文章编号：1001－9014(2018)02－0144－10

DOI:10.11972/j.issn.1001－9014.2018.02.004

Sparse and low-rank abundance estimation with structural information

YUAN Jing[1], ZHANG Yu-Jin[1], YANG De-He[2]

(1. Department of Electrical and Engineering, Tsinghua University, Beijing 100084, China;
2. Institute of Crustal Dynamics, China Earthquake Administration, Beijing 100085, China)

Abstract: Abundance estimation (AE) plays an essential role in the hyperspectral image processing and analysis. Owing to the simplicity and mathematical tractability, various methods based on the constrained linear regression are usually developed to estimate abundance matrix. The obvious limitation of these approaches is that the fitness between the estimated data and ground-truth data does not include the structural information, e. g. row difference and column difference. In this paper, a novel linear regression algorithm is proposed by jointly adding the multi-structured information to the traditional linear regression model. And it is employed to modify sparse and low-rank abundance estimation model to improve estimated accuracy and robustness. Firstly, a new linear regression model is established by taking into account the structural information. Then, mathematical proof of the new linear regression method is presented. Afterwards, it is applied to modify the sparse low-rank abundance estimation model. Finally, Alternating Direction Method of Multipliers(ADMM) technique is adopted to solve the new model. The experimental results demonstrate that the proposed algorithms can capture structural information and improve the estimated performance on the simulated dataset and the real hyperspectral remote sensing images.

Key words: unmixing, sparse and low rank, structural information, abundance matrix, alternating direction method of multipliers (ADMM)

PACS: 89.70.-a

融入结构信息的稀疏低秩丰度估计算法研究

袁 静[1], 章毓晋[1], 杨德贺[2]

(1. 清华大学 电子工程系,北京 100084;
2. 中国地震局地壳应力研究所,北京 100085)

摘要:丰度估计(AE)是从高光谱图像中识别地物的关键预处理技术.鉴于线性混合模型的可解释性以及数学上的可操作性,带约束的线性回归技术在丰度矩阵估计中备受关注.目前,这类方法存在的缺陷是其拟合过程中仅仅考虑到估计数据与真实数据之间的拟合误差,忽略了估计数据的结构与真实数据的结构之间的相似性信息.因此,提出了融合结构信息的线性回归模型,并应用于稀疏低秩丰度矩阵估计领域.首先,通过增加结构信息改进传统的带约束的线性回归模型,并经数学理论证明了增加结构信息的模型较传统模型更加有效;其次,应用该方法改进稀疏低秩丰度估计的数学模型;最后,采用交替乘子法(ADMM)技术求解新模型.实验结果表明,融入结构信息的稀疏低秩丰度估计算法能够有效地提高仿真数据和实际高光谱数据的丰度估计的估计精度,改善其抗噪性能.

关 键 词:解混;稀疏低秩;结构信息;丰度矩阵;交替乘子法(ADMM)

中图分类号:TP751 **文献标识码**:A

Introduction

Hyperspectral remote sensing has gained considerable attention due to its wide range of applications including planetary exploration, precision agriculture, military target identification, forest research, target detection and so on[1-11]. In recent years, it has been widely

Received date: 2017-08-13, **revised date**: 2018-01-13 **收稿日期**:2017-08-13, **修回日期**:2018-01-13
Foundation items:Supported by National Natural Science Foundation of China (61673234,U1636124)
Biography:YUAN Jing(1981-),female, Shijiazhuang, China, Ph. D. candidate. Research fields include remote sensing process and computer vision.

used in environmental monitoring[11-20]. Hyperspectral sensors can sample the electromagnetic spectrum in tens or hundreds of contiguous spectral bands from the visible to the near-infrared region. However, due to their low spatial resolution, more than one different materials can be mixed in a single pixel, which hinders the application and development of hyperspectral technology[1,21-22] and calls for spectral unmixing. In spectral unmixng, the spectrum of a mixed pixel is decomposed into a collection of constituent spectra(endmembers) and a set of corresponding fractions(abundances). However, the identification of the endmembers is challenging due to the insufficient spatial resolution and the unavailability of completely pure pixels in the scene. And an inaccurate endmember leads to unreliable fractional abundance[23-24]. To avoid the limitation, sparse unmixing as a promising approach is the focus of this article. Sparse unmixing is based on the premise hypothesis that: from the large-scale view of point, the observed spectra can be expressed as linear combinations of a number of pure spectral signatures known in advance[24-35]. Therefore, finding the optimal subset of signatures from a large number of spectral samples to best model the mixed pixel in the scene results in a sparse abundance matrix.

The semi-supervised sparse unmixing is termed as sparse abundance estimation. The objective of abundance estimation is to determine how many and which endmembers are present in the mixed pixel under study and to estimate their corresponding abundance. The sparsity is imposed on abundances by means of ℓ_0 norm regularization[25]. However, regularizing the ℓ_0 item is a NP hard optimization problem, which is complex and difficult to be solved. Reference [26] has proved that under the RIP condition, it is reasonable to relax the $\ell_0$norm to ℓ_1 norm. This lays the ground for applying ℓ_1 norm to express the sparseness.

Instead of ℓ_0 norm, the ℓ_1 norm can improve the accuracy of abundance matrix estimation[24,27]. Reference [28] has argued that ℓ_p norm ($0<p<1$) is more effective than the ℓ_1 norm. Meanwhile, a method that approximates the ℓ_0 norm has been put forward in Ref. [29]. It works more better than employingℓ_1 norm at the cost of higher computational complexity. Also, Refs. [30-31] have proposed various sparse regularization item to approximate the ℓ_0 norm to obtain accurate abundance. However, they require much more computational time. On the other hand, in Bayesian schemes, appropriate sparsity inducing priors have been adopted for the abundance vectors[32-33].

The above methods have been based on the prior knowledge about the spectral domain. From the perspective of the spatial domain, Refs. [34-35] have considered that the spectra of homogeneous objects in space are highly similar, so their corresponding abundance vectors have high correlation. Reference [34] has assumed that all pixels belonging to the same window are correlated, i. e., they are composed of the same materials, although maybe in different proportions. This property suggests that the abundance matrix $\boldsymbol{W}$ to be estimated has linearly dependent columns and thus is either low-rank, or it can be well-approximated by a low-rank matrix. Alternatively, Ref. [35] has developed a collaborative deterministic scheme, whereby the information provided by the neighboring pixels is taken into account in the abundance estimation of each single pixel. The method termed as CLSUnSAL uses a wealth of information stemming from all the pixels of the examined HIS. To impose joint-sparsity, CLSUnSAL has applied a $\ell_{2,1}$ norm on the sought abundance matrix, which is then used to penalize a suitably defined quadratic cost function. Following the spirit to develop the localized approaches, Ref. [36] has proposed the use of a 3×3 square window that slides all over the image. Then, the abundance vector of the central pixel has been inferred by taking into account the spectral signatures of the adjacent pixels contained in the window. Based on homogeneous area in local regions[34] and sliding window[36] method, Ref. [37] has proposed to seek for 3^2-column abundance matrices which are simultaneously sparse and low-rank. According to inverse theories in signal processing and machine learning literatures[38-42] and low-rank matrix estimation techniques[43-46], the abundance estimation solved by the alternating directional multiplier method (ADMM)[47] has yielded many advantages over other algorithms.

In all the above abundance estimation algorithms, they cannot consider the structural information. The Oxford English Dictionary tells that a structure is an arrangement and organization of interrelated elements in a material object or system, or the object or system so organized[48]. Abstract structures include data structures in computer science and musical form. Types of structure include a hierarchy (a cascade of one-to-many relationships), a network featuring many-to-many links, or a lattice featuring connections between components that are neighbors in space[48].

As far as hyperspectral image is concerned, "structure" means the pixel-value distance between any group of points. If the map consists of such distant value, it is referred to as structural information map. It is quite clear that each image owns lots of structure information maps. It inspires that if one image matrix $\widehat{\boldsymbol{X}}$ is the best estimation of ground-truth matrix $\boldsymbol{X}$, the fitness is not only from $\|\widehat{\boldsymbol{X}}-\boldsymbol{X}\|_F^2$ but also from $\|\boldsymbol{M}\widehat{\boldsymbol{X}}-\boldsymbol{M}\boldsymbol{X}\|_F^2$. In $\|\boldsymbol{M}\widehat{\boldsymbol{X}}-\boldsymbol{M}\boldsymbol{X}\|_F^2$, $\boldsymbol{M}$ is the operational matrix which could extract structural information map, $\boldsymbol{X}$ and $\widehat{\boldsymbol{X}}$ are termed as original map and $\boldsymbol{M}\boldsymbol{X}$ as well as $\boldsymbol{M}\widehat{\boldsymbol{X}}$ are referred to as structural information map. The structural information map is called structural information for short. In fact, the structural information is firstly used in Ref. [49]. It applies the row-difference information and column-difference as structural information to the auto-encoder neural network to improve the auto-encoding performance. However, it cannot introduce detailed analysis and explanation. To address the issue further, a novel model with structural information is proposed for hyperspectral image unmixing. Firstly, a $k\times k$ square sliding window is extracted. Secondly, the sparse and low rank abundance estimation model in Ref. [37] is modified through adding a penalty term which controls the fitness error be-

tween structural information to estimate k^2-column matrix. Finally, the proposed model is solved by ADMM techniques. Thus, the proposed method is referred to as the alternating direction sparse and low-rank unmixing algorithm with the structural information (ADSpLRU-SI). In order to compare the performances of the proposed approach with the competing methods such as ADSpLRU[37] which prevails over CSUnSAL[35] and BiICE[33] methods, experiments on simulated and real data are performed extensively.

