Endmember Distinguished Low-Rank and Sparse Representation for Hyperspectral Unmixing

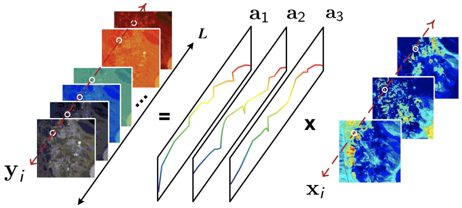
Jin-Liang Xiao
Joint work with *Ruifeng Ren* and *Jie Huang*

University of Electronic Science and Technology of China (UESTC)
Chengdu, Sichuan, China

July 8, 2024

Outline

Hyperspectral unmixing (HU): Signal model



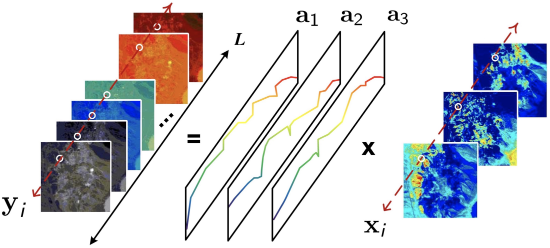
Courtesy to [Wing-Kin Ma *et al.* 2014]

- Linear Mixture Model (LMM):

$$y_i = Ax_i + e_i, \quad i = 1, \ldots, k.$$

where

- $A = [a_1, \ldots, a_n] \in \mathbb{R}^{m \times n}$, $a_i \in \mathbb{R}^m$ is an *endmember* signature vector;
- $y_i \in \mathbb{R}^m$ is the measured hyperspectral vector at pixel i;
- $x_i \in \mathbb{R}^n$ is the *abundance* vector at pixel i; e_i is the noise.



- Let $Y = [y_1, \ldots, y_k]$, $X = [x_1, \ldots, x_k]$, $E = [e_1, \ldots, e_k]$. Then

$$Y = AX + E.$$

- **Goal:** recover X from Y with knowing A.
- **Assumptions:**
  - ANC: abundance nonnegative constraint: $X \geq 0$, i.e., $X_{i,j} \geq 0$.
  - ASC: abundance sum-to-one constraint: $\sum_{i=1}^{m} X_{i,j} = 1$, $\forall j$.

Sparse regression for HU.

- **Idea:** the number of endmembers participating in a mixed pixel is usually very small compared with the (ever-growing) dimensionality (and availability) of spectral libraries.
- The sparse unmixing model [Iordache-Bioucas-Plaza2011] is:

$$\min_{\mathbf{0}\leq X\in\mathbb{R}^{m\times n}} \frac{1}{2}\|Y-AX\|_F^2+\lambda\|X\|_1$$

- solve the model by a fast custom-derived solver (e.g., alternating direction method of multipliers (ADMM))
- SUnSAL: the sparse unmixing via variable splitting and augmented Lagrangian

Collaborative sparse regression.

- **Idea:** use a smallest subset of spectral library to represent all mixed pixels in a hyperspectral image.
- The collaborative sparse model [Iordache-Bioucas-Plaza2014] is:

$$\min_{\mathbf{0} \leq X \in \mathbb{R}^{m \times n}} \frac{1}{2} \|Y - AX\|_F^2 + \lambda \|X\|_{2,1}$$

where $\|X\|_{2,1} = \sum_{i=1}^{m} \left\|x^{[i]}\right\|_2$, and $x^{[i]}$ is the i-th row of X.

- solve it by ADMM.
- the collaborative SUnSAL (CLSUnSAL) algorithm.

- **SUnSAL**: sparsity on each column vector x_j, i.e., X is sparse.
- **CLSUnSAL**: structural sparsity on X, i.e., X is row-sparse.
  - collaborative sparse on a sliding window [Chen-Nasrabadi-Tran2011], [Qu-Nasrabadi-Tran2014]
  - collaborative sparse with general segments (super-pixels) [Huang-Zhang-Pižurica2017], [Wang-Zhong-Zhang-Xu2017]
- **SUnSAL-TV**: the total variation (TV) regularization is included into the classical sparse regression formulation [Iordache-Bioucas-Plaza2012]

$$\min_{X \geq \mathbf{0}} \frac{1}{2} \|Y - AX\|_F^2 + \lambda \|X\|_1 + \lambda_{TV} TV(X),$$

where $\lambda \geq 0$ and $\lambda_{TV} \geq 0$ are regularized parameters.

