it is difficult to extend them to solve MMV problems since multiple sparse vectors exist in the MMV model. Therefore, this paper aims to propose an efficient MOEA-based sparse reconstruction method for solving both SMV and MMV problems.

In the real-world applications, multiple sparse reconstruction tasks may have to be solved at the same time and share similar sparsity pattern. These sparse reconstruction problems can be solved individually by existing sparse reconstruction algorithms. However, these methods may ignore the relevant information available in related tasks. In order to take the advantage of the similar sparsity pattern found across the tasks, this paper investigates the MOEA-based sparse reconstruction method in an evolutionary multitasking setting to facilitate improved convergence characteristics.

In this paper, we propose an MOEA-based sparse reconstruction method for MMV problems and investigate it in an evolutionary multitasking setting. There are two major differences between the previous MOEA-based works and our proposed method. On the one hand, the proposed method can solve both SMV and MMV problems while most of the previous schemes are designed for SMV problems. On the other hand, the proposed method solves multiple sparse reconstruction tasks simultaneously to leverage upon the underlying commonalities between different tasks.

III. Evolutionary Multitasking Sparse Reconstruction

In this section, an evolutionary multitasking framework for sparse reconstruction is presented. First, we show the multiobjective model for SMV problems and then extend it to solve MMV problems. An evolutionary multitasking framework based on the multiobjective model is described in detail. Next, we elaborate the strategies of the representation and the initialization. Then we show the within-task and between-task genetic transfer occurred in the candidate parents. The evaluation of the individuals is also given. Finally, the computational cost of the proposed method is analyzed.

A. Evolutionary Multitasking Framework for Sparse Reconstruction

As described in Section I, it is a difficult task to set the regularization parameter or estimate the true sparsity in sparse reconstruction. In order to detect the true sparsity automatically, sparse reconstruction can be modeled as a multiobjective optimization problem [7], [18], [19]. Some studies have shown that the best sparse solution is located in the knee region of PF [7], [31]. In the MOEA-based sparse reconstruction methods, the measurement error and the sparsity terms are simultaneously optimized to obtain a set of tradeoff solutions. The multiobjective sparse reconstruction model for the SMV problem can be formulated as

$$\min F(\boldsymbol{x}) = \min\left(\|\boldsymbol{x}\|_0, \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2^2\right). \tag{9}$$

Compared with the regularization problem (5), no regularization parameter is needed in the multiobjective model. This paper aims to extend the above multiobjective model for

Algorithm 1: Evolutionary MTSR

Step 1 Initialization
Generate N_p individuals to form the initial population P_0.
for ***every*** p_i *in* P_0 **do**
  Assign the skill factor τ_i.
  Evaluate p_i based on **Algorithm 3**.
end
Compute scalar fitness ϕ_i for every p_i based on the nondominated sorting and crowding distance [20], [24].
Set the generation number $g = 0$.
Step 2 Cycling
Obtain P'_g by the tournament selection method according to P_g.
Acquire offspring C_g based on the within-task and between-task genetic transfer shown in **Algorithm 2**.
for ***every*** c_i *in* C_g **do**
  Evaluate $\boldsymbol{c}_i$ based on **Algorithm 3**.
end
$R_g = P_g \bigcup C_g$.
Update scalar fitness of all individuals in R_g.
Select N_p individuals from R_g to form P_{g+1}.
Step 3 Stopping criteria
If the stopping conditions are not satisfied, then $g = g + 1$ and go to **Step 2**, otherwise, stop the algorithm and output P_g.

solving MMV problems and investigates it in the evolutionary multitasking setting. Inspired by the multiobjective model in (9), we arrive at the following multiobjective model for MMV problems:

$$\min F(\boldsymbol{x}) = \min\left(J(\boldsymbol{X}),\ \|\boldsymbol{AX} - \boldsymbol{Y}\|_F^2\right) \tag{10}$$

where $J(\boldsymbol{X})$ represents the row-l_0 quasi-norm [32], [33].

For the MMV model, the target signals are supported on a common support set and we have

$$\text{supp}(\boldsymbol{X}) = \bigcup_{i=1}^{L} \text{supp}(\boldsymbol{x}_i). \tag{11}$$

Then the row-l_0 quasi-norm of a matrix can be defined as

$$J(\boldsymbol{X}) = |\text{supp}(\boldsymbol{X})|. \tag{12}$$

Similar to the SMV model, many existing methods employ $\|\boldsymbol{X}\|_{0,q}$ $(q \geq 1)$ to measure the number of rows in $\boldsymbol{X}$. The l_q norm is applied to rows of $\|\boldsymbol{X}\|$ and then the l_0 norm is applied to the resulting vector. Similar to the use of the l_p-norm $(0 < p \leq 1)$ in the SMV model, $\|\boldsymbol{X}\|_{p,q}$ can be used to replace $\|\boldsymbol{X}\|_{0,q}$ in the relaxation methods [33].

It has been demonstrated that the transfer of genetic materials among different tasks can speed up the convergence for various optimization problems [23]–[25], [30], [34]. In fact, in many real-world applications, there exist multiple sparse optimization problems that typically share similar sparsity pattern [22]. In order to take full advantage of the parallelism of population-based search, this paper aims to solve multiple sparse reconstruction problems simultaneously using a single population. The algorithm of the proposed evolutionary multitasking framework for sparse reconstruction is shown in Algorithm 1. In the proposed method, we aim to search the locations of the nonzero components/rows instead of searching the sparse vector/matrix directly. Every individual is encoded into a unified space and is assigned to a skill factor to indicate the associated task. During the evolution of the population, the genetic materials can be transferred among the candidate parents belonging to the same or different tasks. The skill factor of the offspring is obtained from one of its parents. In the evaluation phase, we obtain the corresponding sparse vector/matrix based on the acquired locations by solving a simple optimization problem and then evaluate each individual for the associated task. Finally, we arrive at a set of nondominated solutions for each task.

B. Representation and Initialization

In existing EA-based algorithms, the sparse vector $\boldsymbol{x} \in \mathbb{R}^N$ in the SMV model is considered as the decision vector. It is difficult to modify these methods to deal with MMV problems because the solution $\boldsymbol{X} \in \mathbb{R}^{N \times L}$ is a matrix. When either or both of N and L become large, EA-based sparse reconstruction approaches face the curse of dimensionality. Instead of coding the sparse vector $\boldsymbol{x}$ or sparse matrix $\boldsymbol{X}$ directly, we search the locations of the nonzero components in $\boldsymbol{x}$ or nonzero rows of $\boldsymbol{X}$ and then calculate the values of these nonzeros entries by solving a simple optimization problem. In this paper, a binary variable $\boldsymbol{v}$ is used to indicate the nonzero components/rows.

In the multitasking environment with K optimization tasks, it is assumed that the dimensionality of the kth task is D_k. Because $\boldsymbol{v}$ only takes "0" or "1," it is an easy task to construct a unified search space with the maximum dimensionality of the K tasks, which can be expressed as

$$D_{\max} = \max_k \{D_k\},\ k = 1, 2, \ldots, K. \tag{13}$$

In the initialization, every individual has a vector of $D_{\max}$ random variables. For kth task, D_k variables are extracted from the $D_{\max}$ random variables based on some prior domain knowledge [24]. However, in cases where prior domain knowledge is unavailable, a simple default is to use the first D_k variables of the $D_{\max}$ random variables [23], [24].

C. Within-Task and Between-Task Genetic Transfer

In an evolutionary multitasking setting, the transfer of knowledge may occur in the candidate parents belonging to the same or different tasks. The pseudocode of the within-task and between-task genetic transfer is depicted in Algorithm 2, which aims to generate two offspring c_1 and c_2 from two candidate parents p_1 and p_2. It is assumed that the skill factors of p_1 and p_2 are τ_1 and τ_2, respectively. When the two parents are most effective for the same task, i.e., τ_1 is equal to τ_2, the offspring inherit the same skill factor from their parents. When the skill factors of the two parents are different, a random mating probability (rmp) is defined to determine whether the two parents belonging two different tasks conduct genetic transfer or not. Only the mutation operator is employed to generate

Algorithm 2: Within-Task and Between-Task Genetic Transfer

Input: Two candidate parents p_1 and p_2 in P'_g, *rmp*.
Output: Two offspring c_1 and c_2.

```
if τ1 == τ2 then
    %Within-task genetic transfer
    Obtain c1 and c2 by the crossover and mutation operators.
    The skill factors of c1 and c2 are equal to τ1(τ2).
else
    %Between-task genetic transfer
    if rand < rmp then
        Obtain c1 and c2 by the crossover and mutation operators.
        for i = 1, 2 do
            if rand < 0.5 then
                The skill factors of ci is τ1.
            else
                The skill factors of ci is τ2.
            end
        end
    else
        for i = 1, 2 do
            Obtain ci by the mutation operator on pi.
            The skill factors of ci is τi.
        end
    end
end
```

the offspring when the random number is larger than rmp. Otherwise, both crossover and mutation operators are utilized to acquire the offspring and each offspring randomly inherits the skill factor of one of their parents. Note that the two offspring may share the same skill factor. The parameter rmp is used to balance exploitation and exploration of search space. When the value of rmp is close to 0, only culturally alike individuals are allowed to crossover. Therefore, there is always the tendency for solutions to get trapped in local optima. 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. Here, we consider uniform crossover [35] and bit-wise mutation [36] as the mutation and crossover operators.