The rest of this paper is organized as follows. In Sect. 1, the sparse low-rank abundance estimation algorithms are described. And in nature, it is considered as a constrained sparse low-rank linear regression. Section 2 elaborates the proposed model and uses mathematical theory to prove that it has more favorable accuracy than the traditional method. Moreover, it gives the detailed procedure for applying ADMM to solve the new model. Section 3 introduces the experimental results and discussions. Finally, the conclusions are drawn in Sect. 4.

1 Problem formulation

This section will formulate hyperspectral image unmixing. To express our work clearly, the notations are summarized in the Table. 1.

Assume an L-spectral band hyperspectral image whose pixel being composed of N endmembers. The endmembers dictionary is denoted by $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_N]$, where $\theta_i (i = 1, 2, \cdots, N)$ stands for the spectral signature of the *i-th* endmember. As is shown graphically in Fig. 1, define a small $k \times k$ sliding square window, which contains K adjacent pixels ($K = k \times k$). And, $\boldsymbol{Y} = [y_1, y_2, \cdots, y_K] \in R^{L \times K}$ including the spectra of the K pixels in the window as its columns, where $y_m (m = 1, 2, \cdots K)$ termed as measurement spectra, share the same endmember dictionary matrix $\boldsymbol{\theta}$. The mixing process is modeled by a linear mixing model(LMM)

$$\boldsymbol{Y} = \boldsymbol{\theta W} + \boldsymbol{E} \quad . \quad (1)$$

The abundances have to obey two constraints, namely, abundance nonnegativity constraint(ANC) and the abundance sum-to-one constraint(ASC), that is

$$W > 0; 1^{\mathrm{T}} W = 1^{\mathrm{T}} \quad . \quad (2)$$

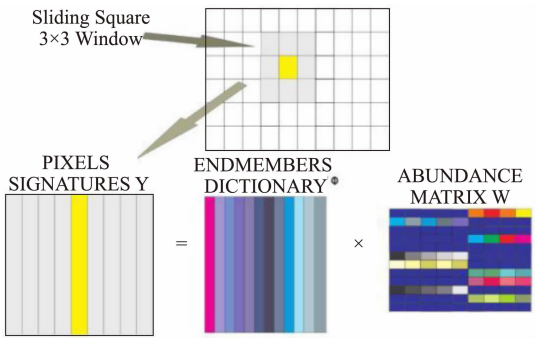

Fig. 1 Graphical illustration of the sliding window approach of the proposed algorithms[38]
图 1 基于滑动窗口的稀疏低秩丰度估计的示例图

Table 1 Notation Table
表 1 符号表

N	number of endmembers
L	number of spectral band
k	the size of the sliding window
K	number of pixels
$\boldsymbol{Y} \in R^{L \times K}$	spectra of the pixels
$\boldsymbol{W} \in R^{N \times K}$	abundance matrix
$\boldsymbol{\theta} \in R^{L \times N}$	endmembers´dictionary
$\boldsymbol{E} \in R^{L \times K}$	noise matrix
R^k	k-dimensional Euclidean space
R^k_+	k-dimensional non-negative orthant
0	zero matrix
1	ones matrix
X	matrix
x	vector
$\|\boldsymbol{X}\|_*$	nuclear norm of a matrix
$\|\boldsymbol{X}\|_1$	1 norm of a matrix
$\|\boldsymbol{X}\|_F$	Frobenius norm of a matrix
X^{T}	transpose of a matrix
tr(.)	the trace of the matrix

As far as sum-to-one constraint is concerned, the sum-to-one constraint is relaxed as following Ref. [37]. In other words, holding the sparsity and low-rank property on the abundance matrix without sum-to-one property, the problem is formulated by Ref. [37] as follows:

$$(\mathrm{P}): \widehat{\boldsymbol{W}} = \mathrm{argmin}_{\boldsymbol{W} \in R_+^{N \times K}} \{ \frac{1}{2} \|\boldsymbol{Y} - \boldsymbol{\theta W}\|_F^2 + \gamma \|\boldsymbol{W}\|_1 + \tau \|\boldsymbol{W}\|_* + I_{R_+}(\boldsymbol{W}) \} \quad , \quad (3)$$

where parameters γ, τ manage the trade-off between the sparsity and rank regularization terms and the data fidelity terms. The non-negativity constraint is denoted by $I_{R_+}(\boldsymbol{W})$.

Structural information: The above model motivates us that if the $\boldsymbol{\theta W}$ is the best estimation of $\boldsymbol{Y}$, it is not only that the original matrix is similar to the ground-truth matrix but also the structural information of the $\boldsymbol{Y}$ is supposed to be similar to the structural information in the $\boldsymbol{\theta W}$. To express our motivation conveniently, $\boldsymbol{M}$ is the operational matrix for calculating structural information. Section IV will present the $\boldsymbol{M}$ in detail. So, the problem in Eq. 3 can be reformulated as follows:

$$(\mathrm{PRM}): \widehat{\boldsymbol{W}} = \mathrm{argmin}_{\boldsymbol{W} \in R_+^{N \times K}} \{ \frac{1}{2} \|\boldsymbol{Y} - \boldsymbol{\theta W}\|_F^2 + \gamma \|\boldsymbol{W}\|_1 + \tau \|\boldsymbol{W}\|_* + \frac{1}{2} \kappa \|\boldsymbol{YM} - \boldsymbol{\theta WM}\|_F^2 + I_{R_+}(\boldsymbol{W}) \} \quad , \quad (4)$$

PRM is used to denote the proposed model. Later, the advantages over the model in (3) are demonstrated by the mathematical theory and experiments.

2 The proposed algorithms

This section presents the understandable mathematical theory to support the novel model. Additionally, it describes that how to solve the PRM model with the splitting strategy of the ADMM[47].

2.1 Proof the model PRM

This subsection presents a proof that the novel model is more effective than the old model. Without sparse

and low-rank constraint, if we guess that $\boldsymbol{W}$ has the $\widehat{\boldsymbol{W}}$, then we are implicitly making the guess that $\boldsymbol{E}$ has the value $\boldsymbol{Y}-\boldsymbol{\theta}\widehat{\boldsymbol{W}}$. Assuming that smaller values of $\boldsymbol{E}$ (measured by $\|.\|_F$ are more plausible than larger values, the most plausible guess for $\boldsymbol{W}$ is $\boldsymbol{Y}-\boldsymbol{\theta}\widehat{\boldsymbol{W}}$[50]. To prove the performance of the model PRM using the theory in Ref. [50], firstly, the old model in Ref. [37] is defined as

$$(\mathrm{P}):\widehat{\boldsymbol{W}}=\operatorname{argmin}_{\boldsymbol{W}\in R_+^{N\times K}}\left\{\frac{1}{2}\|\boldsymbol{Y}-\boldsymbol{\theta}\boldsymbol{W}\|_F^2+\gamma\|\boldsymbol{W}\|_1+\tau\|\boldsymbol{W}\|_*+I_{R_+}(\boldsymbol{W})\right\} \quad (5)$$

Secondly, the proposed model is represented as

$$(\mathrm{PRM}):\widehat{\boldsymbol{W}}=\operatorname{argmin}_{\boldsymbol{W}\in R_+^{N\times K}}\left\{\frac{1}{2}\|\boldsymbol{Y}-\boldsymbol{\theta}\boldsymbol{W}\|_F^2+\gamma\|\boldsymbol{W}\|_1+\tau\|\boldsymbol{W}\|_*+\frac{1}{2}\kappa\|\boldsymbol{Y}\boldsymbol{M}-\boldsymbol{\theta}\boldsymbol{W}\boldsymbol{M}\|_F^2+I_{R_+}(\boldsymbol{W})\right\} \quad (6)$$

Thirdly, since the κ and $\boldsymbol{M}$ are constant, a model which is equivalent to Eq. 5 is formulated as follows.

$$(\mathrm{P2}):\widehat{\boldsymbol{W}}=\operatorname{argmin}_{\boldsymbol{W}\in R_+^{N\times K}}\left\{\frac{1}{2}\|\boldsymbol{Y}-\boldsymbol{\theta}\boldsymbol{W}\|_F^2+\gamma\|\boldsymbol{W}\|_1+\tau\|\boldsymbol{W}\|_*+\frac{1}{2}\kappa\|\boldsymbol{M}\|_F^2\|\boldsymbol{Y}-\boldsymbol{\theta}\boldsymbol{W}\|_F^2+I_{R_+}(\boldsymbol{W})\right\}. \quad (7)$$

Fourthly, using the theory: $\|\boldsymbol{AB}\|_F \leqslant \|\boldsymbol{A}\|_F\|\boldsymbol{B}\|_F$[51], we have:

$$\|\boldsymbol{Y}\boldsymbol{M}-\boldsymbol{\theta}\boldsymbol{W}\boldsymbol{M}\|_F^2 \leqslant \|\boldsymbol{M}\|_F^2\|\boldsymbol{Y}-\boldsymbol{\theta}\boldsymbol{W}\|_F^2 . \quad (8)$$

