Learning a Low-coherence Dictionary to Address Spectral Variability for Hyperspectral Unmixing

Danfeng Hong [a,b], Naoto Yokoya [a,b,c], Jocelyn Chanussot [d], Xiaoxiang Zhu [a,b]

a. Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR)
b. Signal Processing in Earth Observation (SiPEO), Technical University of Munich (TUM)
c. Research Center for Advanced Science and Technology, The University of Tokyo
d. GIPSA-Lab, CNRS, University of Grenoble Alpes

DLR

Roadmap

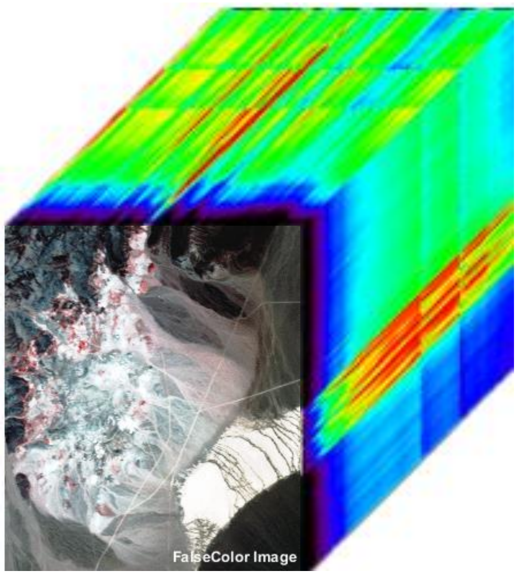
- Background & Motivation
- Proposed Spectral Mixing Model
- Experiments
- Conclusion & Future Work

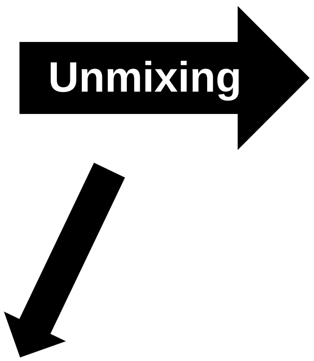
Roadmap

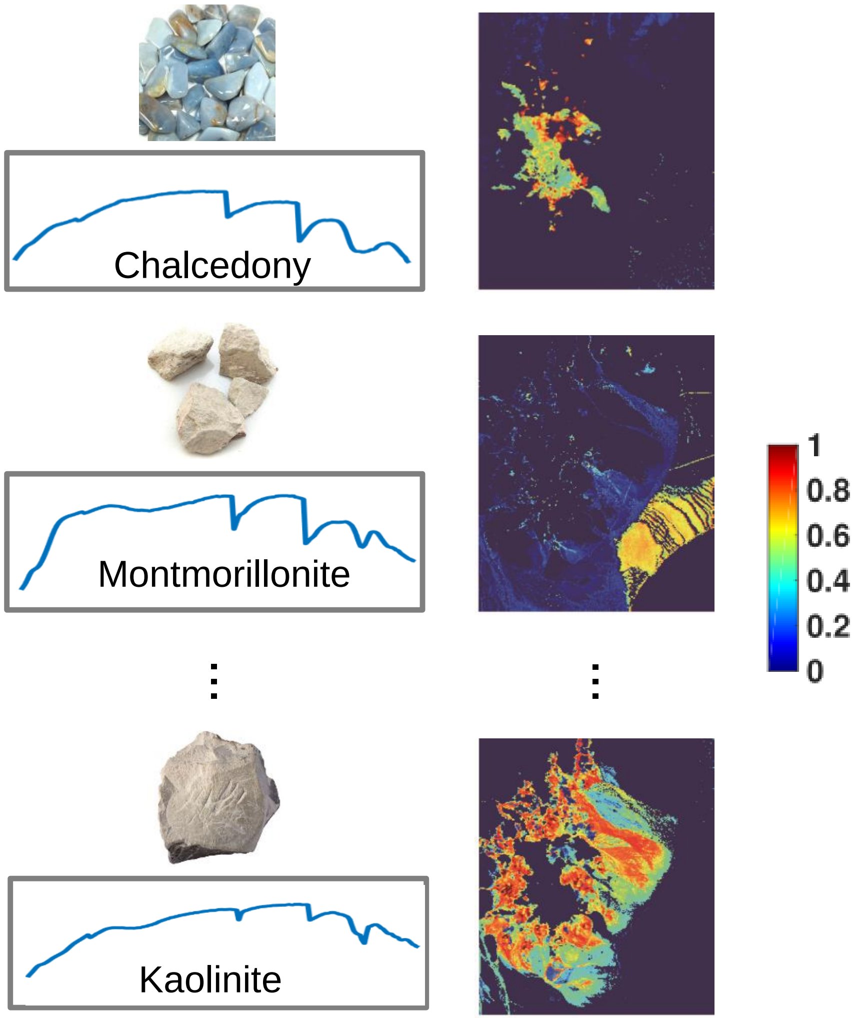
- **Background & Motivation**
- **The Proposed Spectral Mixing Model**
- **Experiments**
- **Conclusion & Future Work**