The between-task genetic transfer shown in Algorithm 2 is able to transfer genetic material between individuals belonging to different tasks. On the one hand, the rmp allows individuals with different skill factors to crossover and the offspring can randomly inherit the skill factor from one of their parents. The transfer turns out to be beneficial if the genetic material from one task happens to be useful for the another task. On the other hand, inferior genes generated due to negative transfer are removed from the population by means of "survival of the fittest." In the special case of one sparse reconstruction task, only within-task transfer is considered and the proposed MTSR framework degenerates to a standard multiobjective method.

Algorithm 3: Obtain the Nonzero Values and Evaluation

Input: The binary vector $\boldsymbol{v}_i$ in the ith individual.
Output: The sparse signal $\boldsymbol{x}_i$ or $\boldsymbol{X}_i$.

```
S = supp(v).
if It is an SMV problem then
    Obtain xi by solving the problem (14) based on the index set S.
else
    Obtain Xi by solving the problem (15) based on the index set S.
end
Evaluate xi or Xi for task τi only.
```

D. Evaluation

As described in Section III-B, a binary vector $\boldsymbol{v}$ is employed to represent the nonzero locations of the sparse vector or matrix. After obtaining $\boldsymbol{v}$, in the case of SMV, we can solve the following optimization problem to acquire the nonzero values of $\boldsymbol{x}$:

$$\boldsymbol{x}_S = \arg\min\{\|\boldsymbol{y} - \boldsymbol{A}_S\boldsymbol{x}_S\|_2\} \tag{14}$$

where $S = \text{supp}(\boldsymbol{v})$. $\text{supp}(\boldsymbol{v})$ is index set of nonzero entries of $\boldsymbol{v}$. Therefore, $\boldsymbol{A}_S$ is the submatrix of $\boldsymbol{A}$ containing the columns specified in the set S. $\boldsymbol{x}_S$ is the subvector of $\boldsymbol{x}$ containing the rows specified in the set S. Furthermore, the MMV problems can be easily extended from (14) to the following:

$$\boldsymbol{X}_S = \arg\min\{\|\boldsymbol{Y} - \boldsymbol{A}_S\boldsymbol{X}_S\|_F\} \tag{15}$$

where $\boldsymbol{X}_S$ is the submatrix of $\boldsymbol{X}$ containing the rows specified in the set S. The problems (14) and (15) can be solved by the least square method [8], [37], [38]. The solutions to the two problems in the noiseless case are expressed as $\boldsymbol{x}_S = (\boldsymbol{A}_S^T\boldsymbol{A}_S)^{-1}\boldsymbol{A}_S^T\boldsymbol{y}$ and $\boldsymbol{X}_S = (\boldsymbol{A}_S^T\boldsymbol{A}_S)^{-1}\boldsymbol{A}_S^T\boldsymbol{Y}$, respectively.

As described in Section III-C, the offspring inherits the genetic material from their parents and is culturally influenced by them via vertical cultural transmission. In the proposed method, the offspring randomly inherits the skill factor of one of their parents and we evaluate each individual only for one task based on the assigned skill factor. This selective evaluation strategy can reduce the computational cost with the increasing number of tasks. After evaluation of the offspring, the next population is selected based on the scalar fitness ϕ, which is calculated according to the skill factor τ and the factorial rank r. As described above, the skill factor of the offspring is obtained from one of its parents. In evolutionary multitasking optimization, the nondominated sorting and crowding distance are employed to order population for acquiring the factorial rank [20], [24]. It is assumed that the nondominated fronts of the individuals p_1 and p_2 are NF_1 and NF_2, respectively. The crowding distances of p_1 and p_2 are CD_1 and CD_2, respectively. The individual p_2 is preferred over p_1 if any one of the following conditions hold.

1) $\text{NF}_2 < \text{NF}_1$.
2) $\text{NF}_2 = \text{NF}_1$ and $\text{CD}_2 > \text{CD}_1$.

E. Analysis on Computational Cost

Function evaluation is a common factor for defining the computation cost. It is known to be computationally expensive to evaluate every individual for every task. Therefore, the selective evaluation strategy proposed in [23] is used to evaluate an individual for only selected tasks. As indicated in [23], for K tasks, function evaluations are reduced almost by a factor of K in comparison with the case that an individual is evaluated for all tasks.

Existing MOEA-based sparse reconstruction methods, including StEMO [7] and MOEA/D [8], [17], have considered the sparse vector as the decision variables of interest. The number of decision variables represents one of the major factors contributing to the complexity and challenges of sparse reconstruction problems [39]. For MMV problems, the number of decision variables explodes with the increase of L. In our proposed method, evolutionary algorithm is used to search the locations of nonzero entries or rows, which are encoded into a binary vector. Therefore, the number of decision variables remains unchanged despite the increase in L. Since the decision variables are binary in nature, our proposed method is computationally more efficient for solving sparse optimization problems when compared to the representation used in StEMO and MOEA/D algorithms.

IV. Case Study on Hyperspectral Unmixing

In this section, we provide a case study on hyperspectral unmixing problems. We begin by briefly introducing the hyperspectral unmixing model and establish the multitasking unmixing model.

A. Hyperspectral Unmixing Model

A hyperspectral camera acquires the data for a large number of spectral bands. It is obvious that hyperspectral images provide much more detailed information about the scene than a color image. A set of images are collected by the hyperspectral sensors and then formed a 3-D hyperspectral data cube for processing and analysis. However, each pixel of a hyperspectral image may be a mixture of several materials. Spectral unmixing aims to separate the mixed pixels into a collection of spectra (endmembers) and estimate their fractional abundances. The traditional spectral unmixing consists of an endmember extraction step and an abundance estimation step [40]. To date, numerous endmember extraction schemes have been proposed to find the pure signatures, such as PPI [41] and N-FINDER [42]. Several other algorithms have been proposed based on the assumption that the input data has no pure signatures [43]. However, such methods do not perform well in highly mixed scenarios [44], [45].

In recent years, several libraries of natural and man-made materials are available for public use [44]. Compared with the number of materials in the spectral library, the number of endmembers existing in a hyperspectral pixel is small and the fractional abundances are sparse. Therefore, hyperspectral unmixing problem can be transformed into a sparse reconstruction model. Iordache *et al.* [44] provided a comparison of several sparse reconstruction algorithms and proposed a constrained sparse unmixing method via variable splitting and augmented Lagrangian (SUnSAL). Overall, a variety of algorithms have been proposed to deal with the sparse unmixing problems [45].

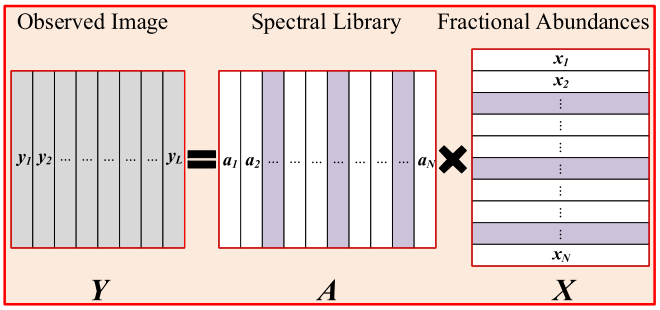


Fig. 1. Graphical illustration of the MMV-based sparse unmixing model. The nonzero rows of the matrix of fractional abundances are shown in purple color.

Let $\boldsymbol{y}$ denote an $M \times 1$ column vector representing an M dimensional vector of a hyperspectral image with M spectral bands. $\boldsymbol{A} = [\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_N]$ is a spectral library with N spectral signatures and each of them has M spectral bands. $\boldsymbol{x}$ is an $N \times 1$ vector representing the estimated fractional abundances. Therefore, the pixel observation $\boldsymbol{y}$ can be expressed as

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{n}. \tag{16}$$

Many algorithms considered the sparse unmixing problem as an SMV problem and obtained the fractional abundances in a pixel-wise manner [44], [45]. In general, the pixels in the hyperspectral image may share a common set of endmembers. If the factional abundances are reformulated as a matrix, there should be only a few nonzero rows. Assuming that the data contains L pixels organized in the matrix $\boldsymbol{Y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \ldots, \boldsymbol{y}_L]$, we have

$$\boldsymbol{Y} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{N} \tag{17}$$

where $\boldsymbol{X} = [\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_L]$ is the abundance fraction matrix. It is obvious that hyperspectral unmixing can be established as an MMV-based sparse reconstruction model. Fig. 1 shows the graphical illustration of the MMV-based sparse unmixing model. Iordache *et al.* [15] improved their previous work and presented a collaborative SUnSAL algorithm, which showcases better performance compared to the pixel-wise independent approach.

