A Robust Multilinear Mixing Model with $l_{2,1}$ norm for Unmixing Hyperspectral Images

Minglei Li, Fei Zhu*, Alan J. X. Guo
Center for Applied Mathematics, Tianjin University, Tianjin, China.
{2017233036, fei.zhu, jiaxiang.guo}@tju.edu.cn

***Abstract*—The unmixing of hyperspectral data is a hot topic in the field of remote sensing. However, in presence of various types of noise, especially the noisy channels, the performance of unmixing approaches is seriously deteriorated. To enhance the robustness of the unmixing method is a subject worth studying. This paper presents a robust unmixing method based on the recently-proposed multilinear mixing model, where the $l_{2,1}$ norm is adopted in the loss function to suppress the influence of noise. The sparseness of abundance is also considered to improve the parameter estimation. The resulting optimization problem is solved by the alternating direction multiplier method (ADMM). Experiments on both synthetic and real images demonstrate the performance of the proposed unmixing strategy.**

***Index Terms*—hyperspectral image, nonlinear unmixing, alternating direction method of multipliers, robust, $l_{2,1}$ norm**

I. Introduction

A hyperspectral image (HSI) is different from the ordinary RGB images, as the former consists of measurements from up to hundreds of spectral channels across a certain wavelength range. Thus, each observed pixel corresponds to a spectral vector. The issue of HSI unmixing plays an important role in many fields related to HSI analysis, such as precision agriculture and mineral exploration [1]. It is assumed that each observed spectrum is mixed by multiple pure material signatures, termed endmembers. The task of unmixing is to extract the endmembers and to estimate their fractional abundances at each pixel.

Over the past decades, numerous unmixing models and related algorithms have been investigated [1]. The linear mixing model (LMM) is the most prevalent because of its simple mathematical modeling and straightforward physical meaning. The LMM supposes that each observed pixel can be expressed as a linear combination of endmember spectra. However, in many real scenarios, the nonlinear effect is serious, thus requiring more complex, nonlinear modeling of the mixing process. The bilinear mixing models assume each photon interacting with two endmembers before reaching the sensor, and are formulated as superposition of second-order terms between endmembers on the LMM [2], [3].

More recently, the multilinear mixing model (MLM) has been proposed in [4], and become a new research hotspot for nonlinear unmixing [5], [6]. The MLM accounts for all degrees of interactions between endmembers, and an additional parameter is introduced at each pixel to characterize the probability of further interactions. As this paper aims to improve the robustness of the MLM, a brief deviation of this model will be given in the subsequent section.

Different unmixing strategies, either the linear or nonlinear ones, suffer a lot from various types of noise in the data. In particular, there exist up to 20% corrupted channels of low signal-to-noise (SNR) ratio in real HSI [7]. To tackle this issue, efforts have been dedicated to enhance the robustness of different unmixing methods against the noise. A category of methods adopt an additional disturbance term to the model, in order to simulate the noise interference, such as [8]. Other methods consider to replace the commonly-used l_2 norm-based loss function by a more robust measure, including the $l_{2,1}$ norm [9] and the maximum correntropy criterion [7]. Of particular note is the use of the $l_{2,1}$ norm in different machine learning tasks to achieve the robustness against noise and outliers, including the PCA [11], the NMF [10] and the regression [12], to name a few.

As far as we know, no previous studies have been proposed to enhance the robustness of the MLM. This paper presents a robust unmxing method based on the MLM, termed R-MLM, mainly by taking advantage of the $l_{2,1}$ norm-based loss function. The resulting optimization problem is solved by the ADMM [13]. Experiments are performed on synthetic and and real datasets, to verify that the proposed R-MLM has strong robustness against noisy channels and outliers in HSIs.