Sparse and low-rank unmixing.

- **ADSpLRU**: The spatial correlation among pixels in an HSI translates into a *low-rank* property of the abundance matrix [Giampouras-Themelis-Rontogiannis-Koutroumbas2016].
- The unmixing model is

$$\min_{\mathbf{0}\leq X\in\mathbb{R}^{m\times K}} \frac{1}{2}\|Y-AX\|_F^2+\lambda\|X\|_{a,1}+\tau\|X\|_{f,*}$$

where $\|X\|_{a,1}=\sum|a_{i,j}X_{i,j}|$, $\|X\|_{f,*}=\sum_{i=1}^{\mathrm{rank}(X)} f(\sigma_i)\sigma_i$ is the weighted nuclear norm of X, and σ_i is the i-th singular value of X.

- **JSpBLRU**: Partition

$$X = [X_1, \cdots, X_s] \in \mathbb{R}^{m \times n},$$

where $X_j \in \mathbb{R}^{m \times d_j}$, for $j = 1, \ldots, s$, $\sum_{j=1}^{s} d_j = n$, and block number s is a positive integer for $1 \leq s \leq n$. Each X_j is assumed joint-sparse [Huang-Huang-Deng-Zhao2019].

- Consider the joint-sparse-blocks and low-rank hyperspectral unmixing model

$$\min_{\mathbf{0} \leq X \in \mathbb{R}^{m \times K}} \frac{1}{2} \|Y - AX\|_F^2 + \lambda \sum_{j=1}^{s} \|X_j\|_{w_j,2,1} + \tau \|X\|_{f,*}$$

where the *weighted $\ell_{2,1}$ norm*, defined as $\|X_j\|_{w_j,2,1} = \sum_{i=1}^{m} w_{i,j} \left\| X_j^{[i]} \right\|_2$, is used to enhance sparsity along rows in each block in X, $X_j^{[i]}$ is the i-th row of the j-th block of X, $w_j = [w_{1,j}, \ldots, w_{m,j}]^T$ is a nonnegative vector.

- **BiJSpLRU**: Consider the mode-3 matricization of $\mathcal{Y}$ along the vertical and horizontal directions simultaneously. [Huang-Di-Wang-Lin-Huang2021]

Endmember distinguished methods

Idea: divide the endmembers into two groups.

- *"Low-rank" endmembers*: they participate in the mixing process and are called *active*. The corresponding abundances are gathered in $X_1 \in \mathbb{R}^{n_1 \times k}$ and we will exploit more spatial information of abundance maps (although sparsity is also important for them).
- *"Sparse" endmembers*: they are *inactive*, resulting in the sparsity of X. The corresponding abundances are gathered in $X_2 \in \mathbb{R}^{n_2 \times k}$.
- Divide the abundance matrix $X \in \mathbb{R}^{n \times k}$ into two row blocks according to the activity. That says

$$PX = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix},$$

where $P = [P_1^T, P_2^T]^T \in \mathbb{R}^{n \times n}$ is a permutation matrix used to reorder the rows of the abundance matrix, $P_1 X = X_1$ and $P_2 X = X_2$ are composed of rows corresponding to n_1 low-rank endmembers and n_2 sparse endmembers with $n_1 + n_2 = n$, respectively.

Endmember distinguished methods

How to distinguish the low-rank and sparse endmembers?

- The key is to determine the permutation matrix P.
- Step 1. Sort the endmembers in descending order by $\|X(i,:)\|_2$, where $X(i,:)$ denotes the i-th row of X for $i = 1, 2, ..., n$.
- Step 2. Use ℓ_2-norm to measure activity. We find the set S of the top K endmembers satisfying that

$$\sum_{i \in S} \|X(i,:)\|_2 \geq \rho \sum_{i=1}^{n} \|X(i,:)\|_2 \, , \tag{1}$$

  where $\rho \in [0, 1]$ is a thresholding parameter.
- Step 3. The endmembers in set S will be considered as *low-rank* endmembers and the rest will be *sparse* ones.