B. Hyperspectral Unmixing in Evolutionary Multitasking

The collaborative sparse unmixing model employed in [15] and [16] assumes that all pixels in the hyperspectral images share the same active set of endmembers [15], [46], [47]. In fact, the active endmembers might be different between pixels. Compared to the pixels in the entire image, the pixels belonging to the same homogeneous region are more likely to share the same active set of endmembers [46], [47]. This paper aims to find a small set of endmembers to unmix all the pixels in a homogeneous region. Assuming that we have T homogeneous

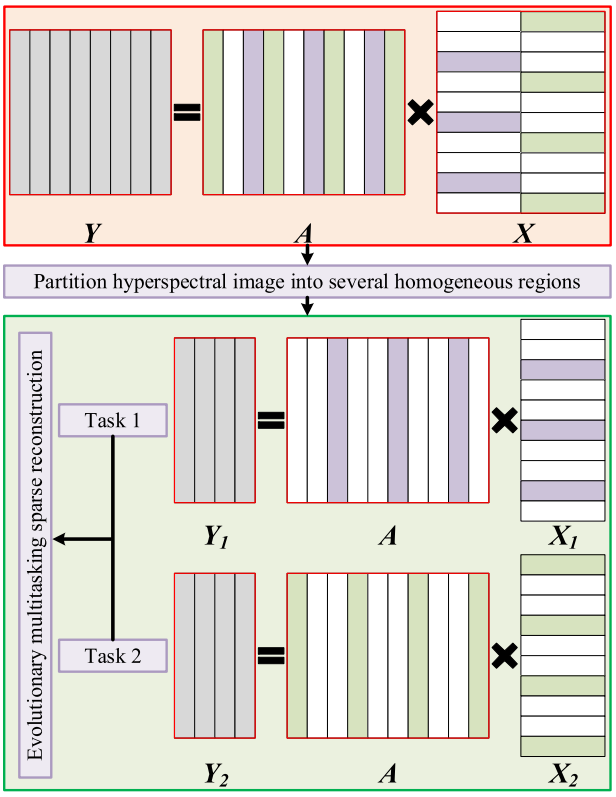


Fig. 2. Evolutionary MTSR for solving hyperspectral unmixing problems.

regions, $\boldsymbol{Y}$ can be represented as

$$\boldsymbol{Y} = \bigcup_{i=1}^{T} \boldsymbol{Y}_i \tag{18}$$

where $\boldsymbol{Y}_i$ is the pixels in the ith homogeneous region. For each region, we have

$$\boldsymbol{Y}_i = \boldsymbol{A}\boldsymbol{X}_i + \boldsymbol{N}_i \tag{19}$$

where $\boldsymbol{X}_i$ is the fractional abundances for the pixels in ith homogeneous region.

It is a natural idea to employ the proposed framework of evolutionary MTSR to solve the above sparse unmixing problems. Fig. 2 illustrates the proposed evolutionary MTSR algorithm for solving hyperspectral unmixing problems. In the proposed method, we partition the hyperspectral image into several homogeneous regions. Various clustering or classification methods for hyperspectral images have been proposed and can be employed in this step [48]. For each homogeneous region, the pixels are very likely to share a same active set of spectral signatures. Therefore, solving the sparse unmixing problem in a homogeneous region can be considered as a task. As shown in Fig. 2, in the original model, some nonzero rows involving all pixels have zero elements because different endmembers might have null abundance in different pixels. The sparsity measured by (12) may not be accurate to consider all pixels in the whole image. The proposed method aims at unmixing all the pixels in a homogeneous region. The sparsity-inducing term in (12) is able to approximately measure the true sparsity of the sparse matrix involving the pixels in a homogeneous region. Finally, the proposed evolutionary MTSR method can be used to solve these tasks simultaneously.

In this paper, hyperspectral unmixing is considered as a case study to investigate MTSR. In general, the potential application in other tasks can also be investigated in the proposed evolutionary multitasking framework. For example, in objective detection and tracking, many sparse-representation-based trackers have been proposed for tracking [49]–[51]. Multitask joint sparse representation has been employed for feature-level fusion [49]. It may exist multiple sparse optimization tasks in objective detection and tracking. These tasks can be optimized simultaneously by the proposed evolutionary multitasking framework. It is promising to improve the overall performance of related tasks by exploiting relevant information available in different tasks.

V. Experimental Study

In the proposed MTSR method, the population size is set to 100. The parameter rmp serves to reflect the possibility of fruitful genetic transfer between tasks. This parameter has been widely investigated in the previous works [23], [24]. As recommended in [24], the parameter rmp is set to 1 to reinforce the exchange of genetic material between tasks. In this section, several experiments are conducted on signal reconstruction and hyperspectral unmixing problems. For signal reconstruction, six problems are generated with different configurations and each of them has five tasks. In the experiments on hyperspectral unmixing, a simulated data set is generated based on the available spectral library. Next, two benchmark problems are used to validate the effectiveness of the proposed multitasking framework.

In order to compare the performance of the sparse reconstruction algorithms, two evaluation indicators are considered in the experiments.

1) *Signal to Reconstruction Error (SRE):* This index can be used to measure the quality of the reconstruction of a signal, which is expressed in dB [44]

$$\text{SRE(dB)} = 10\log_{10}\left(\frac{E\left[\|\boldsymbol{x}_t\|_2^2\right]}{E\left[\|\boldsymbol{x}-\boldsymbol{x}_t\|_2^2\right]}\right) \tag{20}$$

where $\boldsymbol{x}_t$ is the true sparse vector and $\boldsymbol{x}$ stands for the estimated sparse vector. The higher SRE represents the better quality of the reconstruction.

2) *Success Ratio (SR):* If the relative error is smaller than a given threshold τ, the corresponding run of this method is denoted as a successful run [19], [44]. SR under the threshold τ is defined as

$$\text{SR}_\tau = P\left(\frac{\|\boldsymbol{x}-\boldsymbol{x}_t\|_F^2}{\|\boldsymbol{x}_t\|_F^2} \le \tau\right). \tag{21}$$

The probability is the ratio of the successful runs on 100 random instances. If we set $\tau = 10$ and arrive at $\text{SR}_\tau = 1$, this implies that the relative error of the reconstruction result is less than 10 with probability one.

TABLE I
SIX SIMULATED PROBLEMS WITH DIFFERENT SETTINGS (M, N, L, Q, σ) IN OUR EXPERIMENTS

ID	Task	Configuration	ID	Task	Configuration
P1	1	(80,200,1,20,0.05)	P4	1	(120,200,2,20,0.05)
	2	(100,200,1,20,0.05)		2	(120,200,4,20,0.05)
	3	(120,200,1,20,0.05)		3	(120,200,6,20,0.05)
	4	(140,200,1,20,0.05)		4	(120,200,8,20,0.05)
	5	(160,200,1,20,0.05)		5	(120,200,10,20,0.05)
P2	1	(120,200,1,20,0.05)	P5	1	(120,200,4,20,0.05)
	2	(120,200,1,25,0.05)		2	(120,200,4,25,0.05)
	3	(120,200,1,30,0.05)		3	(120,200,4,30,0.05)
	4	(120,200,1,35,0.05)		4	(120,200,4,35,0.05)
	5	(120,200,1,40,0.05)		5	(120,200,4,40,0.05)
P3	1	(80,200,4,20,0.05)	P6	1	(120,200,4,20,0.02)
	2	(100,200,4,20,0.05)		2	(120,200,4,20,0.04)
	3	(120,200,4,20,0.05)		3	(120,200,4,20,0.06)
	4	(140,200,4,20,0.05)		4	(120,200,4,20,0.08)
	5	(160,200,4,20,0.05)		5	(120,200,4,20,0.10)

A. Experiments on Signal Reconstruction

In this experiment, we artificially constructed six sets of test problems and each of them has five optimization tasks. These optimization tasks are associated with the configuration (M, N, L, Q, σ), where Q is the number of nonzero entries or rows and σ is the standard deviation of noise. The six simulated problems with different settings are shown in Table I. P1 and P2 are SMV problems and the others are MMV problems. The settings of M are different in the tasks of P1 and P3 and the others are unchanged. The tasks in P2 and P5 have different sparsity level Q. The tasks in P4 have the same settings of M, N, Q, and σ and different values of L. The noise level σ is different in the tasks of P6.

In the experiments, StEMO proposed in [7] and MOEA/D in [17] are considered as the state-of-the-art benchmarks for comparison. In StEMO, a soft-thresholding method was incorporated into the NSGA-II algorithm. In [17], MOEA based on decomposition [21] was used as the baseline algorithm and iterative hard thresholding (IHT) was considered as the local search. The population size in the two algorithms is configured to 100 for the sake of fair comparison. However, the two methods are only designed for SMV problems. Some existing algorithms for SMV problems have also been extended to deal with MMV problems. In the experiments, IHT [52], focal under-determined system solver for MMV model (MFOCUSS) [9], and compressive sampling matching pursuit [38] represent some of these algorithm benchmarks that we consider here.