II. Review of Multilinear Mixing Model

Let $X = [\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_T] \in \mathbb{R}^{L\times T}$ be a data matrix containing T pixels over L spectral bands, and $E = [\boldsymbol{e}_1, \boldsymbol{e}_2, ..., \boldsymbol{e}_N] \in \mathbb{R}^{L\times N}$ be the endmember matrix composed by N endmembers. The abundance matrix is denoted by $A = [\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_T] \in \mathbb{R}^{N\times T}$, with $\boldsymbol{a}_i$ being the abundance vector for the i-th pixel. Use $P = [p_1, p_2, ..., p_T] \in \mathbb{R}^{1\times T}$ to collect the probability parameters over pixels, where scalar p_j describes the probability of further interactions for the j-th pixel. The MLM [4] accounts for all degrees of interactions between endmembers, and obeys following assumptions: 1) the incoming light will interact with at least one material; 2) for j-th pixel, the probability of undergoing further interactions is p_j, and the probability of escaping the scene and reaching

* Corresponding author: Fei Zhu. The work was supported by the National Natural Science Foundation of China under Grant 61701337 and Grant 61671382, and the Natural Science Foundation of Tianjin City under Grant 18JCQNJC01600.

the sensor is $(1-p_j)$; 3) the probability of interacting with a material is proportional to its corresponding abundance; and 4) the intensity of the light scattered by a material depends on the corresponding material's albedo. As a consequence, the MLM model is formulated as

$$\begin{aligned}\boldsymbol{x}_j &= (1-p_j)\sum_{i=1}^{N}\boldsymbol{e}_i a_{ij} + (1-p_j)p_j\sum_{i,k=1}^{N}(\boldsymbol{e}_i\odot\boldsymbol{e}_k)a_{ij}a_{kj} \\ &\quad + (1-p_j)p_j^2\sum_{i,k,l=1}^{N}(\boldsymbol{e}_i\odot\boldsymbol{e}_k\odot\boldsymbol{e}_l)a_{ij}a_{kj}a_{lj} + \ldots + \boldsymbol{\varepsilon}_j \\ &= (1-p_j)\boldsymbol{y}_j + p_j\boldsymbol{y}_j\odot\boldsymbol{x}_j + \boldsymbol{\varepsilon}_j\end{aligned} \tag{1}$$

where $\boldsymbol{y}_j = \sum_{i=1}^{N}\boldsymbol{e}_i a_{ij} = E\boldsymbol{a}_j$ computes the linear part, and $\boldsymbol{\varepsilon}_j \in \mathbb{R}^{L\times 1}$ is the Gaussian noise.

As for the optimization problem associated to (1), the unsupervised MLMp [5] proposes a simplified loss function by vanishing the denominator in the original loss function, which is given by

$$\begin{aligned}&\underset{E,\boldsymbol{a}_j,p_j}{\arg\min}\sum_{j=1}^{T}\left\|(1-p_j)\boldsymbol{y}_j + p_j\boldsymbol{y}_j\odot\boldsymbol{x}_j - \boldsymbol{x}_j\right\|_2^2 \\ &s.t.\quad 0\le E\le 1;\quad p_j\le 1;\quad \boldsymbol{a}_j\ge 0\quad\text{and}\quad \mathbf{1}_N^\top\boldsymbol{a}_j = 1,\end{aligned} \tag{2}$$

where $\mathbf{1}_N \in \mathbb{N}^{N\times 1}$ is the vector of 1, both the abundance non-negative constraint (ANC) and sum-to-one constraint (ASC) are imposed, and the probability parameter is bounded.

III. Proposed Robust Multilinear Unmixing Model

To enhance the unmixing performance against corrupted channels and outliers, we propose a novel robust MLM method. To estimate the parameters in (1), we consider following optimization problem based on the $l_{2,1}$ norm

$$\begin{aligned}&\underset{A,P}{\arg\min}\,\|(I_{L\times T}-\bar{P})\odot Y + \bar{P}\odot Y\odot X - X\|_{2,1} + \gamma\|A\|_{1,1} \\ &s.t.\quad \boldsymbol{a}_j\ge 0\quad\text{and}\quad 1_N^\top\boldsymbol{a}_j = 1; p_j\le 1\end{aligned} \tag{3}$$

where $I_{L\times T}$ is the matrix of 1, $\bar{P} = \mathbf{1}_L P \in \mathbb{R}^{L\times T}$, $Y = EA$, and $\gamma \ge 0$ is the parameter controlling the influence of abundance sparseness.