Then, add $\frac{1}{2}\|\boldsymbol{Y}-\boldsymbol{\theta}\boldsymbol{W}\|_F^2+\gamma\|\boldsymbol{W}\|_1+\tau\|\boldsymbol{W}\|_*+I_{R_+}(\boldsymbol{W})$ to both sides of the inequality in Eq. 8, the following inequality is obtained:

$$\frac{1}{2}\|\boldsymbol{Y}-\boldsymbol{\theta}\boldsymbol{W}\|_F^2+\gamma\|\boldsymbol{W}\|_1+\tau\|\boldsymbol{W}\|_*+\frac{1}{2}\kappa\|\boldsymbol{Y}\boldsymbol{M}-\boldsymbol{\theta}\boldsymbol{W}\boldsymbol{M}\|_F^2+I_{R_+}(\boldsymbol{W}) \leqslant \frac{1}{2}\|\boldsymbol{Y}-\boldsymbol{\theta}\boldsymbol{W}\|_F^2+\gamma\|\boldsymbol{W}\|_1+\tau\|\boldsymbol{W}\|_*+\frac{1}{2}\kappa\|\boldsymbol{M}\|_F^2\|\boldsymbol{Y}-\boldsymbol{\theta}\boldsymbol{W}\|_F^2+I_{R_+}(\boldsymbol{W}) \quad (9)$$

Finally, observing the inequality in Eq. 9, we found that the left side is an objective function of PRM model, and the other side is an objective function of P2 model. Hence, considering the theory in Ref. [50], we conclude that the estimated $\boldsymbol{W}$ solved from PRM model in Eq. 6 could be more accurate than from P2 model in Eq. 7. And taking into account the fact that the P model in Eq. 5 is equivalent to P2 model in Eq. 7, we could draw the conclusion that the estimated $\boldsymbol{W}$ solved from PRM model in Eq. 6 could be more accurate than from P model in Eq. 5. More interestingly, if $\boldsymbol{Y}=\boldsymbol{\theta}\boldsymbol{W}$ or $\kappa=0$, PRM model degrades to P model.

Proof finished.

2.2 Alternating direction method of multipliers for the novel unmixing model

In this section, ADMM techniques are explored to optimize PRM model.

Firstly, its equivalent ADMM format is reformulated as follows.

$$(\mathrm{PRM2}):\widehat{\boldsymbol{W}}=\operatorname{argmin}_{\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_3\boldsymbol{\Sigma}_4\boldsymbol{\Sigma}_5}\left\{\frac{1}{2}\|\boldsymbol{Y}-\boldsymbol{\Sigma}_1\|_F^2+\gamma\|\boldsymbol{\Sigma}_2\|_1+\tau\|\boldsymbol{\Sigma}_3\|_*+\frac{1}{2}\kappa\|\boldsymbol{Y}\boldsymbol{M}-\boldsymbol{\Sigma}_5\|_F^2+I_{R_+}(\boldsymbol{\Sigma}_4)\right\}$$

$$\text{S. T. } \boldsymbol{\Sigma}_1-\boldsymbol{\theta}\boldsymbol{W}=0,\boldsymbol{\Sigma}_2-\boldsymbol{W}=0,\boldsymbol{\Sigma}_3-\boldsymbol{W}=0,\boldsymbol{\Sigma}_4-\boldsymbol{\theta}\boldsymbol{W}=0,\boldsymbol{\Sigma}_1-\boldsymbol{\theta}\boldsymbol{W}\boldsymbol{M}=0 \quad (10)$$

where $\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_3\boldsymbol{\Sigma}_4\boldsymbol{\Sigma}_5$ of proper dimensions are the auxiliary matrix variables (similar to Ref. [37]).

Secondly, the corresponding Augmented Lagrangian Function (ALF) is given as follows.

$$\mathrm{L}(\boldsymbol{W},\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2\boldsymbol{\Sigma}_3\boldsymbol{\Sigma}_4\boldsymbol{\Sigma}_5)=\frac{1}{2}\|\boldsymbol{Y}-\boldsymbol{\Sigma}_1\|_F^2+\gamma\|\boldsymbol{\Sigma}_2\|_1+\tau\|\boldsymbol{\Sigma}_3\|_*+\frac{1}{2}\kappa\|\boldsymbol{Y}\boldsymbol{M}-\boldsymbol{\Sigma}_5\|_F^2+I_{R_+}(\boldsymbol{\Sigma}_4)+tr(z_1^{\mathrm{T}}(\boldsymbol{\Sigma}_1-\boldsymbol{\theta}\boldsymbol{W}))+tr(z_2^{\mathrm{T}}(\boldsymbol{\Sigma}_2-\boldsymbol{W}))+tr(z_3^{\mathrm{T}}(\boldsymbol{\Sigma}_3-\boldsymbol{W}))+tr(z_4^{\mathrm{T}}(\boldsymbol{\Sigma}_4-\boldsymbol{W}))+tr(z_5^{\mathrm{T}}(\boldsymbol{\Sigma}_5-\boldsymbol{\theta}\boldsymbol{W}\boldsymbol{M}))+\frac{\mu}{2}(\|\boldsymbol{\Sigma}_1-\boldsymbol{\theta}\boldsymbol{W}\|_F^2+\|\boldsymbol{\Sigma}_2-\boldsymbol{W}\|_F^2+\|\boldsymbol{\Sigma}_3-\boldsymbol{W}\|_F^2+\|\boldsymbol{\Sigma}_4-\boldsymbol{W}\|_F^2+\kappa\|\boldsymbol{\Sigma}_5-\boldsymbol{\theta}\boldsymbol{W}\boldsymbol{M}\|_F^2) \quad (11)$$

where the $L\times K$ matrices $\boldsymbol{Z}_1,\boldsymbol{Z}_5$, and the $N\times K$ matrices $\boldsymbol{Z}_2,\boldsymbol{Z}_3,\boldsymbol{Z}_4$ are the Lagrange multipliers and $\mu>0$ is a positive penalty parameter.

Thirdly, the ADMM proceeds by minimizing $\mathrm{L}(\boldsymbol{W},\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2,\boldsymbol{\Sigma}_3,\boldsymbol{\Sigma}_4,\boldsymbol{\Sigma}_5)$ sequentially, each time with respect to a single matrix variable, keeping the remaining variables at their latest values. The dual variables are also updated via a gradient ascend step at the end of each alternating minimization cycle.

To elaborate the steps of the ADMM, the optimization with respect to $\boldsymbol{W}$ is given as following

$$\boldsymbol{W}^{\mathrm{t}+1}=\mathrm{dlyap}(\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}^{\mathrm{T}}) \quad (12)$$

where

$$\boldsymbol{J}^{\mathrm{t}}=\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\Sigma}_1^{\mathrm{t}}+\boldsymbol{\Sigma}_2^{\mathrm{t}}+\boldsymbol{\Sigma}_3^{\mathrm{t}}+\boldsymbol{\Sigma}_4^{\mathrm{t}}+\kappa\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\Sigma}_5^{\mathrm{t}}\boldsymbol{M}^{\mathrm{t}};$$

$$\boldsymbol{H}^{\mathrm{t}}=\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{Z}_1^{\mathrm{t}}+\boldsymbol{Z}_2^{\mathrm{t}}+\boldsymbol{Z}_3^{\mathrm{t}}+\boldsymbol{Z}_4^{\mathrm{t}}+\kappa\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{Z}_5^{\mathrm{t}}\boldsymbol{M}^{\mathrm{t}};$$

$$\boldsymbol{F}=(\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\theta}+3\boldsymbol{I}_N)^{-1};$$

$$\boldsymbol{A}=-\kappa\boldsymbol{F}\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\theta};\boldsymbol{B}=\boldsymbol{M}^{\mathrm{T}}\boldsymbol{M};\boldsymbol{C}^{\mathrm{t}}=\boldsymbol{F}(\boldsymbol{H}^{\mathrm{t}}+\boldsymbol{J}^{\mathrm{t}});$$

and dlyap(.) denotes the solution $\boldsymbol{X}$ in sylvester equation as the form $\boldsymbol{AXB}-\boldsymbol{X}+\boldsymbol{C}=0$[52]; t denotes the *t-th* iteration.

Fourthly, the optimization with respect to remaining variables such as $\boldsymbol{\Sigma}_i(i=1,2,3,4,5)$ and $\boldsymbol{Z}_j(j=1,2,3,4,5)$ is performed as Ref. [37].

At the end, the convergent conditions are set as following Ref. [47].

To represent the proposed method clearly, the proposed algorithm is termed as the alternating direction sparse and low-rank unmixing algorithm with structural information (ADSpLRU-SI). And it is summarized in Algorithm. 1, where the SHR(.) and SVT(.) are described in more detail by Ref. [37].