Fig. 3 shows the results obtained by the proposed method on the five optimization tasks in P5. Note that these tasks possess different sparsity level Q. The red curves in the left figures represent the PF curves acquired by MTSR. The solutions in the knee region have the maximum marginal rates of return [7], i.e., the improvement in one objective brings about a rapid degradation in other objectives. It is easy to distinguish the optimal knee points on the five PFs. In this paper, we use the angle-based method employed in [7] to acquire the knee solution. Because the PF curve is not as smooth as it looks, some points may distort the calculations for finding the knee region. As shown in the blue curves of the left figures, we interpolate the PF using B-splines to arrive at a smooth curve. Then we perform resampling from the smooth curves, which are shown in the dotted curves in the right figures. The curves of the corresponding angle are also shown in the right figures. The dashed line indicates the true sparsity. It can be observed that the point with the largest angle is near to the dashed line. The solution with maximum angle is approximatively considered as the best tradeoff solution. Fig. 4 exploits the box-plots to show the statistical results of the values of f_1, which counts the number of nonzero rows. As shown in Table I, P5 has five MMV tasks with different number of nonzero rows, i.e., $Q =$ 15, 20, 25, 30, and 35. Obviously, for each task, the value of

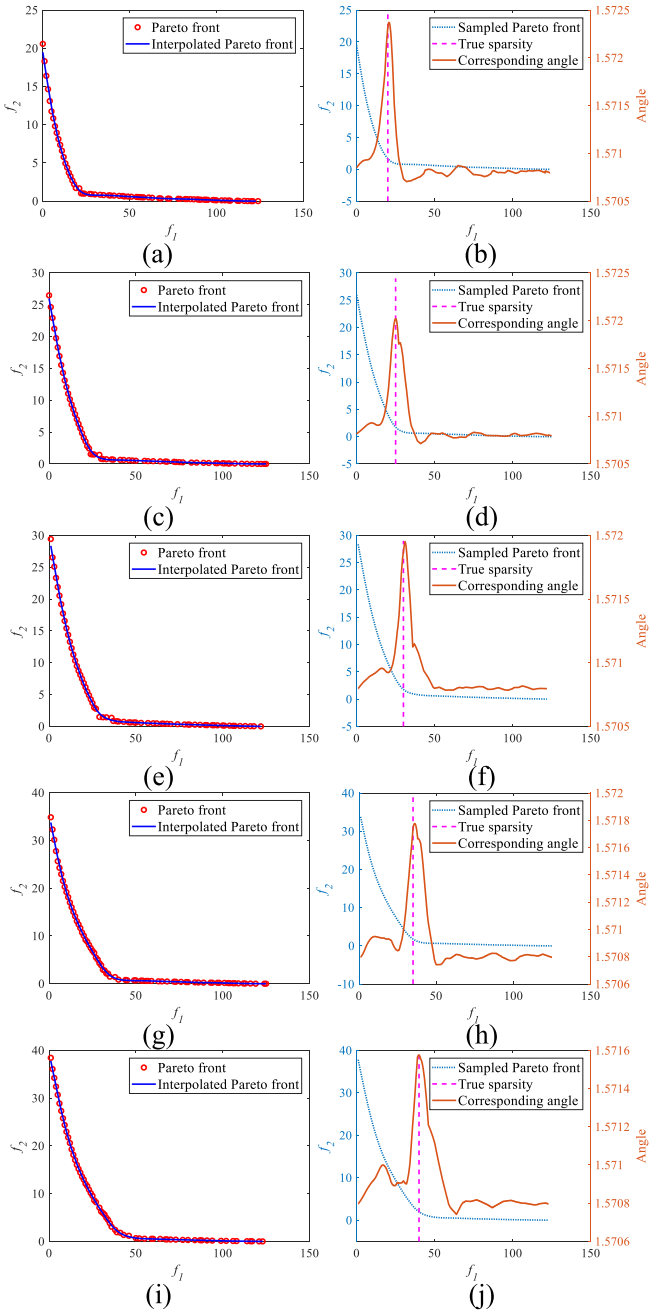


Fig. 3. Results obtained by the proposed method on the optimization tasks in P5. Left: PF. Right: angle of each point on PF.

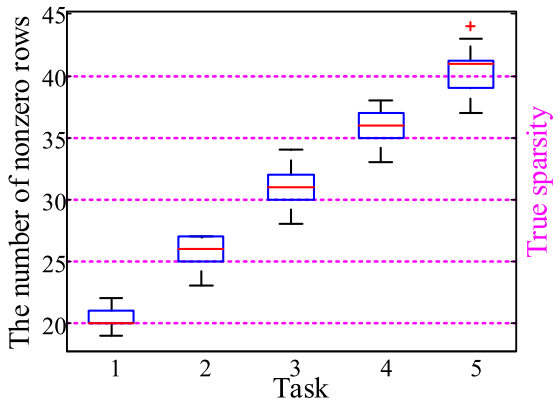


Fig. 4. Box-plots of the number of nonzero rows for five tasks.

TABLE II
MEAN AND STANDARD DEVIATION OF THE HYPERVOLUME OBTAINED BY THE FOUR METHODS ON SIX PROBLEMS. THE RESULTS ARE OBTAINED OVER 50 RUNS

Problem	Statistics	StEMO	MOEA/D	MTSR_0	MTSR
P1	Mean	0.9521	0.9365	0.9634	0.9870
	Std	0.0044	0.0031	0.0048	0.0039
P2	Mean	0.9655	0.9486	0.9752	0.9932
	Std	0.0043	0.0021	0.0035	0.0019
P3	Mean	–	–	0.9611	0.9885
	Std	–	–	0.0045	0.0032
P4	Mean	–	–	0.9719	0.9924
	Std	–	–	0.0044	0.0031
P5	Mean	–	–	0.9644	0.9877
	Std	–	–	0.0039	0.0035
P6	Mean	–	–	0.9716	0.9861
	Std	–	–	0.0039	0.0033

f_1 of the knee solution is very close to the setting of Q in a statistical sense.

As described in Section III, the proposed MTSR method degenerates to a standard multiobjective method when only one task is considered, i.e., only the within-task genetic transfer works. Therefore, MTSR without between-task genetic transfer is denoted as MTSR_0. Hypervolume is used to measure the quality of the PF obtained by StEMO, MOEA/D, MTSR_0, and MTSR. Hypervolume is calculated using a reference point 1% larger in every component than the corresponding nadir point [53], [54]. The nadir point is constructed from the worst objective function values of Pareto optimal solutions. Table II reports the mean and standard deviation of hypervolume on the six problems. The results of P3–P6 obtained by StEMO and MOEA/D are unavailable since these two methods were not designed to cope with MMV problems. The results of StEMO are noted to be inferior to the other methods. It can also be observed that the hypervolume attained by the proposed method are larger than 0.98 across all the problems considered.

Fig. 5 exhibits the successful ratios of different sparse reconstruction methods in 100 runs at different τ on P1–P6. These results indicate that the proposed MTSR method performed best among the algorithms. Table III tabulates the results of SRE(dB) and SR_τ for the different sparse reconstruction

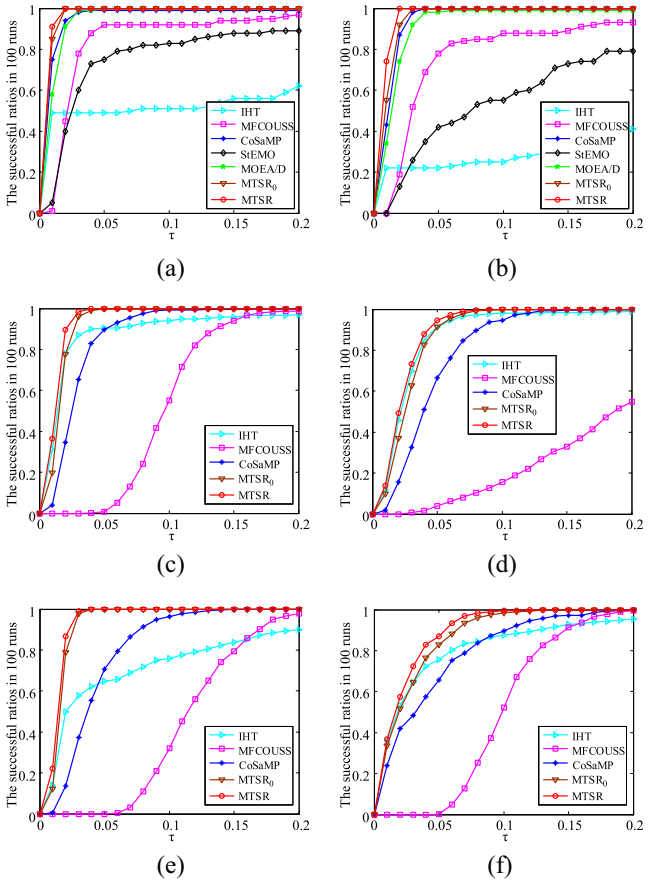


Fig. 5. Successful ratios of different sparse reconstruction methods at different τ on P1–P6. (a) P1. (b) P2. (c) P3. (d) P4. (e) P5. (f) P6.

schemes. The results of SRE(dB) are averaged of 100 runs. Notably, the proposed MTSR method emerges as superior among the others considered.

B. Experiments on Hyperspectral Unmixing

In this section, several experiments on hyperspectral unmixing are conducted to validate the effectiveness of the proposed multitasking framework. The proposed multitasking method for hyperspectral unmixing is also denoted as MTSR. In the experiments, we report the results obtained by the proposed method and pit it against the classical NCLS [55], SUnSAL [44], CLSUnSAL [15], and RCSU [47] algorithms. For all algorithms, the parameters were carefully tuned for optimal performance. In the experiments, the regularization parameter in SUnSAL, CLSUnSAL, and RCSU is tuned: $\lambda \in \{0.01, 9\cdot10^{-3}, 8\cdot10^{-3}, 7\cdot10^{-3}, 6\cdot10^{-3}, 5\cdot10^{-3}, 4\cdot10^{-3}, 3\cdot10^{-3}, 2\cdot10^{-3}, 1\cdot10^{-3}, 9\cdot10^{-4}, 7\cdot10^{-4}, 5\cdot10^{-4}, 3\cdot10^{-4}, 1\cdot10^{-4}\}$. In total, 15 sets of experiments are conducted for each algorithm to select the parameter λ.