A. ADMM algorithm for R-MLM

We apply the well-known ADMM algorithm to solve (3). After introducing the auxiliary variables Q, G, and H, the augmented Lagrangian function of (3) is written by

$$\begin{aligned}&\underset{A,P,Q,G,H}{\arg\min}\,\|Q\|_{2,1} + \gamma\|G\|_{1,1} \\ &+ \iota_{\{1_T^\top\}}(1_N^\top A) + 1_{\mathbb{R}_+^{N\times T}}(G) + \iota_{\{H|H\le 1_T^\top\}}(H) \\ &+ \frac{\mu}{2}\|(1_{L\times T}-\bar{P})\odot Y + \bar{P}\odot Y\odot X - X - Q + V_1\|_F^2 \\ &+ \frac{\mu}{2}\|A - G + V_2\|_F^2 + \frac{\mu}{2}\|P - H + V_3\|_F^2\end{aligned} \tag{4}$$

where $\mu \ge 0$ is the regularization parameter, and V_1, V_2 and V_3 are scaled dual variables. We address problem (4) iteratively, by alternating the optimization over each variable block $\{A, P, Q, G, H\}$, while keeping the others fixed.

1) Update Q: After discarding the irrelevant terms of Q in (4), the reduced subproblem becomes

$$\begin{aligned}Q^{(t+1)} &= \underset{Q}{\arg\min}\,\|Q\|_{2,1} \\ &+ \frac{\mu}{2}\|(1_{L\times T}-\bar{P}^{(t)})\odot Y^{(t)} + \bar{P}^{(t)}\odot Y^{(t)}\odot X - X - Q + V_1^{(t)}\|_F^2.\end{aligned} \tag{5}$$

Similar as in [9], each row of Q is updated by

$$Q^{(t+1)}(r,:) = \text{vect-soft}\Big(\zeta(r,:), \frac{1}{\mu}\Big), \tag{6}$$

where $\zeta = (1_{L\times T}-\bar{P}^{(t)})\odot Y^{(t)} + \bar{P}^{(t)}\odot Y^{(t)}\odot X - X - Q + V_1^{(t)}$, and vect-soft$(b,\tau)$ is the rowwise application of the vector soft threshold operator defined by $g(b,\tau) = b(\max\{\|b\|_2 - \tau, 0\}/\max\{\|b\|_2 - \tau, 0\} + \tau)$.

2) Update A and G: Removing the terms independent of A and G, we obtain

$$\begin{aligned}A^{(t+1)} &= \underset{A}{\arg\min}\,\iota_{\{1_T^\top\}}(1_N^T A) \\ &+ \frac{\mu}{2}\|(1_{L\times T}-\bar{P}^{(t)})\odot Y + \bar{P}^{(t)}\odot Y\odot X \\ &- X - Q^{(t+1)} + V_1^{(t)}\|_F^2 + \frac{\mu}{2}\|A - G^{(t)} + V_2^{(t)}\|_F^2.\end{aligned} \tag{7}$$

Denote $\bar{E}^{(t)} = E\odot((1-P_j^{(t)})1_{L\times N} + P_j^{(t)}\boldsymbol{x}_j 1_N^\top)$, and $\bar{\boldsymbol{x}}_j = \boldsymbol{x}_j + Q_j^{(t+1)} - (V_1)_j^{(t)}$, where $Q_j^{(t+1)}$ and $(V_1)_j^{(t)}$ denotes the j-th column of the corresponding matrix, repectively. Note that similar symbols will be directly used hereafter. The equivalent pixel-wise optimization problem is

$$\begin{aligned}\boldsymbol{a}_j^{(t+1)} &= \underset{a_j}{\arg\min}\,\iota_{\{1\}}(1_N^\top\boldsymbol{a}_j) \\ &+ \frac{\mu}{2}\|\bar{E}^{(t)}\boldsymbol{a}_j - \bar{\boldsymbol{x}}_j\|_F^2 + \frac{\mu}{2}\|\boldsymbol{a}_j - G_j^{(t)} + (V_2)_j^{(t)}\|_F^2.\end{aligned} \tag{8}$$