Hyperspectral Data and Spectral Unmixing

Hyperspectral data (Cuprite)





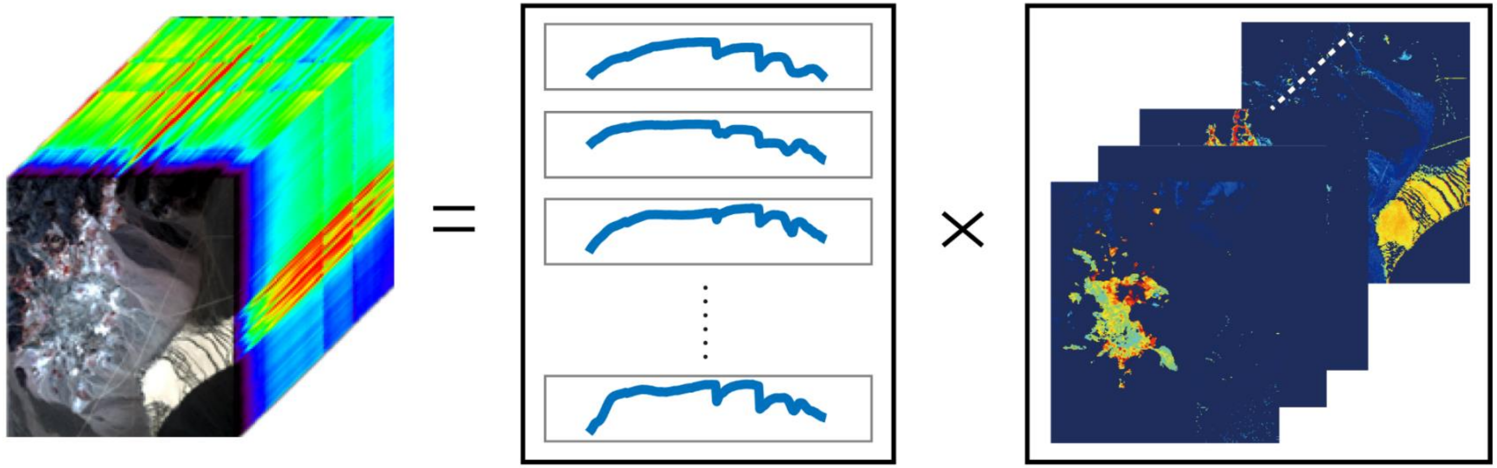


Endmembers **Abundance maps**

separate the hyperspectral data into

- Endmember signatures
- Abundance maps

Linear Mixing Model for Spectral Unmixing



Hyperspectral Cube (Y) *Endmembers Dictionary (A)* *Abundance Maps (X)*

Linear Mixing Model (LMM) : $\mathbf{Y} = \mathbf{A}\mathbf{X} + \boldsymbol{\varepsilon}$ → model errors

$\mathbf{Y} \in \Re^{D\times N}$ $\mathbf{A} \in \Re^{D\times P}$ $\mathbf{X} \in \Re^{P\times N}$ $\boldsymbol{\varepsilon} \in \Re^{P\times N}$