1) Data Sets: The spectral library $\boldsymbol{A}$ used in the experiments is obtained from the USGS library.[1] It contains spectral signatures with reflectance values measured in 224 spectral bands and distributed uniformly in the interval 0.4–2.5 μm. The first data set is denoted as DS1. Based on the known

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

TABLE III
PERFORMANCE COMPARISON WITH THE DIFFERENT SPARSE RECONSTRUCTION METHODS. THE VALUES OF τ FOR SMV AND MMV PROBLEMS ARE SET TO 0.05 AND 0.03, RESPECTIVELY

Method	SRE						SR_τ					
	P1	P2	P3	P4	P5	P6	P1	P2	P3	P4	P5	P6
IHT	15.1056	9.7704	17.8530	17.2061	14.4143	17.6643	0.49	0.22	0.87	0.70	0.58	0.65
MFCOUSS	16.0796	14.2995	10.0994	8.8733	9.2851	9.9607	0.92	0.78	0.00	0.01	0.00	0.00
CoSaMP	21.1771	19.6606	15.9171	15.2547	14.1768	16.1283	0.99	1.00	0.65	0.33	0.37	0.41
StEMO	15.8970	13.4621	–	–	–	–	0.75	0.42	–	–	–	–
MOEA/D	20.5958	18.6326	–	–	–	–	1.00	0.98	–	–	–	–
$MTSR_0$	21.7231	19.1341	18.3476	17.3815	18.0980	18.1252	1.00	1.00	0.96	0.65	0.98	0.65
MTSR	23.5014	19.8186	19.2343	18.2096	18.7563	18.9971	1.00	1.00	0.98	0.73	0.99	0.73

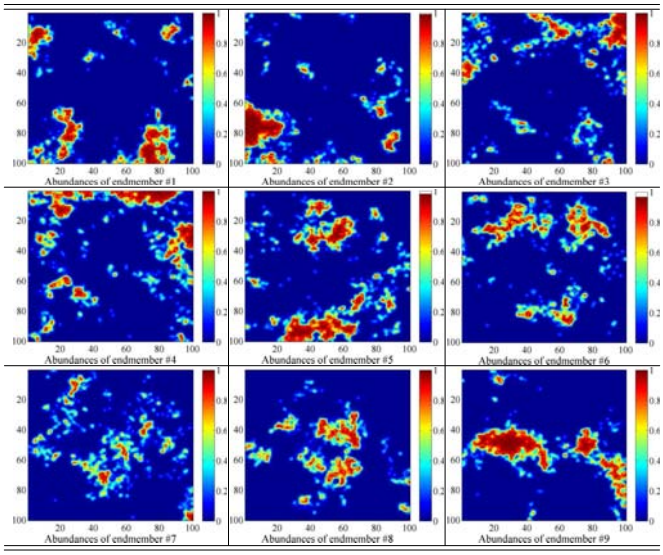


Fig. 6. True fractional abundances of endmembers in DS2.

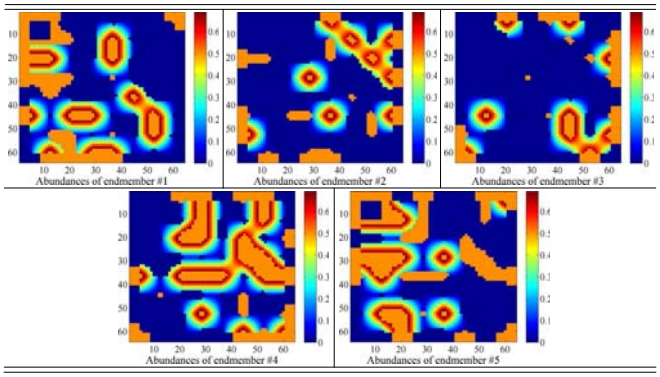


Fig. 7. True fractional abundances of endmembers in DS3.

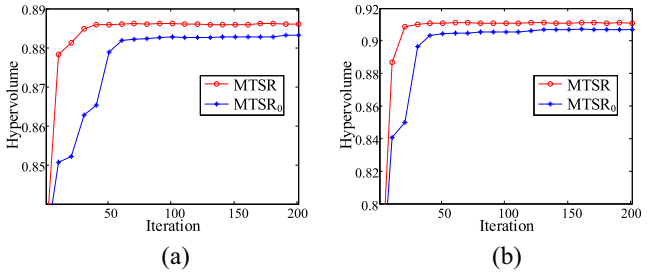


Fig. 8. Comparing the evolution of the hypervolume indicator for DS1. (a) Task 1. (b) Task 2.

library $\boldsymbol{A}$, we can construct different hyperspectral unmixing tasks by randomly selecting signatures from the library to form the endmembers. Therefore, the partition of the hyperspectral data is known beforehand and the procedure of determining the homogeneous regions is not needed. For every simulated pixel, the fractional abundances of the endmembers follow a Dirichlet distribution [56]. For DS1, the number of pixels for each task is set to 500 and the number of endmembers is set to 8.

The following two benchmark data sets are widely used. The second data set (DS2) has 100×100 pixels and is provided by Iordache *et al.* [57]. The fractional abundances are piecewise smooth and the observations show a good spatial homogeneity. Fig. 6 exhibits the true abundances of the endmembers. The third data set (DS3) is a hyperspectral image of 64×64 pixels, which is provided by Tang *et al.* [58]. There is no pure pixel in this hyperspectral data set. The fractional abundances of the five endmembers are shown in Fig. 7. The two data sets are corrupted by different levels of correlated noise (SNR = 20, 30, and 40 dB).

2) Test of the Between-Task Genetic Transfer: In this experiment, the hypervolume indicator is used to compare the convergence characteristics of the MTSR and $MTSR_0$. We prefer hypervolume indicator because the true PF is not known beforehand for real-world problems. The value of hypervolume is expected to increase gradually for an evolving population. Two tasks are constructed in the experiment on DS1. For the two benchmark problems, the number of tasks is set to 3.

The evolution of the hypervolume indicator for DS1 is shown in Fig. 8. Fig. 9 represents the evolution of the hypervolume indicator for DS2 and DS3. With a small number of iterations, the proposed method with between-task genetic transfer obtains higher hypervolume values than those of $MTSR_0$. As shown in Figs. 8 and 9, when multiple relevant sparse reconstruction tasks are optimized simultaneously by the MTSR with the between-task genetic transfer, the overall convergence characteristics can be significantly improved in

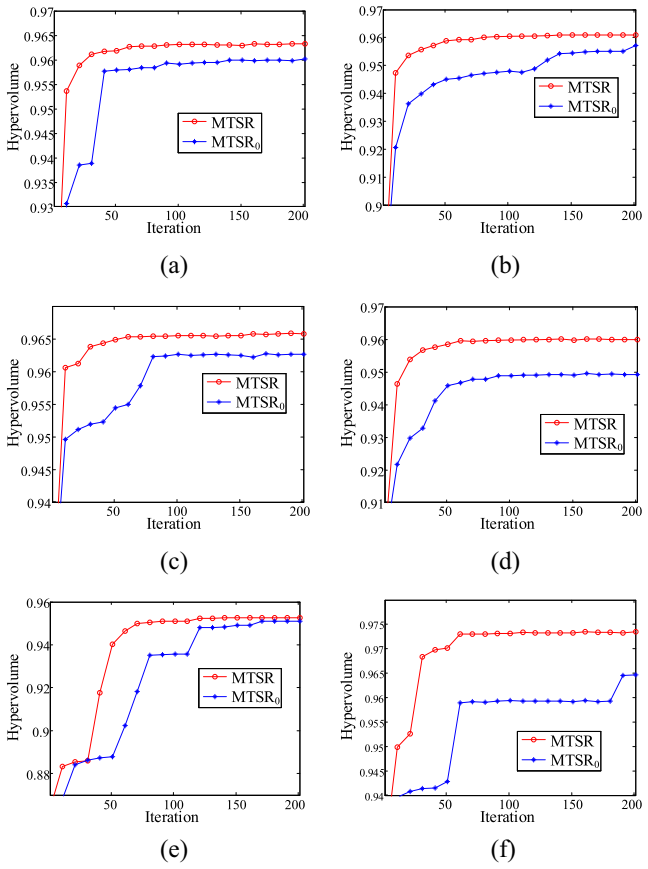


Fig. 9. Comparing the evolution of the hypervolume indicator for DS2 and DS3. (a) DS2, task 1. (b) DS3, task 1. (c) DS2, task 2. (d) DS3, task 2. (e) DS2, task 3. (f) DS3, task 3.

TABLE IV
PERFORMANCE COMPARISON WITH THE DIFFERENT UNMIXING METHODS ON DS1. THE VALUE OF τ IS SET TO 0.15

Method	SRE(dB)			SR_τ		
	20	30	40	20	30	40
NCLS	0.7815	3.5658	9.8945	0.18	0.33	0.67
SUnSAL	0.8215	3.6745	9.8945	0.18	0.35	0.69
CLSUnSAL	0.8413	3.8039	10.9431	0.19	0.35	0.73
RCSU	0.8749	3.9336	11.1352	0.19	0.36	0.79
$MTSR_0$	1.2587	5.0751	14.9283	0.21	0.41	0.88
MTSR	1.3806	5.4894	15.5121	0.22	0.44	0.91

comparison with the MTSR without the between-task genetic transfer.