Weighted nuclear norm regularization

- We apply the weighted nuclear norm regularization on the abundance map of each low-rank endmember, instead of applying on X_1 directly, which can be described as

$$\sum_{i=1}^{n_1} \|Re(X_1(i,:))\|_{g,*}$$

where n_1 is the number of low-rank endmembers, $Re(\cdot)$ is a function that reshapes the row vector of the i-th low-rank endmember into abundance map.

- Here the weighted nuclear norm is defined as

$$\|T\|_{g,*} = \sum_i g(\sigma_i)\sigma_i,$$

where σ_i is the i-th singular value of the matrix T and g is a weighting function for singular values defined as $g(\sigma) = 1/(\sigma + \varepsilon)$, and ε is a small number to avoid singularities [Gu-Xie-Meng-Zuo-Feng-Zhang2017].

Weighted sparse regularization

- We consider the sparse property for the abundance matrix X.
- In particular, we introduce a weight factor $B \in \mathbb{R}^{n \times k}$ into the ℓ_1-norm of X.
- The weight factor B consists of a spectral factor $B_1 \in \mathbb{R}^{n \times k}$ and a spatial factor $B_2 \in \mathbb{R}^{n \times k}$ to make full use of spectral and spatial information. [Zhang-Li-Li-Deng-Plaza2018]
- The weighted ℓ_1-norm regularization on X is as follows

$$\|B \odot X\|_1, \quad B = \mathrm{sqrt}(B_1 B_2), \tag{2}$$

where the operator $\odot$ denotes the Hadamard product and $\mathrm{sqrt}(\cdot)$ is the elementwise square root function.

Weighted sparse regularization

- The spectral factor B_1 is used to enhance the sparsity of the endmembers in the spectral library, which is defined as

$$B_1 = \mathrm{diag}\left[\frac{k}{\|X(1,:)\|_1 + \varepsilon}, ..., \frac{k}{\|X(n,:)\|_1 + \varepsilon}\right].$$

- The spatial factor B_2 utilizes the spatial correlation between the pixel and its neighbors to promoting sparsity, which can be described as

$$B_2(i,j) = \frac{\sum_{k \in \mathcal{N}(j)} w_k}{\sum_{k \in \mathcal{N}(j)} w_k X(i,k) + \varepsilon},$$

where $\mathcal{N}(j)$ is a collection of the column indices of the j-th pixel's neighbors (including itself) in the abundance matrix and w_k is corresponding weight of neighbors.

Proposed model

- In conclusion, our proposed unmixing model is given as

$$\min_{X} \frac{1}{2} \|Y - AX\|_F^2 + \lambda \sum_{i=1}^{n_1} \|Re(X_1(i,:))\|_{g,*} + \tau \|B \odot X\|_1$$
$$\text{s.t.} \quad X \geq 0, \quad PX = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix},$$

- To solve this model under the ADMM framework. We transform this model into a constrained problem as

$$\min_{X} \; \frac{1}{2} \|Y - AX\|_F^2 + \lambda \sum_{i=1}^{n_1} \|Re(V_1(i,:))\|_{g,*} + \tau \|B \odot V_2\|_1 + \iota_{R_+}(V_3)$$
$$\text{s.t.} \quad P_1 X = V_1, \; X = V_2, \; X = V_3,$$

where V_1, V_2, V_3 are auxiliary varaibles, $\iota_{R_+}(x)$ is the indicator function, λ and τ are nonnegative regularization parameters.