$\mathbf{A}$ → extracted from **Y** by PPI [1] , VCA [2]

Fully Constrained Least Square Unmixing (FCLSU) : $\arg\min_{\mathbf{X}} \|\mathbf{Y} - \mathbf{A}\mathbf{X}\|_F^2 \quad s.t. \quad \mathbf{X} \geq 0, \ \mathbf{1}_{P\times 1}^T \mathbf{X} = \mathbf{1}_{1\times N}$

$\mathbf{X} \geq 0$ → Non-negative constraint

$\mathbf{1}_{P\times 1}^T \mathbf{X} = \mathbf{1}_{1\times N}$ → Sum-to-one constraint

[1] C. Chang et al, IEEE GRL 2006

[2] J. M. P. Nascimento et al, IEEE TGRS 2005.

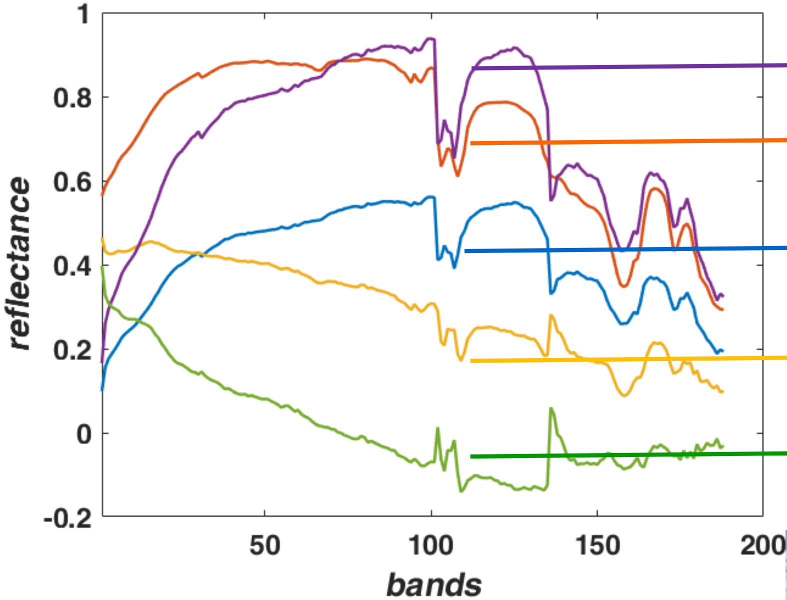
Spectral Variability

$\mathbf{Y}_{real} \longrightarrow \mathbf{Y}_{observed}$

- Scaling (mainly) : illumination, topography change
- Complex Noise : environmental condition , sensor noise
- Atmospheric Effects : physically or chemically
- Nonlinear Mixing : intimate or multilayered mixing
- Others: Offset, etc.

Spectral Variability (SV): $\mathbf{V} \in \Re^{D \times N}$

$$\mathbf{Y}_{observed} = \mathbf{Y}_{real} + \mathbf{V}$$



④ Scaled Alunite

② Alunite in AVIRIS Cuprite data

① Alunite from USGS library (ground measurement)

③ ②-① all spectral variabilities

⑤ ②-④ spectral variability without scaling factors

Variations of LMM

- LMM with spectral variability: $\mathbf{Y}_{observed} \neq \mathbf{AX} + \boldsymbol{\varepsilon}$

$$= \mathbf{AX} + \mathbf{V} + \boldsymbol{\varepsilon}$$

leading to inaccurate estimation of **X**, since **SV** is absorbed by **X.**

- Scaled LMM [1] : $\mathbf{y}_k = s_k \mathbf{A}\mathbf{x}_k + \boldsymbol{\varepsilon}_k$

s_k → a scalar shared for all endmembers

Advantage: speed up to estimate scaling factors
obtain a relatively robust solution

Disadvantage: endmembers share the same scaling factor
ignore other variabilities

- Extended LMM [2] : $\mathbf{y}_k = \mathbf{A}\mathbf{S}_k \mathbf{x}_k + \boldsymbol{\varepsilon}_k$