3) Comparison of MTSR Against Other Methods: Fig. 10 graphically depicts the sparse solutions obtained by the different algorithms on the simulated data that comprises two tasks. Note that only 40 pixels are shown to clearly distinguish the differences between these figures. The two columns in Fig. 10 show the results of the two sparse unmixing tasks, respectively. The first row exhibits the ground-truth sparse solutions. From Fig. 10, the maps generated by the proposed method are more similar to the ground-truth maps than those of the other algorithms. Next, we compute the sum of the

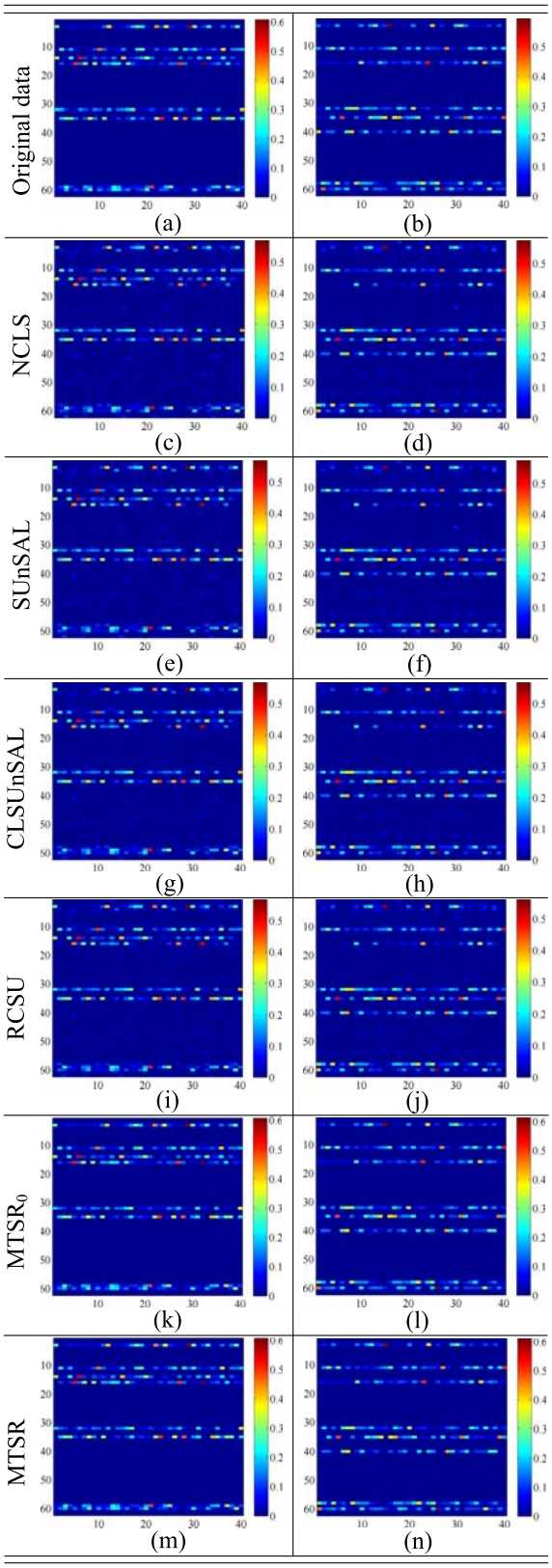


Fig. 10. Sparse solutions obtained by the different algorithms on the simulated data with two tasks.

elements in each row of the sparse matrix. The results are shown in Fig. 11. The blue lines represent the true results and the red lines denote the results obtained by different methods. Obviously, the results obtained by the proposed method

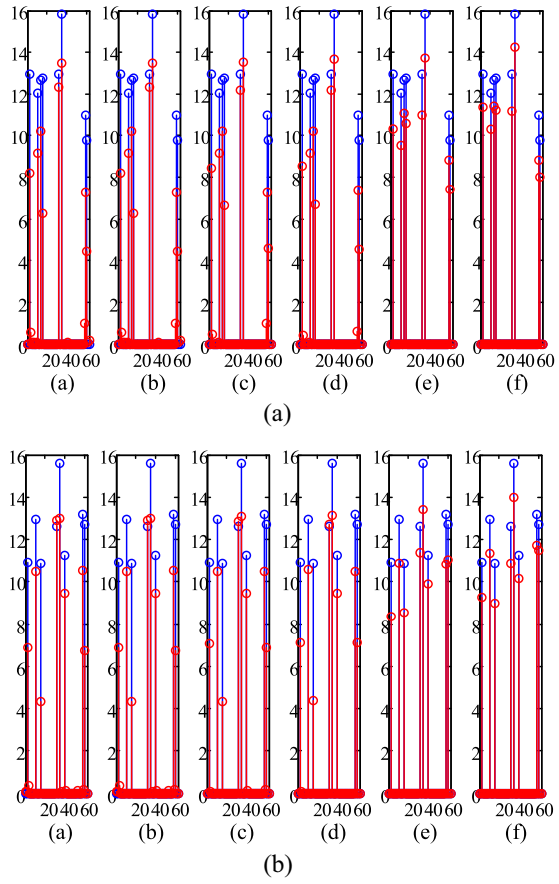


Fig. 11. Sum of the elements in each row of the sparse matrix. The blue lines represent the true results. The red lines denote the results obtained by (a) NCLS, (b) SUnSAL, (c) CLSUnSAL, (d) RCSU, (e) MTSR_0, and (f) MTSR.

are more close to the true results in comparison with those of the other schemes. Table IV shows the performance comparison with the different unmixing schemes on the simulated data with different noise levels. It can be observed that the proposed approach can achieve the highest values of SRE(dB) and SR_τ on the simulated data. The results of MTSR and MTSR_0 are much better than those of the competing methods because the proposed multiobjective framework can find a better compromise between the measurement error and the sparsity terms automatically. From Table IV, the proposed MTSR has better performance compared with that of the MTSR_0 method. The multitasking framework is able to enhance exploration of the entire search space and accelerate the convergence.

Figs. 12 and 13 graphically exhibit the estimated abundance maps for the selected endmembers in the two data sets as the estimated abundance maps. The other endmembers also show similar behavior. Fig. 12 shows the estimated abundance maps of endmember 1, 8, and 9 for DS1 obtained by the six sparse unmixing methods. The fractional abundance maps of endmember 1, 4, and 5 for DS2 are presented in Fig. 13. It can be observed that the CLSUnSAL, RCSU, and MTSR algorithms show better performance compared with the other two schemes. As shown in Fig. 12, the estimated abundance maps of endmember 1 obtained by NCLS and SUnSAL

TABLE V
PERFORMANCE COMPARISON WITH THE DIFFERENT UNMIXING METHODS ON DS2. THE VALUE OF τ IS SET TO 0.15

Method	SRE(dB)			SR_τ		
	20	30	40	20	30	40
NCLS	-7.3250	-1.6087	5.7983	0.23	0.35	0.54
SUnSAL	-7.0565	-1.5587	5.9169	0.34	0.45	0.65
CLSUnSAL	5.5164	11.4842	18.7935	0.62	0.75	0.91
RCSU	5.5829	11.5508	18.7923	0.63	0.75	0.91
MTSR_0	6.5744	12.9472	22.2723	0.67	0.78	1.00
MTSR	7.0496	13.7802	22.7329	0.70	0.80	1.00

TABLE VI
PERFORMANCE COMPARISON WITH THE DIFFERENT UNMIXING METHODS ON DS3. THE VALUE OF τ IS SET TO 0.15

Method	SRE(dB)			SR_τ		
	20	30	40	20	30	40
NCLS	-4.7118	3.8059	13.2893	0.15	0.43	0.74
SUnSAL	-4.4264	4.0443	13.9636	0.20	0.50	0.86
CLSUnSAL	8.2382	13.0988	14.3502	0.65	0.83	0.88
RCSU	8.3170	13.2029	14.3730	0.67	0.85	0.88
MTSR_0	9.1933	14.0197	16.9749	0.68	0.86	0.92
MTSR	10.7254	14.6143	17.6775	0.75	0.89	0.93

contain many pixels with cool color. It can be seen that the abundance maps acquired by the proposed MTSR method are more close to the true abundance maps in comparison with the abundance maps obtain by the other schemes. Although the abundance maps in Fig. 13 acquired by different methods are similar, the abundance maps obtained by the proposed MTSR method contain much less noise than those of the other algorithms because the pixels in the homogeneous region are likely to share similar endmembers. Table V exhibits the results of SRE(dB) and SR_τ obtained by the six sparse unmixing methods on DS2. The results of SRE(dB) and SR_τ obtained by the six sparse unmixing methods on DS3 are shown in Table VI. It is obvious that the results of MTSR and MTSR_0 are significantly better than those of the other methods. It is appropriate to consider the unmixing problem involving the pixels in a homogeneous region as a task. The experimental results with two hyperspectral data sets demonstrate that the proposed MTSR method can obtain satisfactory solutions by establishing multiple optimization tasks and optimizing them simultaneously.