Proposed algorithm

- Let $\Omega_1 \in \mathbb{R}^{n_1 \times k}$, $\Omega_2 \in \mathbb{R}^{n \times k}$, and $\Omega_3 \in \mathbb{R}^{n \times k}$ are Lagrange multipliers.
- Define $G = [P_1^T, I, I]^T$, $E = \mathrm{diag}(-I, -I, -I)$, $V = [V_1^T, V_2^T, V_3^T]^T$, $\Omega = [\Omega_1^T, \Omega_2^T, \Omega_3^T]^T$, then the constraints become $GX + EV = 0$.
- Define

$$\mathcal{L}_\mu(X, V; \Omega) = \frac{1}{2} \|Y - AX\|_F^2 + \lambda \sum_{i=1}^{n_1} \|Re(V_1(i,:))\|_{g,*} + \tau \|B \odot V_2\|_1 + \iota_{R_+}(V_3) + \frac{\mu}{2} \|GX + EV - \Omega\|_F^2,$$

  where $\mu > 0$ is a penalty parameter.
- The ADMM framework is derived

$$\begin{cases} X^{k+1} = \underset{X}{\operatorname{argmin}}\, \mathcal{L}_\mu(X, V^k; \Omega^k), \\ V^{k+1} = \underset{V}{\operatorname{argmin}}\, \mathcal{L}_\mu(X^{k+1}, V; \Omega^k), \\ \Omega^{k+1} = \Omega^k - (GX^{k+1} + E\Omega^{k+1}). \end{cases}$$

Proposed algorithm

- To make the notations clearly, we introduce the soft-thresholding (SHR) operator and singular value thresholding (SVT) operator.
- Let $X = U\Sigma V^T$ be the singular value decomposition of X and recall the weighting function. Define

$$\mathrm{SHR}_{g,\alpha}(x) = \mathrm{sign}(x)\max\left(0, x - \alpha g(x)\right),$$
$$\mathrm{SVT}_{g,\beta}(X) = U\mathrm{SHR}_{g,\beta}\left(\Sigma\right)V^T.$$

- Now we can propose the final algorithm named as Endmember Distinguished Low-rank and Sparse Representation Unmixing (EDLSpRU).

Proposed algorithm

Algorithm 1: EDLSpRU

Input: Y and A.
Selected parameters: λ, τ, μ and maximum iterations
Initialization: X^0, V^0, Ω^0, and $t = 1$
Repeat:

$$X^t = (A^T A + 2\mu I + P_1^T P_1)^{-1}[A^T Y + \mu P_1^T (V_1^{t-1} + \Omega_1^{t-1}) + \mu(V_2^{t-1} + \Omega_2^{t-1} + V_3^{t-1} + \Omega_3^{t-1})]$$

for $i = 1, 2, ..., n_1$

$$Re(V_1^t(i,:)) = \mathrm{SVT}_{g,\frac{\lambda}{\mu}}(Re(P_1 X^t(i,:) - \Omega_1^{t-1}(i,:)))$$

end for

$V_2^t = \mathrm{SHR}_{B,\frac{\tau}{\mu}}(X^t - \Omega_2^{t-1})$ with B calculated by (2)

$V_3^t = \max(X^t - \Omega_3^{t-1}, 0)$

$\Omega_1^t = \Omega_1^{t-1} - P_1 X^t + V_1^t$

$\Omega_2^t = \Omega_2^{t-1} - X^t + V_2^t$

$\Omega_3^t = \Omega_3^{t-1} - X^t + V_3^t$

Update P_1 and n_1 according to the set S satisfying (1)

until some stopping criterion is satisfied.
Output: $\hat{X} = X^t$

Experiment setting.

- We will compare EDLSpRU with three state-of-the-art algorithms: ADSpLRU, JSpBLRU and BiJSpLRU.
- RMSE and the signal-to-reconstruction error (SRE) are used to evaluate the performance of unmixing results. They are defined by, respectively,

$$\text{RMSE} = \sqrt{\frac{1}{nk}\sum_{i=1}^{k} \|\hat{x}_i - x_i\|_2^2}, \quad \text{SRE (dB)} = 10\log_{10}\left(\frac{\frac{1}{k}\sum_{i=1}^{k}\|\hat{x}_i\|_2^2}{\frac{1}{k}\sum_{i=1}^{k}\|\hat{x}_i - x_i\|_2^2}\right),$$

where k is the number of pixels, n is the number of endmembers, and $\hat{x}_i$ and x_i are estimated and exact abundance vectors of the i-th pixel, respectively.