$\mathbf{S}_k$ → diagonal matrix for different endmembers

Advantage: model the scaling factor for each endmember

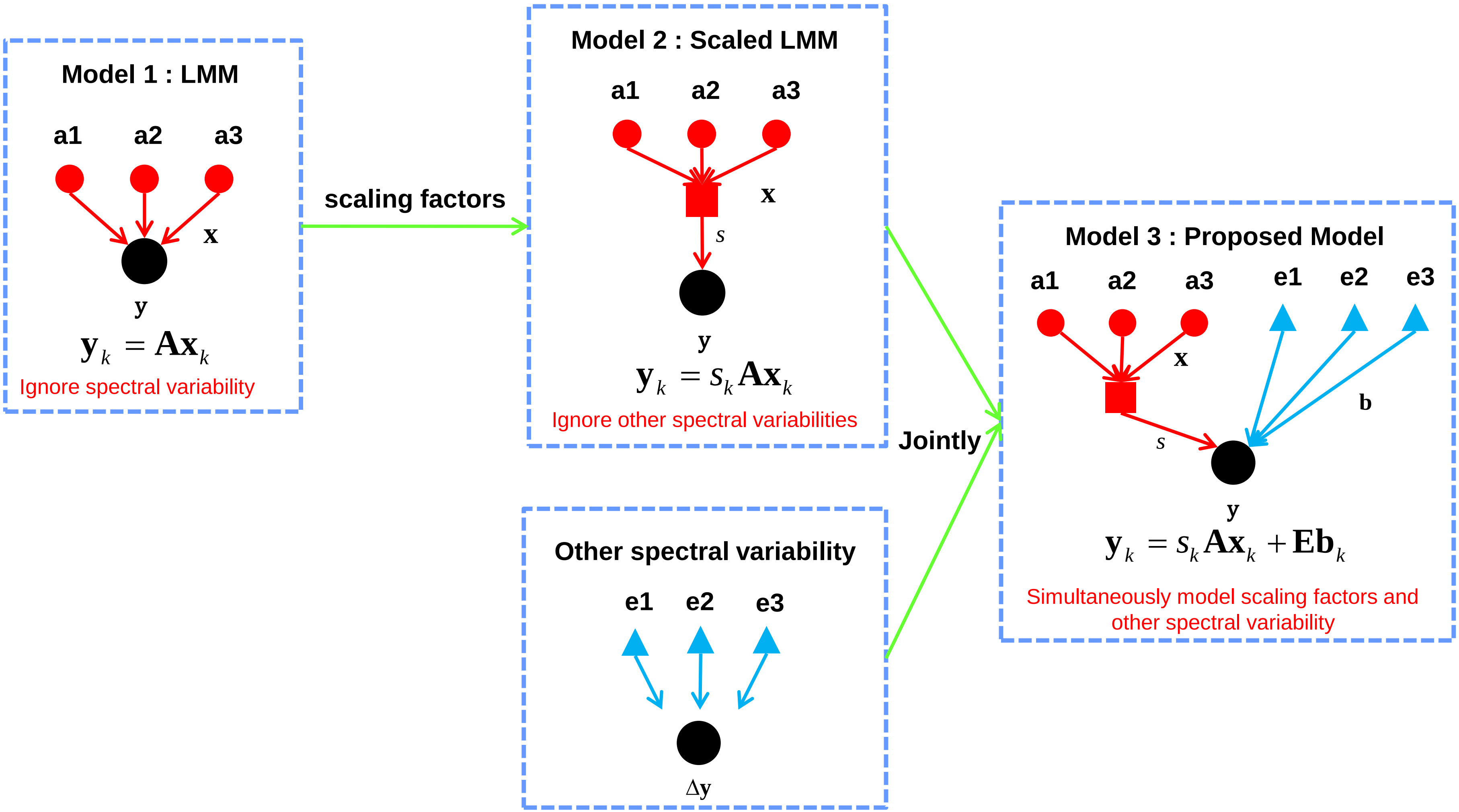
Disadvantage: inaccurate solution: non-convexity
ignore other variabilities

[1] M. A. Veganzones et al, WHISPERS2014. [2] L. Drumetz, et.al, IEEE TIP, 2016.