VI. CONCLUSION

Recently, evolutionary multiobjective optimization has been successfully used to solve SMV-based sparse reconstruction problems for eliminating the estimation of the sparsity. The real-world applications always exist multiple sparse reconstruction tasks that have to be solved simultaneously. This paper has proposed an MTSR framework to optimize multiple sparse optimization tasks simultaneously. To deal with MMV problems, evolutionary algorithms have been employed to

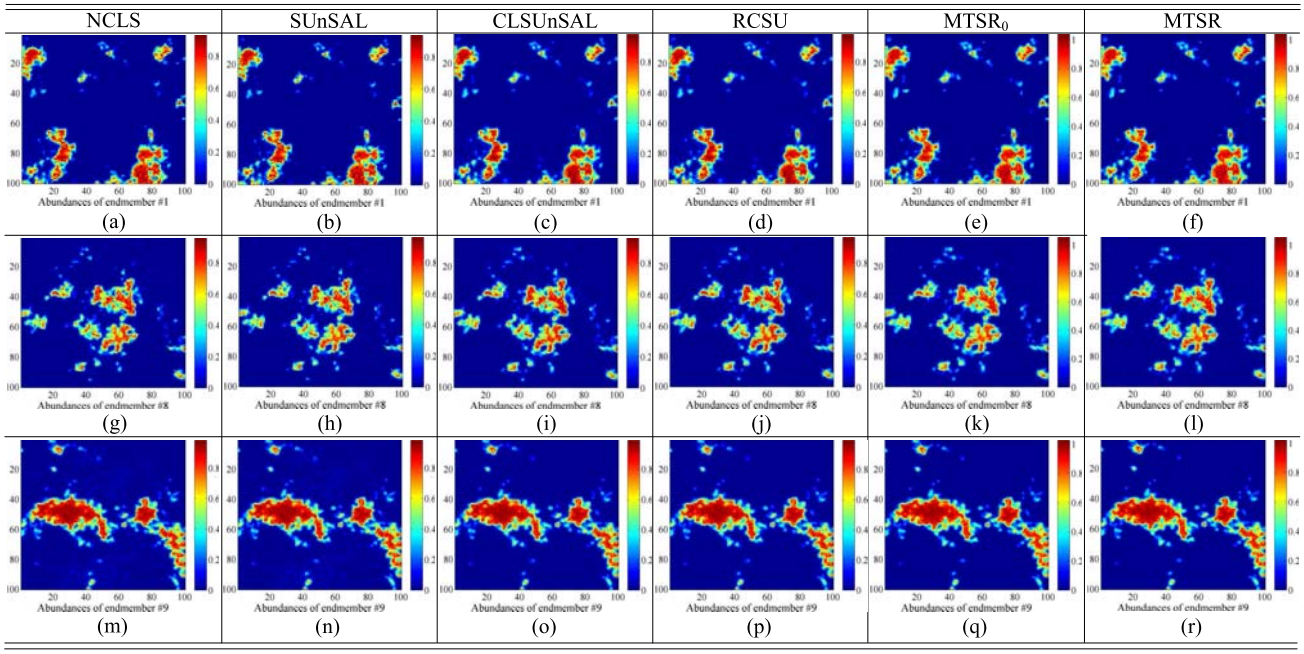


Fig. 12. Fractional abundance maps of endmember 1, 8, and 9 for DS1 estimated by NCLS, SUnSAL, CLSunSAL, RCSU, $MTSR_0$, and MTSR.

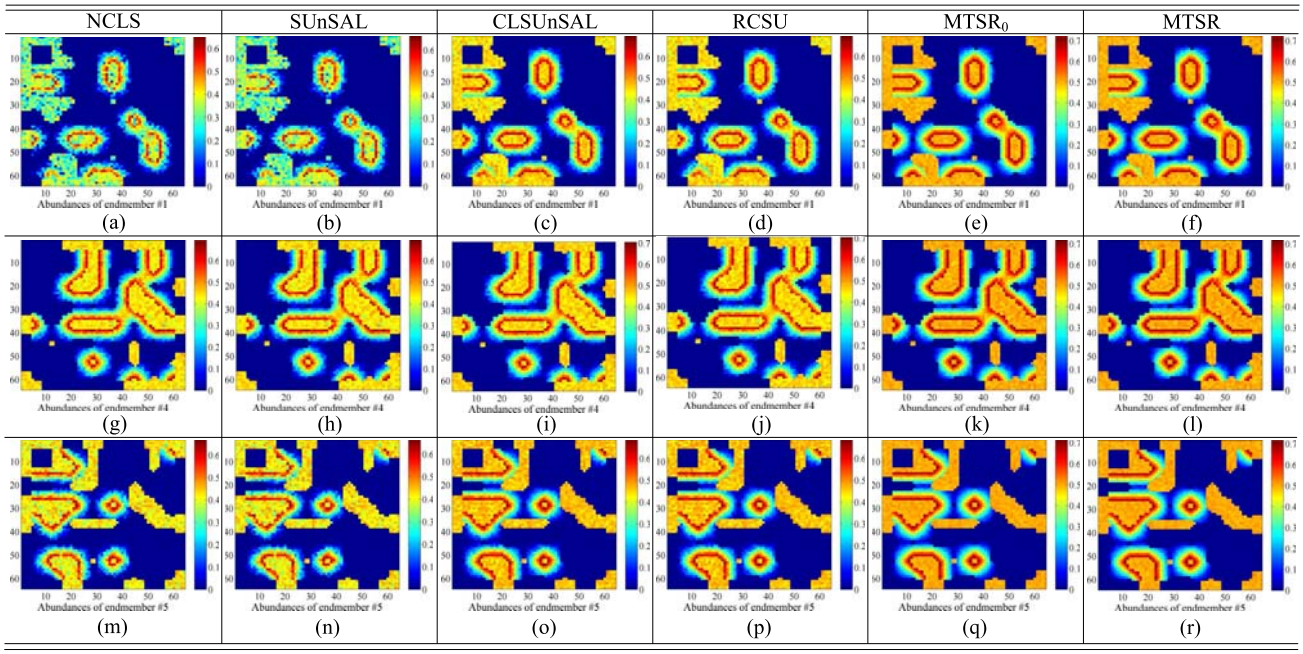


Fig. 13. Fractional abundance maps of endmember 1, 4, and 5 for DS2 estimated by NCLS, SUnSAL, CLSunSAL, RCSU, $MTSR_0$, and MTSR.

search the locations of the nonzero rows since MMV problems suffer from the curse of dimensionality with the increasing number of measurement vectors. For a same sensing matrix or over-complete dictionary, the decision variables for the MMV model have the same length as the SMV model in the proposed technique. Genetic materials belonging to the same or different tasks have been, respectively, exchanged via the within-task or between-task genetic transfer to exploit the similar sparsity pattern and accelerate convergence.

A real-world application, hyperspectral unmixing, has been considered as a case study in this paper. A sparse unmixing problem involving the pixels in a homogeneous region is considered as a task and then the proposed MTSR framework is employed to solve these tasks simultaneously. Experimental studies on signal reconstruction and hyperspectral unmixing have demonstrated the superiority of the proposed algorithm. In the future, we hope to propose an efficient multitasking optimization algorithm to deal with sequential sparse reconstruction tasks.

REFERENCES

[1] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2479–2493, Jul. 2009.

[2] J. Wright *et al.*, "Sparse representation for computer vision and pattern recognition," *Proc. IEEE*, vol. 98, no. 6, pp. 1031–1044, Jun. 2010.

[3] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.

[4] G. Davis, "Adaptive nonlinear approximations," Ph.D. dissertation, Graduate School Arts Sci., New York Univ., New York, NY, USA, 1994.

[5] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.

[6] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, 2001.

[7] L. Li, X. Yao, R. Stolkin, M. Gong, and S. He, "An evolutionary multiobjective approach to sparse reconstruction," *IEEE Trans. Evol. Comput.*, vol. 18, no. 6, pp. 827–845, Dec. 2014.

[8] Y. Zhou, S. Kwong, H. Guo, X. Zhang, and Q. Zhang, "A two-phase evolutionary approach for compressive sensing reconstruction," *IEEE Trans. Cybern.*, vol. 47, no. 9, pp. 2651–2663, Sep. 2017.

[9] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2477–2488, Jul. 2005.

[10] J. Chen and X. Huo, "Theoretical results on sparse representations of multiple-measurement vectors," *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4634–4643, Dec. 2006.

[11] M. F. Duarte and Y. C. Eldar, "Structured compressed sensing: From theory to applications," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4053–4085, Sep. 2011.

[12] Z. Zhang and B. D. Rao, "Sparse signal recovery with temporally correlated source vectors using sparse Bayesian learning," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 5, pp. 912–926, Sep. 2011.

[13] M. E. Davies and Y. C. Eldar, "Rank awareness in joint sparse recovery," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 1135–1146, Feb. 2012.

[14] J. A. Zhang, Z. Chen, P. Cheng, and X. Huang, "Multiple-measurement vector based implementation for single-measurement vector sparse Bayesian learning with reduced complexity," *Signal Process.*, vol. 118, pp. 153–158, Jan. 2016.

[15] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Collaborative sparse regression for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 341–354, Jan. 2014.

[16] 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 Trans. Geosci. Remote Sens.*, vol. 52, no. 7, pp. 4364–4382, Jul. 2014.

[17] H. Li, X. Su, Z. Xu, and Q. Zhang, "MOEA/D with iterative thresholding algorithm for sparse optimization problems," in *Proc. 12th Int. Conf. Parallel Problem Solving Nat.*, Taormina, Italy, 2012, pp. 93–101.