As in [16], the solution of (8) is

$$\boldsymbol{a}_j^{(t+1)} = B^{-1}\omega - C(1_N^\top B^{-1}\omega - 1_N) \tag{9}$$

where $B = (\bar{E}^{(t)})^\top\bar{E}^{(t)} + I_{N\times N}$, $C = B^{-1}1_N(1_N^\top B^{-1}1_N)^{-1}$, $\omega = (\bar{E}^{(t)})^T\bar{x}_j + G_j^{(t)} - (V_2)_j^{(t)}$, and $I_{N\times N}$ is identity matrix.

To update G, we consider the optimization subproblem

$$G^{(t+1)} = \underset{G}{\arg\min}\,\gamma\|G\|_{1,1} + \iota_{\mathbb{R}_+^{N\times T}}(G) + \frac{\mu}{2}\left\|A^{(t+1)} - G + V_2^{(t)}\right\|_F^2 \tag{10}$$

Without considering the ANC constraint, which is given by $\iota_{\mathbb{R}_+^{N\times T}}(G)$, the solution of (10) is

$$G^{(t+1)} = \text{soft}_{\gamma/\mu}(A^{(t+1)} + V_2^{(t)}), \tag{11}$$

where

$$\text{soft}_\tau(b) = \begin{cases}\zeta - b & b > \tau \\ 0 & \|b\| < \tau \\ b + \tau & b < -\tau\end{cases} \tag{12}$$

is the soft threshold operator [13]. To further impose ANC, we project the result onto the first orthant by

$$G^{(t+1)} = \max\{0_{N\times T}, G^{(t+1)}\}. \tag{13}$$

3) Update P and H: The reduced optimization problem in terms of P is formulated as

$$P^{(t+1)} = \arg\min_{P} \frac{\mu}{2}\|P - H^{(t)} + V_3^{(t)}\|_F^2 + \frac{\mu}{2}\|(1_{L\times T}-\bar{P})\odot Y^{(t+1)}+\bar{P}\odot Y\odot X - X - Q^{(t+1)}+V_1^{(t)}\|_F^2, \tag{14}$$

and the solution is

$$P^{(t+1)} = \Big(1_L^\top\big((Y^{(t+1)}-Y^{(t+1)}\odot X) \odot (Y^{(t+1)}-X-Q^{(t+1)}+V_1^{(t)})\big)+H^{(t)}-V_3^{(t)}\Big) ./\big(1_L^\top(Y^{(t+1)}-Y^{(t+1)}\odot X)^2+1_T^\top\big), \tag{15}$$

where matrix $Y^{(t+1)}=EA^{(t+1)}$ records the linear mixing part.

Similarly, the subproblem w.r.t H is given by

$$H^{(t+1)} = \arg\min_{H} \iota_{\{H\le 1_T^\top\}}(H) + \left\|P^{(t+1)}-H+V_3^{(t)}\right\|_F^2. \tag{16}$$