3 Experimental results and discussion

This section introduces experimental tests performed both on simulated and real data to reveal the performance of the proposed ADSpLRU-SI algorithms. Our techniques are compared with state-of-the art unmixing algorithms, namely, ADSpLRU[37], CSUnSAL[35] and Bi-ICE[33]. The computational time of all the tested algorithms is given in Table 2. The experiment is carried out on the computer with 4CPUs and 32G RAM to get the total time for unmixing the 9 pixels. Each experiment is run for 10 times. The result is calculated by (total time)/10. Table 2 shows that the ADSpLRU-SI costs more time than other methods since the abundance is estimated by involving the structural information. Next, the parameters'

Table 1 Algorithm I
表 1 算法 I

Algorithm 1: The proposed ADSpLRU-SI

Input: $\boldsymbol{Y}, \boldsymbol{\theta}$

Set:

$t=0$; $\boldsymbol{M}; \boldsymbol{F}=(\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\theta}+3\boldsymbol{I}_N)^{-1}; \boldsymbol{A}=-\kappa\boldsymbol{F}\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\theta}; \boldsymbol{B}=\boldsymbol{M}^{\mathrm{T}}\boldsymbol{M}$;

Initialize: $\boldsymbol{W}^0, \boldsymbol{\Sigma}_1^0, \boldsymbol{\Sigma}_2^0, \boldsymbol{\Sigma}_3^0, \boldsymbol{\Sigma}_4^0, \boldsymbol{\Sigma}_5^0, \boldsymbol{Z}_1^0, \boldsymbol{Z}_2^0, \boldsymbol{Z}_3^0, \boldsymbol{Z}_4^0, \boldsymbol{Z}_5^0$;

Repeat:

$\boldsymbol{J}^{\mathrm{t}}=\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\Sigma}_1^{\mathrm{t}}+\boldsymbol{\Sigma}_2^{\mathrm{t}}+\boldsymbol{\Sigma}_3^{\mathrm{t}}+\boldsymbol{\Sigma}_4^{\mathrm{t}}+\kappa\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\Sigma}_5^{\mathrm{t}}\boldsymbol{M}^{\mathrm{t}}$;

$\boldsymbol{H}^{\mathrm{t}}=\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{Z}_1^{\mathrm{t}}+\boldsymbol{Z}_2^{\mathrm{t}}+\boldsymbol{Z}_3^{\mathrm{t}}+\boldsymbol{Z}_4^{\mathrm{t}}+\kappa\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{Z}_5^{\mathrm{t}}\boldsymbol{M}^{\mathrm{t}}$;

$\boldsymbol{F}=(\boldsymbol{\theta}^{\mathrm{T}}\boldsymbol{\theta}+3\boldsymbol{I}_N)^{-1}$;

$\boldsymbol{C}^{\mathrm{t}}=\boldsymbol{F}(\boldsymbol{H}^{\mathrm{t}}+\boldsymbol{J}^{\mathrm{t}}); \boldsymbol{W}^{\mathrm{t}+1}=\mathrm{dlyap}(\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}^{\mathrm{T}})$;

$\boldsymbol{\Sigma}_1^{\mathrm{t}+1}=\frac{1}{1+\mu}(\boldsymbol{Y}+\mu\boldsymbol{\theta}\boldsymbol{W}^{\mathrm{t}+1}-\boldsymbol{Z}_1^{\mathrm{t}}); \boldsymbol{\Sigma}_5^{\mathrm{t}+1}=\frac{1}{1+\mu}(\boldsymbol{Y}+\mu\boldsymbol{\theta}\boldsymbol{W}^{\mathrm{t}+1}-\boldsymbol{Z}_5^{\mathrm{t}})$;

$\boldsymbol{\Sigma}_2^{\mathrm{t}+1}=\boldsymbol{SHR}_{\gamma}(\mu\boldsymbol{W}^{\mathrm{t}+1}-\boldsymbol{Z}_2^{\mathrm{t}}); \boldsymbol{\Sigma}_3^{\mathrm{t}+1}=\boldsymbol{SVT}_{\tau}(\mu\boldsymbol{W}^{\mathrm{t}+1}-\boldsymbol{Z}_3^{\mathrm{t}})$;

$\boldsymbol{\Sigma}_4^{\mathrm{t}+1}=\pi_{R+}(\mu\boldsymbol{W}^{\mathrm{t}+1}-\boldsymbol{Z}_4^{\mathrm{t}}); \boldsymbol{\Sigma}_5^{\mathrm{t}+1}=\boldsymbol{Z}_5^{\mathrm{t}}-\kappa(\mu\boldsymbol{\theta}\boldsymbol{W}^{\mathrm{t}+1}\boldsymbol{M}+\boldsymbol{\Sigma}_5^{\mathrm{t}+1})$;

$\boldsymbol{\Sigma}_{\mathrm{i}}^{\mathrm{t}+1}=\boldsymbol{Z}_{\mathrm{i}}^{\mathrm{t}}-\mu\boldsymbol{\theta}\boldsymbol{W}^{\mathrm{t}+1}+\boldsymbol{\Sigma}_{\mathrm{i}}^{\mathrm{t}+1})\ (i=1,2,3,4)$

Until: convergence

Output: $\widehat{\boldsymbol{W}}=\boldsymbol{W}^{\mathrm{t}+1}$

setting and performance evaluation criteria are given. To unveil the advantages of the work over other methods, four different types of experiments on synthetic data are executed. Meanwhile, one experimental test is performed on real hyperspectral image. Since Ref. [37] has demonstrated that ADSpLRU prevails over CSUnSAL[35] and Bi-ICE[33] methods, we only compare the proposed method with ADSpLRU on the rest of experiments.

Table 2 Running time comparison
表 2 运行时间比较

Algorithm	CSUnSAL	ADSpLRU	ADSpLRUsI	Bi-ICE
Costing time	0.12 ns	0.06 ns	0.2 ns	0.42 ns

3.1 The structural matrix for extracting multiple-structure information

To extract structural information, the structural matrix is defined as $\boldsymbol{M}_s=[\boldsymbol{m}_1,\boldsymbol{m}_2,\cdots,\boldsymbol{m}_K]\in\boldsymbol{R}^{K\times K}$, where $\boldsymbol{m}_i$ is a column vector and composed of -1 and 1. The matrix $\boldsymbol{M}_s$, the so called "single structure matrix (SSM)", can extract the structure information of the original image. To illustrate it clearly, we give an example in Fig. 2(a). In Fig. 2(a), $\boldsymbol{I}$ is an image vector (obtained from converting the $K=2\times2$ image to a vector), $\boldsymbol{M}_s=[\boldsymbol{m}_1,\boldsymbol{m}_2,\cdots,\boldsymbol{m}_4]$ is a single-structured matrix (SSM), where $\boldsymbol{m}_1=[0,0,1,-1]^{\mathrm{T}}$, $\boldsymbol{m}_2=[0,1,-1,0]^{\mathrm{T}}$, $\boldsymbol{m}_3=[1,0,-1,0]^{\mathrm{T}}$, $\boldsymbol{m}_4=[0,-1,0,1]^{\mathrm{T}}$. The structural information (SI) is obtained by $\boldsymbol{I}\times\boldsymbol{M}_s$. Figure 2(a) elaborates that the elements in $\boldsymbol{SI}$ denote the distance between any pair of pixels in image $\boldsymbol{I}$. And we find that the number of column-vector could reach $2^4=16$. Meanwhile, each single-structure matrix (SSM) is composed of four column-vectors. It is obviously found that there are various SSMs which is composed of -1 and $+1$. In order to use these various structural information, a number of SSMs are summed up to form a new matrix called a multi-structured matrix (MSM). As shown in Fig. 2(b), the multi-structured information is defined as follows:

$$\boldsymbol{SI}=\frac{1}{q}\times(\boldsymbol{SI}_1+\boldsymbol{SI}_2+\cdots+\boldsymbol{SI}_q)=\boldsymbol{I}\times(\Sigma_{i=1}^{q}\boldsymbol{M}_i)/q \quad (13)$$

Thus, it is quite clear that the MSM denoted by matrix $\boldsymbol{M}$ is formulated as

$$\boldsymbol{M}=(\Sigma_{i=1}^{q}\boldsymbol{M}_i)/q \quad (14)$$

where q denotes the number of structural matrices.

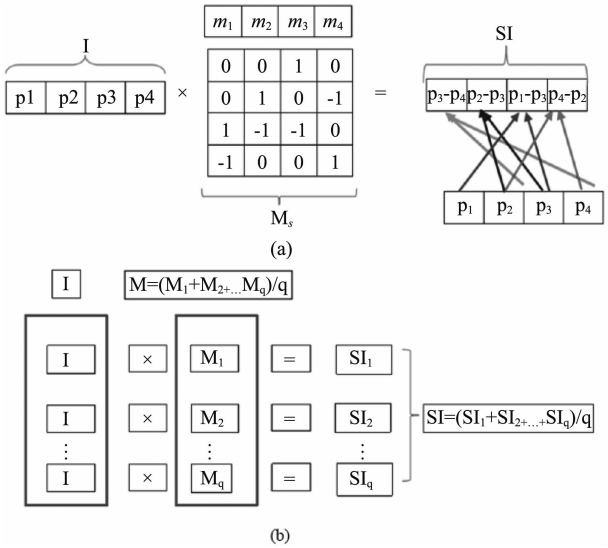

Fig. 2 Structural matrix. (a) Single-structure matrix $\boldsymbol{M}$ (b) Multi-structured matrix $\boldsymbol{M}$
图 2 结构矩阵,(a) 单结构矩阵 (b)多结构矩阵

In this article, $K=3\times3$ is fixed, the proposed method with structural information is termed as ADSpLRU-SI. We use ADSpLRU-SI-q means that ADSpLRU-SI uses the matrix $\boldsymbol{M}$ composed of q SSMs.