- We set the maximum number of iterations to 500.
- We select optimal regularization parameters for EDLSpRU as follows:

$$\lambda, \tau \in \{0.0001, 0.0005, 0.001, 0.005, 0.01, 0.05, 0.1, 0.5, 1, 5\},$$

and select $\mu \in \{0.1, 0.5, 1, 5\}$. In addition, we set $\rho = 0.9$ in all experiments. All possible combinations are considered and the optimal parameters are chosen to get maximum SREs.

- Our test were done by using MATLAB R2021a on a laptop with 2.30 GHz Intel Core i7 and 12 GB memory.

Experiments on simulated data

- Data Cube 1 (DC1): The first data cube contains 75×75 pixels with 224 spectral bands. The spectral dictionary $A_1 \in \mathbb{R}^{224 \times 240}$ is extracted from the U.S. Geological Survey (USGS). Five endmembers are randomly chosen from A_1 according to LMM. Finally, the generated data cube is contaminated by white Gaussian i.i.d. noise with signal-to-noise ratio (SNR) of 20, 30, and 40dB respectively.
- Data Cube 2 (DC2): This data cube contains 100×100 pixels with 99 spectral bands and the spectral dictionary $A_2 \in \mathbb{R}^{99 \times 120}$ is from the National Aeronautics and Space Administration Johnson Space Center Spacecraft Materials Spectral Database. Nine endmember signatures are randomly chosen from A_2 and the corresponding abundance maps are used to generate the true data cube, which is also contaminated by Gaussian noise with the same SNR values adopted for DC1.

Table: SRE(dB) and RMSE by different algorithms.

Data Cube 1 (DC1)						
Algorithm	SNR = 20 dB		SNR = 30 dB		SNR = 40 dB	
	RMSE	SRE	RMSE	SRE	RMSE	SRE
ADSpLRU	0.0168	6.24	0.0079	12.82	0.0014	27.70
JSpBLRU	0.0117	9.42	0.0059	15.40	0.0008	32.50
BiJSpLRU	0.0115	9.53	0.0067	14.30	0.0008	32.82
EDLSpRU	**0.0108**	**10.13**	**0.0016**	**26.55**	**0.0005**	**36.70**

Data Cube 2 (DC2)						
Algorithm	SNR = 20 dB		SNR = 30 dB		SNR = 40 dB	
	RMSE	SRE	RMSE	SRE	RMSE	SRE
ADSpLRU	0.0375	5.75	0.0114	16.10	0.0034	26.66
JSpBLRU	0.0230	9.98	0.0087	18.42	0.0027	28.44
BiJSpLRU	0.0180	12.12	0.0069	20.48	**0.0024**	**29.55**
EDLSpRU	**0.0155**	**13.39**	**0.0058**	**22.01**	0.0029	28.02

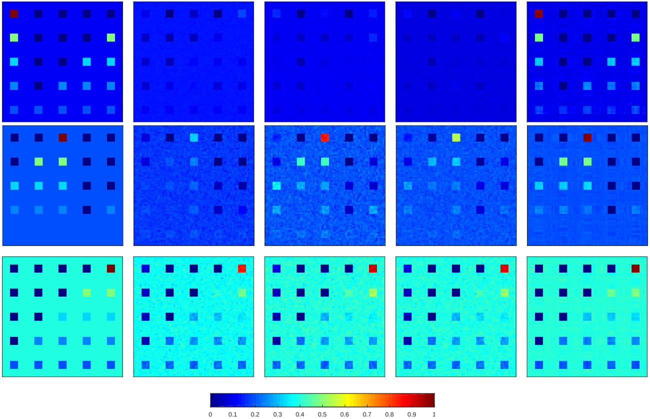


Figure: True and estimated abundance maps for endmembers #1, #3 and #5 by different unmixing algorithms for DC1 with SNR = 30 dB. From left to right: True, ADSpLRU, JSpBLRU, BiJSpLRU, and EDLSpRU.