If we omit the boundary constraint with $\iota_{\{H\le 1_T^\top\}}(H)$, the solution of (16) would be

$$H^{(t+1)} = P^{(t+1)} + V_3^{(t)}. \tag{17}$$

The final solution of $H^{(t+1)}$ is given by

$$H^{(t+1)} = \min\{1_T^\top, H^{(t+1)}\}. \tag{18}$$

4) Update V_1, V_2 and V_3: The scaled dual variable blocks V_1, V_2 and V_3 represent the running sum of the primal residuals, and are updated as follows

$$V_1^{(t+1)} = V_1^{(t)} + (1_{L\times T} - \bar{P}^{(t+1)})\odot Y^{(t+1)} + \bar{P}^{(t+1)}\odot Y^{(t+1)}\odot -X - Q^{(t+1)} \tag{19}$$

$$V_2^{(t+1)} = V_2^{(t)} + A^{(t+1)} - G^{(t+1)} \tag{20}$$

$$V_3^{(t+1)} = V_3^{(t)} + P^{(t+1)} - H^{(t+1)}. \tag{21}$$

B. Initialization and stopping criterion

As the proposed R-MLM is a supervised unmixing method, the endmembers are extracted by the VCA [14]. The abundance matrix is initialized by the FCLS [15], the nonlinear parameter vector is initialized by (15), with the boundary constraint satisfied. Other variable blocks are initialized with zero matrix/vector.

The stopping criterion of the R-MLM is two-folds: 1) the tolerances of primal and dual residuals satisfy $\mathrm{res_p} \le \mathrm{tol1}$ and $\mathrm{res_d} \le \mathrm{tol2}$, with $\mathrm{tol1} = \mathrm{tol2} = \sqrt{\mathrm{NT}}\times 10^{-5}$, as in [16]; 2) the iteration number does not exceed the preset maximum value.

Algorithm 1 Algorithm for R-MLM

Input: $X \in \mathbb{R}^{L\times T}$: HSI data; $E \in \mathbb{R}^{L\times N}$: endmember matrix
Output: $A \in \mathbb{R}^{N\times T}$: abundance matrix; $P \in \mathbb{R}^{1\times T}$: probability parameter vector

1: **set** $t = 0$, $\mu > 0$, and $\gamma > 0$.
2: **initialize** $A^{(0)}$,$P^{(0)}$, $Q^{(0)}$, $G^{(0)}$, $H^{(0)}$, $V_1^{(0)}$, $V_2^{(0)}$and $V_3^{(0)}$.
3: **while** stopping criterion not met, **do**
4: Update $Q^{(t+1)}$ with (6).
5: Update $S^{(t+1)}$, $G^{(t+1)}$ with (9), (13), respectively.
6: Update $P^{(t+1)}$, $H^{(t+1)}$ with (15), (18), respectively.
7: Update $V_1^{(t+1)}$, $V_2^{(t+1)}$, and $V_3^{(t+1)}$ with (19), (20), and (21), respectively.
8: $t = t + 1$
9: **end while**

IV. Experiments

The performance of the proposed R-MLM is demonstrated on unmixing both the synthetic and real datasets, which are all seriously corrupted by noise. Four state-of-the-art unmixing approaches are compared, including the linear assumption based FCLS [15], a bilinear-based method GBM [3], and two multilinear unmixing methods, *i.e.*, MLM [4] and MLMp [5].

A. Experiments on synthetic DC data

The synthetic DC data is generated as follows: Five random endmembers are selected from the USGS spectral library [17]. The MLM is used to generate 50×50 mixed pixels, with the abundance vectors following the Dirichlet distribution and meeting both ANC and ASC. The nonlinear probability parameter P is randomly selected from the half-normal distribution with $\sigma = 0.3$. To simulate the noisy bands and outliers, in addition to the Gaussian noise (SNR = 20, 25, and 30dB), we also impose the impulse noise and dead lines [18].

In the experiments, the maximum iteration number is set to be 500, the tolerance value is chosen as previously stated, and the penalty parameter which controls the convergence rate is set as $\mu = 0.5$. The sparseness parameter γ is selected from the candidate values set $\{0.0001, 0.005, 0.001, 0.005, 0.01, 0.05, 0.1\}$, and we roughly set $\gamma = 0.01$ in this experiment. For the comparing methods, we keep the same parameter settings as in their original papers. As the unsupervised MLMp [5] is performed in a supervised manner, the endmembers are fixed to the result of the VCA. With groundtruth information on the actual abundances, the unmixing performance on DC image is evaluated by the root-mean-square error, defined by RMSE $= \sqrt{\frac{1}{NT}\sum_{t=1}^{T}\|\boldsymbol{a}_t - \overline{\boldsymbol{a}}_t\|^2}$ for abundance estimation.