3.2 Setting of parameters and performance evaluation criteria

In the matter of parameters, γ denotes the sparsity parameter in all examined algorithms which controls the convergent behavior of ADSpLRU-SI and ADSpLRU, takes positive values. In all our experiments, $\mu=0.01$ is fixed. The low-rank parameter is represented by τ. And κ manages the structural information. Parameters γ, τ, and κ are fined tuned with 10 different values, as shown in Tabel 3.

Table 3 Parameters Setting
表 3 模型参数

Algorithm	γ	τ	μ	κ ·
ADSpLRU	0,10^[−10,−9,⋯,−2]	0,10^[−10,−9,⋯,−2]	0.01	Not available
ADSpLRUSI	0,10^[−10,−9,⋯,−2]	0,10^[−10,−9,⋯,−2]	0.01	0,10^[−10,−9,⋯,−2]

To evaluate the performance of the proposed algorithms and the competing one, the signal-to-reconstruction error (SRE) is adopted to measure the power between the signal and error[37], which is defined as follows:

$$\mathrm{SRE} = 10\log_{10}\left(\frac{\sum_{i=1}^{n}\|\boldsymbol{w}_i\|_2^2}{\sum_{i=1}^{n}\|\boldsymbol{w}_i - \boldsymbol{w}'_i\|_2^2}\right), \quad (15)$$

where n stands for the total number of the pixels, $\boldsymbol{w}_i$ is the actual abundance vector of the *i-th* pixel, and $\boldsymbol{w}_i'$ is the estimated one.

SRE is used instead of the classical root mean square error (RMSE) because Ref. [53] has provided the suggestion that SRE gives more information regarding the power of the error in relation with the power of the signal. The higher the SRE (dB), the better the unmixing performance[53].

3.3 Experiments on simulated data

For the synthetic data experiments, we use the spectral library randomly selected from the United States Geological Survey (USGS) digital spectral library[54], which has 224 spectral bands uniformly ranging from 0.4 ~2.5 μm, and contains 498 different spectral signatures of endmembers. In this work, the endmembers' dictionary $\boldsymbol{\theta}$ is constructed by selecting randomly the N endmembers from the USGS library. Then, given $\boldsymbol{\theta}$ *and actual abundance matrix* $\boldsymbol{W}$ and noise $\boldsymbol{E}$, the spectra are generated by utilizing the LMM in Eq. (1).

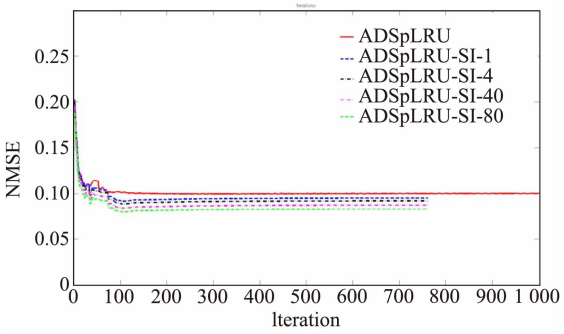

Fig. 3　Convergent behavior
图 3　收敛性

1) Convergent behavior of ADSpLRU-SI and other competing algorithms. In order to reveal the convergent behavior of the proposed algorithm, firstly, a rank 2 and sparsity level 20% abundance matrix $\boldsymbol{W} \in \boldsymbol{R}^{N\times K}$ (N = 50 endmembers, K = 9 pixels) is generated. Secondly, spectra $\boldsymbol{Y} \in \boldsymbol{R}^{L\times K}$ (L = 224 bands, K = 9 pixels) is derived by adding Gaussian noise(e. g. SNR = 28 dB) to the LMM in Eq. (1). Thirdly, the abundance matrix is estimated by four different algorithms namely ADSpLRU and ADSpLRU-SI-1 and ADSpLRU-SI-4 and ADSpLRU-SI-40 and ADSpLRU-SI-80. To compare the behavior with the method in Ref. [37], normalized mean squared estimation error (NMSE) defined in Ref. [37] is used to evaluate their convergent property.

$$\mathrm{NMSE}(t) = \frac{1}{p}\sum_{i=1}^{p}\frac{\|\boldsymbol{W}_i^t - \boldsymbol{W}_i\|_F^2}{\|\boldsymbol{W}_i\|_F^2}, \quad (16)$$

where $\boldsymbol{W}_i$ is the actual matrix of the *i-th* realizations, $\boldsymbol{W}_i^t$ is the estimated matrix of the $\boldsymbol{W}_i$ at the *t-th* iteration. In nature, the NMSE is equivalent to RMSE. Namely, if $\boldsymbol{W}_i$ is constant, the NMSE equals to RMSE.

In this work, a number of 10 independent realizations is run for each experiment and the average NMSE is used to measure the convergence. The convergent property is illustrated in Fig. 3 clearly: ADSpLRU-SI-1 achieves slightly better estimation accuracy than ADSpLRU in terms of NMSE. It results from that only one structure is employed by ADSpLRU-SI-1. Next, ADSpLRU-SI-4 reaches a smaller value than ADSpLRU-SI-1 and ADSpLRU in terms of NMSE. It stems from that ADSpLRU-SI-4 uses four structures. Further, the ADSpLRU-SI-40 could obtain smaller NMSE than ADSpLRU and ADSpLRU-SI-1 and ADSpLRU-SI-4. Moreover, ADSpLRU-SI-80 shows higher estimation accuracy than other methods in terms of NMSE. It illustrates that the greater the amount of structural information, the better the convergence. It is noteworthy that the better convergence of ADSpLRU-SI is at the cost of much more time than the ADSpLRU algorithm.

2) A toy example: To highlights the significance of the structural information, we take the row-difference structures and column-difference structures for example. The procedure is designed as following steps: firstly, a $N \times K$ (where N = 50 and K = 9) abundance W with rank 2 and sparsity level 20% is generated and illustrated in Fig. 4(a). And Fig. 4(b) illustrated a ground-truth abundance $\boldsymbol{W}$ with rank 2 and sparsity level 10%. The $L \times K$ spectra $\boldsymbol{Y}$ is obtained as the LMM in Eq. (1). Then, the abundance matrix (sparsity level 20%) estimated by ADSpLRU is shown in Fig. 4(c). And the abundance (sparsity level 20%) resulted from ADSpLRU-SI is illustrated in Fig. 4(e). Similarly, Fig. 4(d) illustrates the estimated abundance matrix (sparsity level 10%) obtained by ADSpLRU. Figure 4(f) displays the abundance (sparsity level 10%) estimated by ADSpLRU-SI. Looking into the Fig. 4, we could observe that most parts of the estimated abundance from ADSpLRU-SI and ADSpLRU share the common abundance values except for the circled area. The circled area in abundance maps displays that the abundance matrix estimated by ADSpLRU-SI is closer to the actual abundance than the abundance estimated from ADSpLRU method. In conclusion, as a result of the structural information, ADSpLRU-SI improve the performance in sparse low-rank abundance estimation.

3) The key role of the parameter κ: The parameters γ, τ control the imposition of sparsity and low-rankness which has been discussed in Ref. [37]. This experiment mainly explores the parameter κ which control the structural information as constraint counterparts. To unveil dependency of the optimal set of κ values, row-difference map as the structural information is used to experiment. Nine types of abundance matrices reflect different row-difference information of the matrix with rank 2 and sparsity level 10%. A number of 10 independent realizations is run for each of the nine experiments, and the average NMSE is demonstrated as function of κ. As shown in Fig. 5, with the amplitudes of the row-difference map increasing, the parameter κ could play more effective influence on the estimation performance. On the

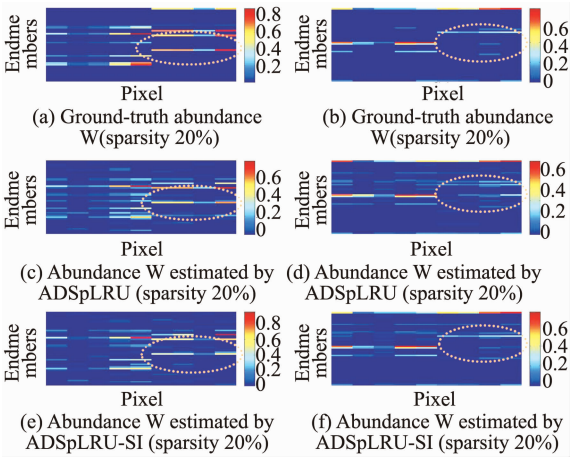

Fig. 4 Toy experiments, (a) Ground-truth abundance $\boldsymbol{W}$ (sparsity 20%), (b) Ground-truth abundance $\boldsymbol{W}$ (sparsity 10%), (c) Abundance $\boldsymbol{W}$ estimated by ADSpLRU (sparsity 20%), (d) Abundance $\boldsymbol{W}$ estimated by ADSpLRU (sparsity 10%), (e) Abundance $\boldsymbol{W}$ estimated by ADSpLRU-SI (sparsity 20%), (f) Abundance $\boldsymbol{W}$ estimated by ADSpLRU-SI (sparsity 10%)

图 4 仿真结果,(a) 稀疏度为 20% 的真值 $\boldsymbol{W}$,(b) 稀疏度为 10% 的真值 $\boldsymbol{W}$,(c) 由 ADSpLRU 算法估计(a)的结果,(d) 由 ADSpLRU 估计(b)的结果,(e) 由 ADSpLRU-SI 估计(a)结果,(f) 由 ADSpLRU-SI 估计(b)的结果

contrary, if the row-difference amplitude is too small, it is easy for the κ playing a negative impact on the estimation performance. Considering the equations in algorithm 1, at the condition of $\boldsymbol{\theta M}$ becoming smaller, the $\boldsymbol{\theta M}$ plays less influence on the abundance matrix estimation. And the κ could make the solution escaping away from the convergent area easily. In conclusion, provided the amplitude of the row-difference map in the abundance matrix is not too small, the ADSpLRU-SI could achieve better performance than ADSpLRU.