Roadmap

- Background & Motivation
- **The Proposed Spectral Mixing Model**
- Experiments
- Conclusion & Future Work

Proposed Spectral Mixture Model



Model Formulation and Optimization

The object function of spectral variability dictionary learning can be formulated as

$$\arg\min_{\mathbf{X},\mathbf{B},\mathbf{S},\mathbf{E}} \frac{1}{2}\left\|\mathbf{Y}-\mathbf{AXS}-\mathbf{EB}\right\|_F^2+\alpha\Phi(\mathbf{X})+\beta\Psi(\mathbf{B})+\gamma\Upsilon(\mathbf{E})$$

$$s.t.\quad \mathbf{X}\geq\mathbf{0},\ \mathbf{1}_{P\times 1}^T\mathbf{X}=\mathbf{1}_{1\times N},\ \mathbf{S}\geq\mathbf{0}$$

- Sub-problem (1) : Pixel-wise spectral unmixing

$$\arg\min_{\mathbf{X},\mathbf{B},\mathbf{S}} \sum_{k=1}^{N}\left(\frac{1}{2}\left\|\mathbf{y}_k-s_k\mathbf{A}\mathbf{x}_k-\mathbf{E}\mathbf{b}_k\right\|_2^2\right)+\alpha\Phi(\mathbf{X})+\beta\Psi(\mathbf{B})$$

$$s.t.\quad \mathbf{X}\geq\mathbf{0},\ \mathbf{1}_{P\times 1}^T\mathbf{X}=\mathbf{1}_{1\times N},\ \mathbf{S}\geq\mathbf{0}$$

- Sub-problem (2) : Spectral variability dictionary updating

$$\arg\min_{\mathbf{E}} \frac{1}{2}\left\|\mathbf{Y}-\mathbf{AXS}-\mathbf{EB}\right\|_F^2+\gamma\Upsilon(\mathbf{E})$$

- Regularization terms :

$$\Phi(\mathbf{X})=\frac{1}{2}\left\|\mathbf{X}\right\|_{1,1}$$

$$\Psi(\mathbf{B})=\frac{1}{2}\left\|\mathbf{B}\right\|_F^2$$

$$\Upsilon(\mathbf{E})=\frac{1}{2}\left(\left\|\mathbf{A}^T\mathbf{E}\right\|_F^2+\left\|\mathbf{E}^T\mathbf{E}-\mathbf{I}\right\|_F^2\right)$$

Pseudocode

Learning Spectral Variability Dictionary

Input: Endmember dictionary **(A)**, Mixture spectral signature **(Y)**, and parameters (α , β , γ).

Initialize: Variability dictionary **(E)** , Abundance map **(X)**, Coefficient for Variability dictionary **(B=0)**
Scaling factor **(S=1)**

Main steps:

while *not converged or t > maxIter* **do**

    Dictionary Updating

    fix $\mathbf{X}$, $\mathbf{B}$, $\mathbf{S}$ and update **E** by solving the problem (2) using ADMM

    if conditions are satisfied

    then

        Spectral Unmixing

        fix **E** and update $\mathbf{X}$, $\mathbf{B}$, $\mathbf{S}$ by solving the problem (1) using ADMM

        if conditions are satisfied

        then

            Stop iteration;

    else

        $t=t+1$

    end

end

Output : $\mathbf{X}$, $\mathbf{B}$, $\mathbf{E}$, **S**

Note: $\mathbf{E}$: Random Orthogonal Matrix ; Initial $\mathbf{X}$: obtained by Scaled LMM