As in TABLE I, the proposed R-MLM always yields the optimal unmixing results on DC data, at three noise levels. It shows the robustness of the proposed method against different types of noise, including the Gaussian noise, impulse noise and dead lines. Also, as the SNR value increases, the performances of all the unmixing strategies improve to some extent. It is not

TABLE I
AVERAGED RMSE OVER 5 RUNS ON DC.

	SNR = 20dB	SNR = 25dB	SNR = 30dB
R-MLM	**0.0365**	**0.0195**	**0.0122**
MLMp	0.0701	0.0592	0.0575
MLM	0.0532	0.0506	0.0507
GBM	0.1386	0.1380	0.1383
FCLS	0.1375	0.1369	0.1372

surprised that the MLM-based methods are generally superior to the others here, as the DC image is generated by MLM.

B. Experiments on Cuprite data

The Cuprite data is returned by AVIRIS and is well-investigated. It has 250×190 pixels and contains 224 bands covering a wavelength range of 0.4-2.5 μm. To test the algorithm robustness against noisy channels, all the 224 bands are retained for analysis, including the water vapor absorption and low signal-to-noise ratio bands (4-104, 116-149 and 171-187). This field is known to be mainly composed by 12 endmembers, which are extracted by the VCA.

Since the groundtruth information is unknown for real HSI, the unmixing results are evaluated by the reconstruction error (RE) and the averaged spectral angle distance (SAD) between the observed and reconstructed pixels, as reported in TABLE II. As observed, the proposed R-MLM algorithm leads to the best results, demonstrating its robustness when processing HSIs containing outlier bands. Fig. 1 compares the abundance maps of endmember #alunite, where the estimation by R-MLM is affected by the corrupted bands least.

TABLE II
COMPARISON OF RE AND SAD ON CUPRITE, WITHOUT CONSIDERING BANDS 4-104, 116-149 AND 171-187.

	R-MLM	MLMp	MLM	GBM	FCLS
RE	**0.0020**	0.0074	0.0060	0.0079	0.0082
SAD	**0.0366**	0.0871	0.0738	0.0739	0.0751

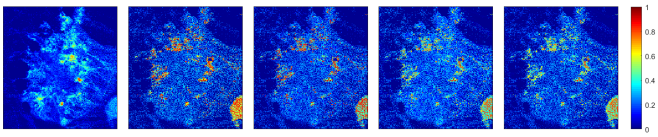

Fig. 1. Abundance maps of #alunite by R-MLM, MLMp, MLM, GBM and FCLS on Cuprite.

V. CONCLUSION

This paper proposed a novel robust multilinear unmixing model. Taking advantage of the $l_{2,1}$ norm, the interference of different noises to the unmixing process was suppressed. To improve the parameter estimation, the abundance sparseness was also introduced. The associated optimization problem was addressed by the ADMM. Results on the synthetic and real HSIs demonstrated the robustness of the proposed unmixing method against different types of noise. Future works include the investigation of other loss functions, *e.g.*, Welsch function, for robust unmixing.

REFERENCES

[1] J. M. P. Nascimento and J. M. Bioucas-Dias. Nonlinear mixture model for hyperspectral unmixing. Image and Signal Processing for Remote Sensing XV, vol. 7477, pp. 157-164, Sept. 2009.

[2] W. Fan, B. Hu, J. Miller and M. Li. Comparative study between a new nonlinear model and common linear model for analysing laboratory simulate-forest hyperspectral data. International Journal of Remote Sensing, vol. 30, no. 11, pp. 2951-2962, Jun. 2009.