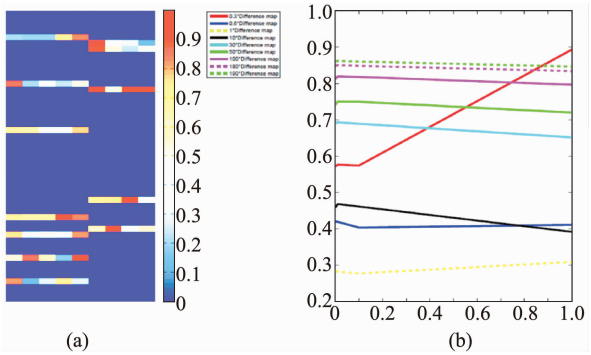

Fig. 5 Experiments of the parameter κ, (a) initial abundance matrix, (b) NMSE function of the parameter κ

图 5 关于参数 κ 的实验,(a) 初始丰度矩阵 (b) 对不同行差幅度的 W 进行估计,其估计性能(NMSE)随参数 κ 的变化曲线

4) Robustness to noise: In this experiment we aim at exhibiting the robustness of the proposed algorithms to white and correlated noise corruption. To this end, we stick with a specific simultaneously sparse and low-rank abundance matrix $\boldsymbol{W}$ of sparsity level 20% and rank 3. Based on this $\boldsymbol{W}$, $K=9$ linearly mixed pixels are generated, in the same way as described above. Then, depending on the case, white or colored Gaussian noise contaminates the data. 16 SNR values are considered ranging from 10 to 40 dB, while 10 realizations are run for each SNR value, and the mean of SRE metrics is calculated.

White Gaussian Noise: The SRE resulting from the proposed ADSpLRU-SI (ADSpLRU-SI-1, ADSpLRU-SI-4, ADSpLRU-SI-40, ADSpLRU-SI -80) and the ADSpLRU are shown in the Fig. 6(a). In all the examined SNR values, ADSpLRU-SI-80 obtains the most accurate results comparing with ADSpLRU. Additionally, the ADSpLRU-SI with a larger number of single structural matrices performs better than ADSpLRU. In conclusion, the proposed methods are more robust to the majority of white noise than others.

. *Colored Gaussian Noise*: Actually, in real hyperspectral images the noise that corrupts the data is rather structured than white. Thus, to assess the behavior of the proposed methods in such realistic conditions, we simulate correlated Gaussian noise that adds up to the linearly mixed pixels. Figure 6 (b) illustrates the effectiveness of the tested algorithms in terms of SRE, for different SNR values. It is clearly seen that ADSpLRU-SI-80 achieves superior results to other algorithms in all the examined SNR values. Additionally, it is noted that ADSpLRU-SI-1 performs slightly better than ADSpLRU. As a result, the robustness of our proposed methods is also corroborated in the presence of correlated noise with different magnitude.

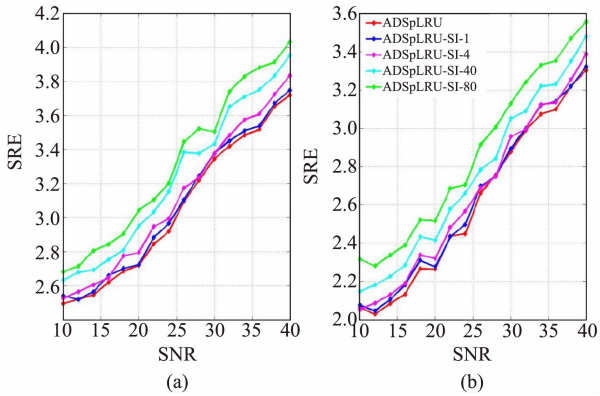

Fig. 6 Robustness, (a) white noise, (b) colored noise

图 6 抗噪声,(a)白噪声,(b) 彩色噪声

3.4 Experiment on real data

This section illustrates the performance of the proposed algorithms when applied on a real hyperspectral image. The hyperspectral scene under examination is from the Urban datasets[55]. There are 307 × 307 pixels, each of which corresponds to a 3 × 3 m^2 area. In this image, there are 210 wavelengths ranging from 400 nm to 2 500 nm, resulting in a spectral resolution of 10 nm. After the channels 1 ~ 4, 76, 87, 101 ~ 111, 136 ~ 153 and 198 ~ 210 are removed (due to dense water vapor and atmospheric effects), we remain 162 channels (this

is a common preprocess for hyperspectral unmixing analyses)[55]. This scene consists of six different species, namely Asphalt, Grass, Tree, Roof, Metal, Dirt. In order to assess the effectiveness of the proposed algorithms, the experiment is carried out on the 50×50 sub-window in hyperspectral image.

The results are shown in Fig. 7. Figure 7(a) illustrates the ground-truth abundance of roof component in the sub-window. Figure 7(b) shows the estimated abundance which is obtained by the ADSpLRU method. The abundance which is estimated by ADSpLRU-SI algorithm with one structural information (ADSpLRU-SI-1) is displayed in Fig. 7(c). Figure 7(d) illustrates the abundance estimated by ADSpLRU-SI-4. The abundances derived from ADSpLRU-SI-40 and ADSpLRU-SI-80 are shown in Fig. 7(e) and (f) respectively. From a careful visual inspection of the generated maps, the estimated abundances from the methods share many common values, except for the ellipse area. Focusing on the ellipse region, we can see that the abundances obtained ADSpLRU-SI-1and ADSpLRU-SI-4 present patterns that are slightly closer to the ground-truth abundance in Fig. 7(a) than abundance obtained by ADSpLRU. The little effectiveness results from a small number of structural information which is used by the ADSpLRU-SI-1 and ADSpLRU-SI-4. Looking into Fig. 7(e), we find that the abundance obtained by ADSpLRU-SI-40 is much closer to the ground-truth data than by these methods (ADSpLRU, ADSpLRU-SI-1 and ADSpLRU-SI-4).

Compared with ADSpLRU-SI-40, the ADSpLRU-SI-80 obtains the better performance as shown in Fig. 7(f). This is due to the fact that 80-strucured information brings more effectiveness on abundance estimation than 40-strutured information. Figure 8 shows the SRE(dB) results achieved by the above methods. It suggests that the proposed algorithm ADSpLRU-SI achieves higher estimation accuracy than ADSpLRU.

In conclusion, ADSpLRU-SI could obtain more accurate estimation abundance than ADSpLRU. Further, the above experiments have unveiled the dependency that a large amount of SSMs applied in ADSpLRU-SI could achieve higher estimation accuracy than a small amount of SSMs employed in ADSpLRU-SI.

4　Conclusion and future direction

In this paper, a novel linear constrained regression model is put forward by adding the multi-structured information to the traditional linear regression model. Moreover, the mathematical theory is employed to prove that the new model can reach more favorable accuracy than the old one. This proof suggests that such methods could be popularized and applied to other related fields. The novel model is used to estimate the sparse and low-rank abundance matrix for spectral unmixing. However, when the novel model is solved by the ADMM techniques, the structural information brings challenges to handling the optimization with respect to abundance matrix. Faced with the challenges, the article discovers that sylvester equation opens up a way to easily solve the abundance matrix. In the field of the sparse low-rank abundance matrix estimation, the proposed algorithm is termed as the alternating direction sparse and low-rank unmixing algorithm with structural information (ADSpLRU-SI). The comparisons between ADSpLRU-SI and ADSpLRU indicate that the proposed algorithm by adding a large number of multi-structured information, can improve the regression performance, keep more structural information, improve the accuracy of abundance estimation, speed up