Regularization Terms

$$\Phi(\mathbf{X}) = \frac{1}{2}\|\mathbf{X}\|_{1,1}$$

A limited number of materials —promote→ Sparsity

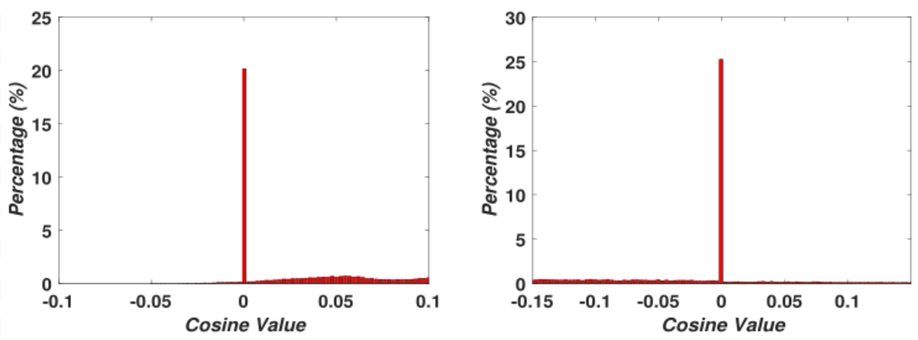
$$\Psi(\mathbf{B}) = \frac{1}{2}\|\mathbf{B}\|_F^2$$

Many atoms in **E** should be activated —promote→ Density

$$\Upsilon(\mathbf{E}) = \frac{1}{2}\left(\|\mathbf{A}^T\mathbf{E}\|_F^2 + \|\mathbf{E}^T\mathbf{E} - \mathbf{I}\|_F^2\right)$$

promote

- ➢ Enough to represent various **SV**
- ➢ Avoid the trivial solution



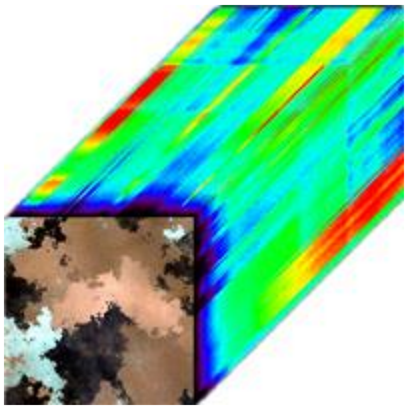
(a) simulated dataset (b) real dataset (Urban)

Statistics of Cosine Value between **A** and **E**

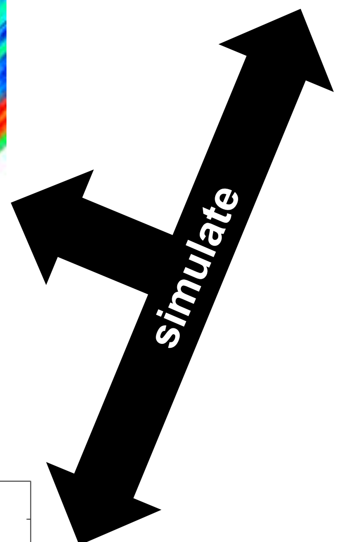
Roadmap

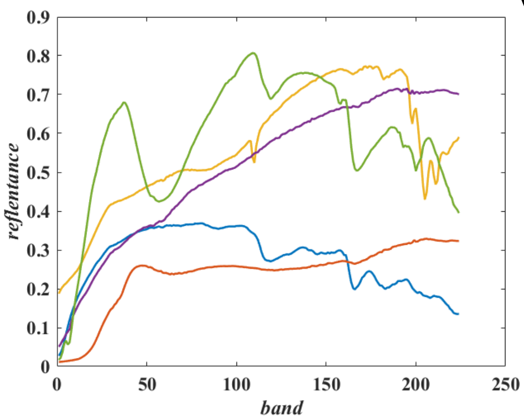
- Background & Motivation
- The Proposed Spectral Mixing Model
- **Experiments**
- Conclusion & Future Work