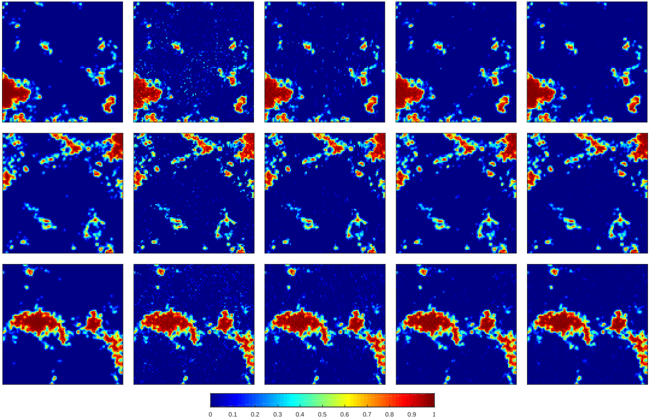
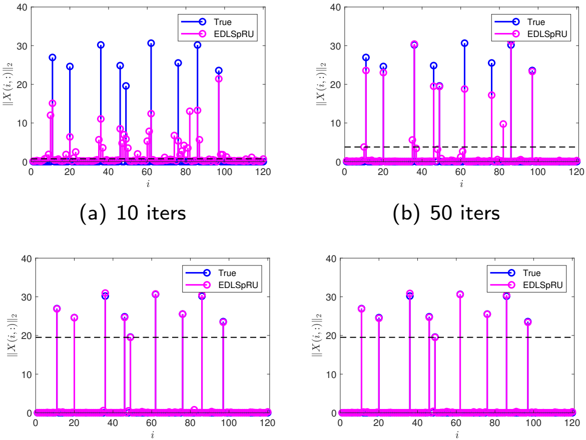


Figure: True and estimated abundance maps for endmembers #2 , #3 and #9 by different unmixing algorithms for DC2 with SNR = 30 dB. From left to right: True, ADSpLRU, JSpBLRU, BiJSpLRU, and EDLSpRU.



(a) 10 iters (b) 50 iters

(c) 100 iters (d) 200 iters

Figure: Comparison of each endmember's abundance of the ground truth and the estimations by EDLSpRU for DC2 with SNR = 30 dB.

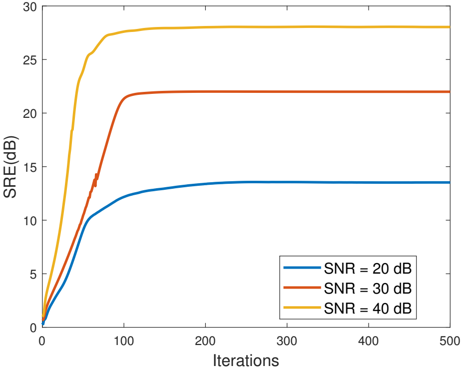


Figure: Plots of SRE (dB) against iterations by EDLSpRU for DC2.

Experiments on real data

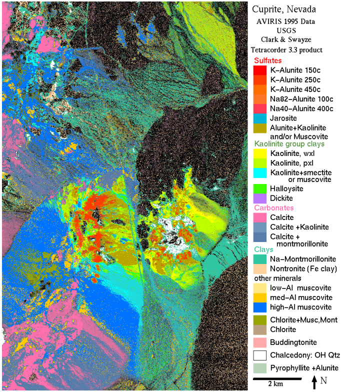


Figure: USGS map showing the location of different minerals in the Cuprite mining district in Nevada.

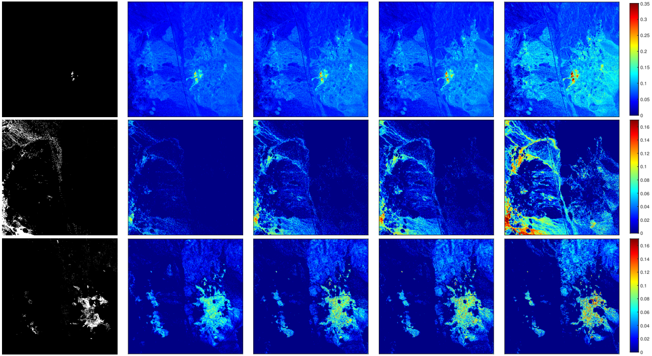


Figure: Estimated abundance maps for minerals buddington, muscovite, and chalcedony by different unmixing algorithms. From left to right: Tetracorder, ADSpLRU, JSpBLRU, BiJSpLRU, and EDLSpRU.