[3] A. Halimi, Y. Altmann, N. Dobigeon and J. Tourneret. Nonlinear Unmixing of Hyperspectral Images Using a Generalized Bilinear Model. IEEE Transactions on Geoscience and Remote Sensing, vol. 49, no. 11, pp. 4153-4162, Nov. 2011.

[4] R. Heylen and P. Scheunders. A Multilinear Mixing Model for Nonlinear Spectral Unmixing. IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 1, pp. 240-251, Jan. 2016.

[5] Q. Wei, M. Chen, J. Tourneret and S. Godsill. Unsupervised Nonlinear Spectral Unmixing Based on a Multilinear Mixing Model. IEEE Transactions on Geoscience and Remote Sensing, vol. 55, no. 8, pp. 4534-4544, Aug. 2017.

[6] B. Yang and B. Wang. Band-Wise Nonlinear Unmixing for Hyperspectral Imagery Using an Extended Multilinear Mixing Model. IEEE Transactions on Geoscience and Remote Sensing, vol. 56, no. 11, pp. 6747-6762, Nov. 2018.

[7] F. Zhu, A. Halimi, P. Honeine, B. Chen and N. Zheng. Correntropy Maximization via ADMM: Application to Robust Hyperspectral Unmixing. IEEE Transactions on Geoscience and Remote Sensing, vol. 55, no. 9, pp. 4944-4955, Sept. 2017.

[8] J. Li, J. M. Bioucas-Dias, A. Plaza and L. Liu. Robust Collaborative Nonnegative Matrix Factorization for Hyperspectral Unmixing. IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 10, pp. 6076-6090, Oct. 2016.

[9] Y. Ma, C. Li, X. Mei, C. Liu and J. Ma. Robust Sparse Hyperspectral Unmixing With $\ell_{2,1}$ Norm. IEEE Transactions on Geoscience and Remote Sensing, vol. 55, no. 3, pp. 1227-1239, Mar. 2017.

[10] D. Kong, C. Ding and H. Huang. Robust Nonnegative Matrix Factorization using $l_{2,1}$ norm. Proceedings of the 20th ACM International Conference on Information and Knowledge Management. pp. 673-682, Oct. 2011.

[11] C. Ding, D. Zhou, X. He and H. Zha. R1-PCA: Rotational Invariant l_1 norm Principal Component Analysis for Robust Subspace Factorization. Proceedings of the 23rd International Conference on Machine Learning, pp. 281-288, Jun. 2006.

[12] C.X. Ren, D.Q. Dai and H. Yan. Robust Classification Using $l_{2,1}$ norm Based Regression Model. Pattern Recognition, vol. 45, no. 7, pp. 2708-2718, Jul. 2012.

[13] S. Boyd, N. Parikh, E. Chu, B. Peleato and J. Eckstein. Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers. Found. Trends Mach. Learn, vol. 3, no. 1, pp. 1-122, Jan. 2011.

[14] J. M. P. Nascimento and J. M. B. Dias. Vertex component analysis: a fast algorithm to unmix hyperspectral data. IEEE Transactions on Geoscience and Remote Sensing, vol. 43, no. 4, pp. 898-910, Apr. 2005.

[15] D. C. Heinz and C. I. Chang. Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery. IEEE Transactions on Geoscience and Remote Sensing, vol. 39, no. 3, pp. 529-545, Mar. 2001.

[16] J. M. Bioucas-Dias and M. A. T. Figueiredo. Alternating direction algorithms for constrained sparse regression: Application to hyperspectral unmixing. 2010 2nd Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, pp. 1-4, 2010.

[17] J. M. Bioucas-Dias and J. M. P. Nascimento. Hyperspectral Subspace Identification. IEEE Transactions on Geoscience and Remote Sensing, vol. 46, no. 8, pp. 2435-2445, Aug. 2008.

[18] X. Mei, Y. Ma, C. Li, et al. Robust GBM hyperspectral image unmixing with superpixel segmentation based low rank and sparse representation. Neurocomputing, vol. 275, pp. 2783-2797, 2018.