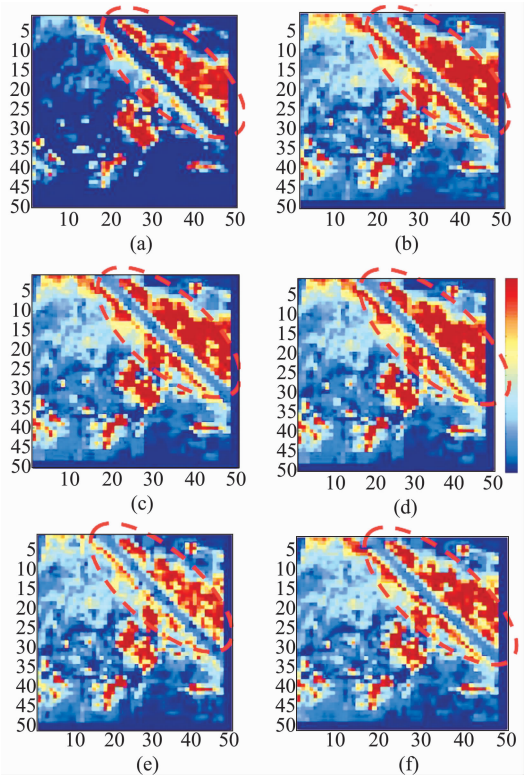

Fig. 7　Abundance map for roof. (a) Original Abundance, (b) Abundance obtained from ADSpLRU, (c) Abundance obtained from ADSpLRU-SI-1, (d) Abundance obtained from ADSpLRU-SI-4, (e) Abundance obtained from ADSpLRU-SI-40, (f) Abundance obtained from ADSpLRU-SI-80

图 7　草地区域的丰度图,(a)真值图,(b)ADSpLRU 算法估计的丰度图,(c)ADSpLRU-SI-1 算法估计的丰度图,(d)ADSpLRU-SI-4 算法估计的丰度图,(e)ADSpLRU-SI-40 算法估计的丰度图,(f)ADSpLRU-SI-80 算法估计的丰度图

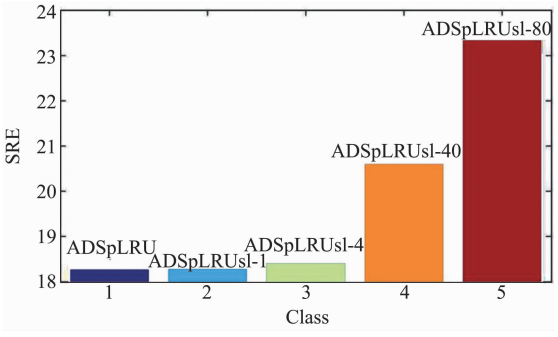

Fig. 8　SRE For different estimated Abundance

图 8　不同算法的 SRE 评测指标

the convergent steps and enhance the robustness to noise. Since the multi-structured matrix plays an important role in unmixing, how to search the best structural matrix will be the future research work. Moreover, since the performance is influenced easily by parameters which control sparse and low-rank and structural information terms, the future researches would be carried out on the adaptive parameters selection.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (NNSF: 61673234 and NNSF: U1636124). The authors wish to thank the anonymous reviewers for their constructive comments that helped improve the scholarly quality of the paper. And the authors thank Paris V. Giampouras who offered some materials, which supported a part of experiments in the paper.

References

[1] Keshava N, Mustard J F. Spectral unmixing[J]. *IEEE Signal Processing Magazine*, 2002, **19**(1):44-57.

[2] Smith M O, Johnson P E, Adams J B. Quantitative determination of mineral types and abundances from reflectance spectra using principal components analysis[C]. *Lunar and Planetary Science Conference Proceedings*, 1985:C797-C804.

[3] Adams J B, Smith M O, Johnson P E. Spectral mixture modeling: a new analysis of rock and soil types at the Viking Lander I site [J]. *Journal of Geophysical Research Atmospheres*, 1986, **91**(B8): 8098-8112.

[4] Gillespie A R, Smith M O, Adams J B, *et al*, Interpretation of residual images: Spectral mixture analysis of AVIRIS images[C]. *Proceedings of the 2nd AVIRIS Workshop*, 1990:243-270.

[5] Vane G, Green R O, Chrien T G, *et al*. The airborne/infrared imaging spectrometer (AVIRIS)[J], *Remote Sensing of Environment*, 1993, **44**:127-143.

[6] Swayze G, Clark R, Sutley S, *et al*. Ground-truthing AVIRIS mineral mapping at Cuprite, Nevada[C]// *Summaries of the Third Jpl Airborne Geosciences Workshop*. 1992.

[7] Green R O, Eastwood M L, Sarture C M, *et al*. Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS)[J]. *Remote Sensing of Environment*, 1998, **65**(3):227-248.

[8] Shaw G, Manolakis D. Signal processing for hyperspectral image exploitation[J]. *IEEE Signal Processing Magazine*, 2002, **19**(1):12-16.

[9] Landgrebe D. Hyperspectral image data analysis[J]. *IEEE Signal Processing Magazine*, 2002, **19**(1):17-28.

[10] Manolakis D, Shaw G. Detection algorithms for hyperspectral imaging applications[J]. *IEEE Signal Processing Magazine*, 2002, **19**(1):29-43.

[11] Stein D W J, Beaven S G, Hoff L E, *et al*. Anomaly detection from hyperspectral imagery[J]. *IEEE Signal Processing Magazine*, 2002, **19**(1):58-69.

[12] Lu G, Qin X, Wang D, *et al*. *Estimation of tissue optical parameters with hyperspectral imaging and spectral unmixing*[M]. 2015.

[13] Matsuki T, Yokoya N, Iwasaki A. Hyperspectral tree species classification of Japanese complex mixed forest with the aid of lidar data[J]. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, 2015, **8**(5):2177-2187.

[14] Brook A, Dor E B. Quantitative detection of settled dust over green canopy using sparse unmixing of airborne hyperspectral data[J]. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, 2016, **9**(2):884-897.

[15] Alam M S, Sidike P. Trends in oil spill detection via hyperspectral imaging[C]// *IEEE International Conference on Electrical & Computer Engineering*, 2013:858-862.

[16] Lin H, Zhang X. Retrieving the hydrous minerals on Mars by sparse unmixing and the Hapke model using MRO/CRISM data[J]. *Icarus*, 2017, **288**:160-171.

[17] Baskurt D O, Omruuzun F, Cetin Y Y. Hyperspectral unmixing based analysis of forested areas[C]// *IEEE Signal Processing and Communications Applications Conference*. 2015:2329-2332.

[18] Das B S, Sarathjith M C, Santra P, *et al*. Hyperspectral remote sensing: opportunities, status and challenges for rapid soil assessment in India[J]. *Current Science*, 2015, **108**(5):860-868.

[19] Ji C C, Jia Y H, Li X S, *et al*. Estimation of vegetation coverage of Nitraria tangutorum by linear / non-linear spectral mixing model[J]. *Journal of Remote Sensing*, 2016. **20**(6): 1402-1412.

[20] Zhang B, Shen Q, Junsheng L I, *et al*. Retrieval of three kinds of representative water quality parameters of Lake Taihu from hyperspectral remote sensing data[J]. *Journal of Lake Sciences*, 2009, **21**(2): 182-192.

[21] Rosin P L. Robust pixel unmixing[J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2001, **39**(9):1978-1983.

[22] Bioucas-Dias J M, Plaza A, Dobigeon N, *et al*. Hyperspectral unmixing overview: Geometrical, statistical, and sparse regression-based approaches[J]. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, 2012, **5**(2):354-379.

[23] Themelis K E, Rontogiannis A A, Koutroumbas K D. A novel hierarchical Bayesian approach for sparse semisupervised hyperspectral unmixing[J]. *IEEE Transactions on Signal Processing*, 2012, **60**(2): 585-599.

[24] Feng R, Zhong Y, Zhang L. Adaptive spatial regularization sparse unmixing strategy based on joint MAP for hyperspectral remote sensing imagery[J]. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, 2016, **9**(12):5791-5805.

[25] Iordache M D, Bioucas-Dias J, Plaza A. Unmixing sparse hyperspectral mixtures[C]// *IEEE International Geoscience and Remote Sensing*, 2009:IV-85-IV-88.

[26] Candès E J, Romberg J, Tao T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information[J]. *IEEE Transactions on Information Theory*, 2006, **52**(2):489-509.

[27] Iordache M D, Bioucas-Dias J M, Plaza A. Sparse unmixing of hyperspectral data[J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2011, **49**(6):2014-2039.

[28] Chen F, Zhang Y. Sparse hyperspectral unmixing based on constrained $\ell_p - \ell_2$ optimization[J]. *IEEE Geoscience & Remote Sensing Letters*, 2013, **10**(5):1142-1146.

[29] Tang W, Shi Z, Duren Z. Sparse hyperspectral unmixing using an approximate ℓ_0 norm[J]. *Optik - International Journal for Light and Electron Optics*, 2014, **125**(1):31-38.

[30] Sun L, Jeon B, Zheng Y, *et al*. Hyperspectral unmixing based on $\ell_2 - \ell_1$ sparsity and total variation[C]// *IEEE International Conference on Image Processing. IEEE*, 2016:4349-4353.

[31] He W, Zhang H, Zhang L. Hyperspectral unmixing using total variation regularized reweighted sparse non-negative matrix factorization [C]// *Geoscience and Remote Sensing Symposium. IEEE*, 2016:7034-7037.

[32] Themelis K E, Rontogiannis A A, Koutroumbas K D. A Novel hierarchical Bayesian approach for sparse semisupervised hyperspectral unmixing[J]. *IEEE Transactions on Signal Processing*, 2012, **60**(2): 585-599.

[33] Rontogiannis A, Themelis K, Sykioti O, *et al*. A fast variational Bayes algorithm for sparse semi-supervised unmixing of OMEGA/Mars Express data[C]// *IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, 2013.