Conclusion and future work

- We have proposed a new unmixing model which can distinguish endmembers into low-rank ones and sparse ones so that different regularizations can be applied to each type of them. The model aims at considering both low-rank and sparse characteristics for low-rank endmembers, while only considering sparsity for sparse ones.
- We solve our model under the ADMM framework and propose a new algorithm named as EDLSpRU. The experiment results for both simulated and real-data demonstrate the efficacy of the proposed unmixing structure.
- In the future, we will extend the proposed endmember distinguished structure to blind unmixing or its tensor edition for HSI processing.

References

[Iordache-Bioucas-Plaza2011] M.-D. Iordache, J. Bioucas-Dias, and A. Plaza, "Sparse unmixing of hyperspectral data," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 6, pp. 2014–2039, Jun. 2011.

[Iordache-Bioucas-Plaza2014] D. Iordache, J. 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.

[Chen-Nasrabadi-Tran2011] Y. Chen, N. Nasrabadi, and T. Tran, "Hyperspectral image classification using dictionary-based sparse representation", *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 3973–3985, Oct. 2011.

[Qu-Nasrabadi-Tran2014] Q. Qu, N. Nasrabadi, and T. Tran, "Abundance estimation for bilinear mixture models via joint sparse and low-rank representation", *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 7, pp. 4404–4423, July 2014.

[Huang-Zhang-Pižurica2017] S. Huang, H. Zhang, and A. Pižurica, "A robust sparse representation model for hyperspectral image classification," *Sensors*, vol. 17, no. 9, pp. 2087, Sept. 2017.

[Wang-Zhong-Zhang-Xu2017] X. Wang, Y. Zhong, L. Zhang, and Y. Xu, "Spatial group sparsity regularized nonnegative matrix factorization for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 11, pp. 6287–6304, Nov. 2017.

[Iordache-Bioucas-Plaza2012] D. Iordache, J. 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.

[Huang-Huang-Deng-Zhao2019] J. Huang, T.-Z. Huang, L.-J. Deng, X.-L. Zhao, "Joint-Sparse-Blocks and Low-Rank Representation for Hyperspectral Unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 2419–2438, April 2019.

[Giampouras-Themelis-Rontogiannis-Koutroumbas2016] P. Giampouras, K. Themelis, A. Rontogiannis, and K. Koutroumbas, "Simultaneously sparse and low-rank abundance matrix estimation for hyperspectral image unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4775–4789, Aug. 2015.

[W.-K. Ma *et al.*2014] W.-K. Ma, J. M. Bioucas-Dias, T.-H. Chan, N. Gillis, et al., "A signal processing perspective on hyperpsectral unmixing," *IEEE Signal Process. Mag.*, vol. 31, no. 1, pp. 67–81, Jan. 2014.

[Wright-Nowak-Figueiredo2009] S. Wright, R. Nowak, and M. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2479–2493, July 2009.

[W. Dong *et al.* 2014]. W. Dong, G. Shi, X. Li, Y. Ma, and F. Huang, "Compressive sensing via nonlocal low-rank regularization," *IEEE Trans. Image Process.*, vol. 23, no. 8, pp. 3618–3632, June 2014.

[Gu-Zhang-Zuo-Feng2014]. S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm minimization with application to image denoising," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 2862–2869.

[Zhang-Li-Li-Deng-Plaza2018]. J. S. Zhang, J. Li, H.-C. Li, C. Deng, and A. Plaza, "Spectral-spatial weighted sparse regression for hyperspectral image unmixing," in *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 6, pp. 3265–3276, 2018.

[Huang-Di-Wang-Lin-Huang2021]. J. Huang, W.-C. Di, J.-J. Wang, J. Lin, and T.-Z. Huang, "Bilateral joint-sparse regression for hyperspectral unmixing," in *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 10147–10161, 2021.

Thank you very much!