[18] J. Luo, L. Jiao, and J. A. Lozano, "A sparse spectral clustering framework via multiobjective evolutionary algorithm," *IEEE Trans. Evol. Comput.*, vol. 20, no. 3, pp. 418–433, Jun. 2016.

[19] H. Li, Q. Zhang, J. Deng, and Z.-B. Xu, "A preference-based multiobjective evolutionary approach for sparse optimization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 5, pp. 1716–1731, May 2018.

[20] 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, Apr. 2002.

[21] 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, Dec. 2007.

[22] K. Lounici, M. Pontil, A. Tsybakov, and S. Van De Geer, "Taking advantage of sparsity in multi-task learning," in *Proc. 22nd Conf. Learn. Theory*, Montreal, QC, Canada, 2009, p. 10.

[23] A. Gupta, Y.-S. Ong, and L. Feng, "Multifactorial evolution: Toward evolutionary multitasking," *IEEE Trans. Evol. Comput.*, vol. 20, no. 3, pp. 343–357, Jun. 2016.

[24] A. Gupta, Y.-S. Ong, L. Feng, and K. C. Tan, "Multiobjective multifactorial optimization in evolutionary multitasking," *IEEE Trans. Cybern.*, vol. 47, no. 7, pp. 1652–1665, Jul. 2017.

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

[26] K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms*. New York, NY, USA: Wiley, 2001.

[27] N. Beume, B. Naujoks, and M. Emmerich, "SMS-EMOA: Multiobjective selection based on dominated hypervolume," *Eur. J. Oper. Res.*, vol. 181, no. 3, pp. 1653–1669, 2007.

[28] R. Caruana, "Multitask learning," *Mach. Learn.*, vol. 28, no. 1, pp. 41–75, 1997.

[29] A. T. W. Min, Y.-S. Ong, A. Gupta, and C.-K. Goh, "Multi-problem surrogates: Transfer evolutionary multiobjective optimization of computationally expensive problems," *IEEE Trans. Evol. Comput.*, preprint. [Online]. Available: https://ieeexplore.ieee.org/document/8207638, doi: 10.1109/TEVC.2017.2783441.

[30] A. Gupta, Y.-S. Ong, and L. Feng, "Insights on transfer optimization: Because experience is the best teacher," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 2, no. 1, pp. 51–64, Feb. 2018.

[31] M. Gong, H. Li, E. Luo, J. Liu, and J. Liu, "A multiobjective cooperative coevolutionary algorithm for hyperspectral sparse unmixing," *IEEE Trans. Evol. Comput.*, vol. 21, no. 2, pp. 234–248, Apr. 2017.

[32] J. A. Tropp, A. C. Gilbert, and M. J. Strauss, "Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit," *Signal Process.*, vol. 86, no. 3, pp. 572–588, 2006.

[33] J. A. Tropp, "Algorithms for simultaneous sparse approximation. Part II: Convex relaxation," *Signal Process.*, vol. 86, no. 3, pp. 589–602, 2006.

[34] L. Feng, Y.-S. Ong, S. Jiang, and A. Gupta, "Autoencoding evolutionary search with learning across heterogeneous problems," *IEEE Trans. Evol. Comput.*, vol. 21, no. 5, pp. 760–772, Oct. 2017.

[35] G. Syswerda, "Uniform crossover in genetic algorithms," in *Proc. 3rd Int. Conf. Genet. Algorithms*, 1989, pp. 2–9.

[36] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading, MA, USA: Addison-Wesley, 1989.

[37] M. Mishali and Y. C. Eldar, "Reduce and boost: Recovering arbitrary sets of jointly sparse vectors," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4692–4702, Oct. 2008.

[38] J. D. Blanchard, M. Cermak, D. Hanle, and Y. Jing, "Greedy algorithms for joint sparse recovery," *IEEE Trans. Signal Process.*, vol. 62, no. 7, pp. 1694–1704, Apr. 2014.

[39] X. Ma *et al.*, "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, Apr. 2016.

[40] A. Plaza, P. Martínez, R. Pérez, and J. Plaza, "A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 3, pp. 650–663, Mar. 2004.

[41] C.-I. Chang and A. Plaza, "A fast iterative algorithm for implementation of pixel purity index," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 63–67, Jan. 2006.

[42] M. Zortea and A. Plaza, "A quantitative and comparative analysis of different implementations of N-FINDR: A fast endmember extraction algorithm," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 4, pp. 787–791, Oct. 2009.

[43] M. Berman *et al.*, "ICE: A statistical approach to identifying endmembers in hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 10, pp. 2085–2095, Oct. 2004.

[44] 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, Jun. 2011.

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

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

[47] J. Li, Q. Du, and Y. Li, "Region-based collaborative sparse unmixing of hyperspectral imagery," in *Proc. Remotely Sens. Data Compression Commun. Process. XII*, vol. SPIE 9874. Baltimore, MD, USA, 2016.

[48] A. Paoli, F. Melgani, and E. Pasolli, "Clustering of hyperspectral images based on multiobjective particle swarm optimization," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 12, pp. 4175–4188, Dec. 2009.

[49] X. Lan, A. J. Ma, P. C. Yuen, and R. Chellappa, "Joint sparse representation and robust feature-level fusion for multi-cue visual tracking," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5826–5841, Dec. 2015.

[50] X. Lan, M. Ye, S. Zhang, H. Zhou, and P. C. Yuen, "Modality-correlation-aware sparse representation for RGB-infrared object tracking," *Pattern Recogn. Lett.*, Oct. 9, 2018, in press. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167865518307633
[51] X. Lan, S. Zhang, P. C. Yuen, and R. Chellappa, "Learning common and feature-specific patterns: A novel multiple-sparse-representation-based tracker," *IEEE Trans. Image Process.*, vol. 27, no. 4, pp. 2022–2037, Apr. 2018.
[52] T. Blumensath and M. E. Davies, "Normalized iterative hard thresholding: Guaranteed stability and performance," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 298–309, Apr. 2010.
[53] H. Seada and K. Deb, "A unified evolutionary optimization procedure for single, multiple, and many objectives," *IEEE Trans. Evol. Comput.*, vol. 20, no. 3, pp. 358–369, Jun. 2016.
[54] H. Ishibuchi, R. Imada, Y. Setoguchi, and Y. Nojima, "How to specify a reference point in hypervolume calculation for fair performance comparison," *Evol. Comput.*, vol. 26, no. 3, pp. 411–440, 2018.
[55] D. Heize and C. Chang, "Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 3, pp. 529–545, Mar. 2001.
[56] J. M. P. Nascimento and J. M. B. Dias, "Vertex component analysis: A fast algorithm to unmix hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 898–910, Apr. 2005.
[57] M.-D. Iordache, J. M. Bioucas-Dias, and A. Plaza, "Total variation spatial regularization for sparse hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 11, pp. 4484–4502, Nov. 2012.
[58] W. Tang, Z. Shi, Y. Wu, and C. Zhang, "Sparse unmixing of hyperspectral data using spectral a priori information," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 2, pp. 770–783, Feb. 2015.

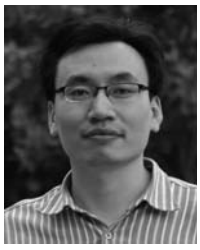
Hao Li received the B.S. degree in electronic engineering from Xidian University, Xi'an, China, in 2013, where he is currently pursuing the Ph.D. degree in pattern recognition and intelligent systems with the School of Electronic Engineering.

His current research interests include computational intelligence and machine learning.

Yew-Soon Ong (M'99–SM'12–F'18) received the Ph.D. degree in artificial intelligence in complex design from the Computational Engineering and Design Center, University of Southampton, Southampton, U.K., in 2003.

He is a Professor with the School of Computer Science and Engineering, Nanyang Technological University (NTU), Singapore, where he is also the Director of the Data Science and Artificial Intelligence Research Center and a Principal Investigator of the Data Analytics and Complex Systems Programme with the Rolls-Royce@NTU Corporate Laboratory, Singapore. His current research interests include computational intelligence, memetic computing, complex design optimization, and big data analytics.

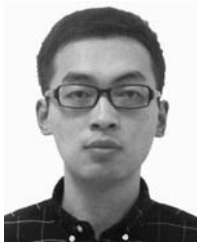
Dr. Ong is the Founding Editor-in-Chief of the IEEE TRANSACTIONS ON EMERGING TOPICS IN COMPUTATIONAL INTELLIGENCE, the Founding Technical Editor-in-Chief of *Memetic Computing*, and an Associate Editor of the IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION, the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, the IEEE TRANSACTIONS ON CYBERNETICS, and the IEEE TRANSACTIONS ON BIG DATA.



Maoguo Gong (M'07–SM'14) received the B.S. degree in electronic engineering (First Class Hons.) 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, where he was promoted as an Associate Professor and as a Full Professor in 2008 and 2010, respectively, both with exceptive admission. His current research interests include computational intelligence with applications to optimization, learning, data mining, and image understanding.

Dr. Gong was a recipient of 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 an Associate Editor of the IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION and the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS.



Zhenkun Wang received the Ph.D. degree in circuits and systems from Xidian University, Xi'an, China, in 2016.

He is currently a Research Fellow with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. His current research interests include multiobjective optimization, evolutionary computation, test problem construction, and some machine learning techniques.