[34] Zhao Y, Yang J. Hyperspectral image denoising via sparsity and low rank[C]// *Geoscience and Remote Sensing Symposium. IEEE*, 2014: 1091-1094.

[35] Iordache M D, Bioucas-Dias J M, Plaza A. Collaborative sparse regression for hyperspectral unmixing[J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2014, **52**(1):341 - 354.
[36] Qu Q, Nasrabadi N M, Tran T D. Abundance estimation for bilinear mixture models via joint sparse and low-rank representation[J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2014, **52**(7):4404 - 4423.
[37] Giampouras P V, Themelis K E, Rontogiannis A A, *et al*. Simultaneously sparse and low-rank abundance matrix estimation for hyperspectral image unmixing[J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2016, **54**(8):4775 - 4789.
[38] Richard E, Savalle P A, Vayatis N. Estimation of simultaneously sparse and low rank matrices[J]. *Computer Science*, 2012.
[39] Oymak S, Jalali A, Fazel M, *et al*. Simultaneously structured models with application to sparse and low-rank matrices[J]. *IEEE Transactions on Information Theory*, 2012, **61**(5):2886 - 2908.
[40] Golbabaee M, Vandergheynst P. Compressed sensing of simultaneous low-rank and joint-sparse matrices[J]. *Mathematics*, 2012, 1 - 32.
[41] Richard E, Obozinski G, Vert J P. Tight convex relaxations for sparse matrix factorization[J]. *Advances in Neural Information Processing Systems*, 2014, **4**:3284 - 3292.
[42] Chen L S, Huang J H Z. Sparse reduced - rank regression for simultaneous dimension reduction and variable selection[J]. *Journal of the American Statistical Association*, 2012, **107**(500):1533 - 1545.
[43] Bach F R. Consistency of trace norm minimization[J]. *Journal of Machine Learning Research*, 2008, **9**(2):2008.
[44] Bioucas-Dias J M, Figueiredo M A T. Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing[C]// *The Workshop on Hyperspectral Image & Signal Processing: Evolution in Remote Sensing. IEEE*, 2012:1 - 4.
[45] Negahban S, Wainwright M J. Estimation of (near) low-rank matrices with noise and high-dimensional scaling[J]. *Annals of Statistics*, 2011, **39**(2):1069 - 1097.
[46] Babacan S D, Luessi M, Molina R, *et al*. Sparse Bayesian methods for low-rank matrix estimation[J]. *IEEE Transactions on Signal Processing*, 2011, **60**(8):3964 - 3977.
[47] Boyd S, Parikh N, Chu E, *et al*. Distributed optimization and statistical learning via the alternating direction method of multipliers[J]. *Foundations & Trends in Machine Learning*, 2010, **3**(1):1 - 122.
[48] https://en. wikipedia. org/wiki/Structure. "structure, n.". Oxford English Dictionary (Online ed.). Retrieved 1 October 2015.
[49] Yuan J, Zhang Y J. Application of sparse denoising auto encoder network with gradient difference information for abnormal action detection [J]. *Journal of Automatica Sinica*, 2017, **43**(4):604 - 610.
[50] Boyd S, Vandenberghe L. *Convex optimization*[M]. Cambridge University Press, 2004.
[51] Zhang X D, *Matrix analysis and applications*[M]. Beijing: Tsinghua university press, 2004:44.
[52] Lu C S. Solution of the matrix equation AX + XB = C[J]. *Electronics Letters*, 1986, **37**(4):351 - 355.
[53] Iordache M D, Bioucas-Dias J M, Plaza A. Total variation spatial regularization for sparse hyperspectral unmixing[J]. *IEEE Transactions on Geoscience & Remote Sensing*, 2012, **50**(11):4484 - 4502.
[54] Clark R N, Swayze G A, Wise R, et al. U. S Geological Survey Library. splib06b[EB/OL]. 2007, http://speclab. cr. usgs. gov/spectral. lib06b.
[55] Zhu F, Wang Y, Fan B, *et al*. Effective spectral unmixing via robust representation and learning-based sparsity[EB/OL]. 2014, http://arxiv. org/abs/1409. 0685.

(上接第 143 页)

[10] Apiratikul P, He L, Richardson C J K. 2μm laterally coupled distributed-feedback GaSb-based metamorphic laser grown on a GaAs substrate [J]. *Applied Physics Letters*, 2013, **102**(23):031107.
[11] Telkkälä J, Karinen J, Viheriälä J, *et al*. Narrow-linewidth distributed feedback lasers with laterally coupled ridge-waveguide surface gratings fabricated using nanoimprint lithography [J]. *Semiconductor Conference. IEEE Xplore*, 2010:131 - 141.
[12] Dridi K, Benhsaien A, Hall T, *et al*. 1.55μm laterally coupled ridge-waveguide DFB lasers with third-order surface grating [J]. *Proceedings of SPIE-The International Society for Optical Engineering*, 2012, **8412**:84121R - 84121R - 6.
[13] Rößner K, Hümmer M, Benkert A, *et al*. Long-wavelength GaInAsSb/AlGaAsSb DFB lasers emitting near 2.6 μm [J]. *Physica E: Low-dimensional Systems and Nanostructures*, 2005, **30**(1 - 2):159 - 163.
[14] Martin R D, Forouhar S, Keo S, *et al*. InGaAs-GaAs-AlGaAs laterally-coupled distributed feedback (LC-DFB) ridge laser diode [J]. *Electronics Letters*, 1994, **30**(13):1058 - 1060.
[15] Chen T R, Hsin W, Chen B, *et al*. Wide temperature range operation of DFB Lasers at 1310 and 1490nm [J]. *Semiconductor Laser Conference*, 2008. *Islc* 2008. *IEEE*, *International. IEEE*, 2008:81 - 82.
[16] LIAO Yong-Ping, ZHANG Yu, YANG Cheng-Ao, *et al*. High-power, high-efficient GaSb-based quantum well laser diodes emitting at 2 μm [J]. *Journal of Infrared and Millimeter Waves* (廖永平, 张宇, 杨成奥,等. 大功率高效率2μm 锑化镓基量子阱激光器. **红外与毫米波学报**), 2016, **35**(6):672 - 675.
[17] Briggs R M, Frez C, Ksendzov A, *et al*. Laterally coupled distributed-feedback GaSb-based diode lasers for atmospheric Gas detection at 2 μm [J]. *Lasers and Electro-Optics. IEEE*, 2012:1 - 2.
[18] Viheriälä J, Haring K, Suomalainen S, *et al*. High spectral purity high-power GaSb-based DFB laser fabricated by nanoimprint lithography [J]. *IEEE Photonics Technology Letters*, 2016, **28**(11):1233 - 1236.
[19] Jang S J, Yu J S, Lee Y T. Laterally coupled DFB lasers with self-aligned metal surface grating by holographic lithography [J]. *IEEE Photonics Technology Letters*, 2008, 20(7):514 - 516.
[20] Zhang Y, Wang G W, Tang B, *et al*. Molecular beam epitaxy growth of InGaSb/AlGaAsSb strained quantum well diode lasers [J]. *Journal of Semiconductors*, 2011, **32**(10):103002.
[21] Fan Z F, Luo J S, Ye W H. Compressible Rayleigh-Taylor instability with preheat in inertial confinement fusion[J]. *Chinese Physics Letters*, 2007, **24**(8):2308 - 2311.
[22] Deng L G, Rahman M, Berg J A V D, *et al*. Contribution of atomic and molecular ions to dry-etch damage [J]. *Applied Physics Letters*, 1999, **75**(2):211 - 213.
[23] Rahman M. Channeling and diffusion in dry-etch damage [J]. *Journal of Applied Physics*, 1997, **82**(5):2215 - 2224.
[24] Rahman M, Deng L G, Wilkinson C D W, *et al*. Studies of damage in low-power reactive-ion etching of III - V semiconductors[J]. *Journal of Applied Physics*, 2001, **89**(4):2096 - 2108.
[25] Streifer W, Scifres D R, Burnham R. Coupled wave analysis of DFB and DBR lasers [J]. *Quantum Electronics IEEE Journal of*, 1977, **13**(4):134 - 141.
[26] Yang C A, Zhang Y, Liao Y P, *et al*. 2-μm single longitudinal mode GaSb-based laterally coupled distributed feedback laser with regrowth-free shallow-etched gratings by interference lithography [J]. *Chinese Physics B*, 2016, **25**(2):181 - 185.
[27] Zhang Y G, Zheng Y L, Lin C, *et al*. Continuous wave performance and tunability of MBE grown 2.1μm InGaAsSb/AlGaAsSb MQW lasers[J]. *Chinese Physics Letters*, 2006, **23**(8):2262 - 2265.
[28] Reboul J R, Cerutti L, Rodriguez J B, *et al*. Continuous-wave operation above room temperature of GaSb-based laser diodes grown on Si [J]. *Applied Physics Letters*, 2011, **99**(12):511.
[29] Jallipalli A, Nunna K, Kutty M N, *et al*. Compensation of interfacial states located inside the "buffer-free" GaSb/GaAs (001) heterojunction via δ-doping [J]. *Applied Physics Letters*, 2009, **95**(7):683.