Abundance Map for Simulated Data

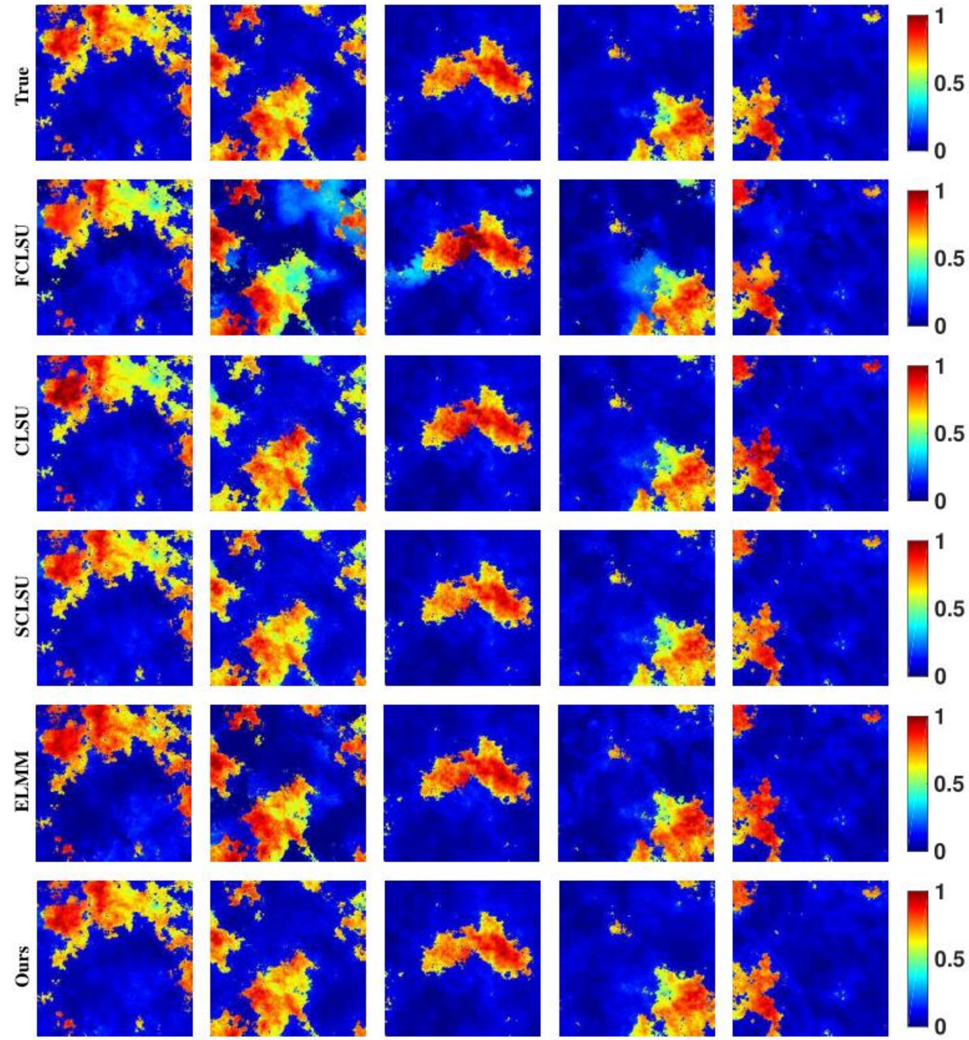


A false color simulated data



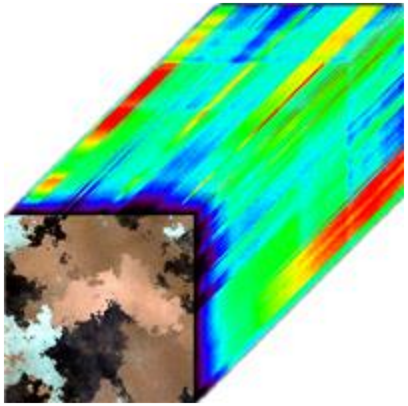


Endmembers used for generating the simulated data

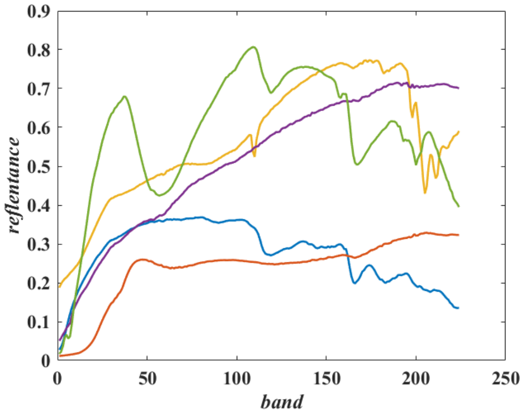


The abundance maps (each column corresponds to one endmember extracted by VCA) and the first row shows the ground truth

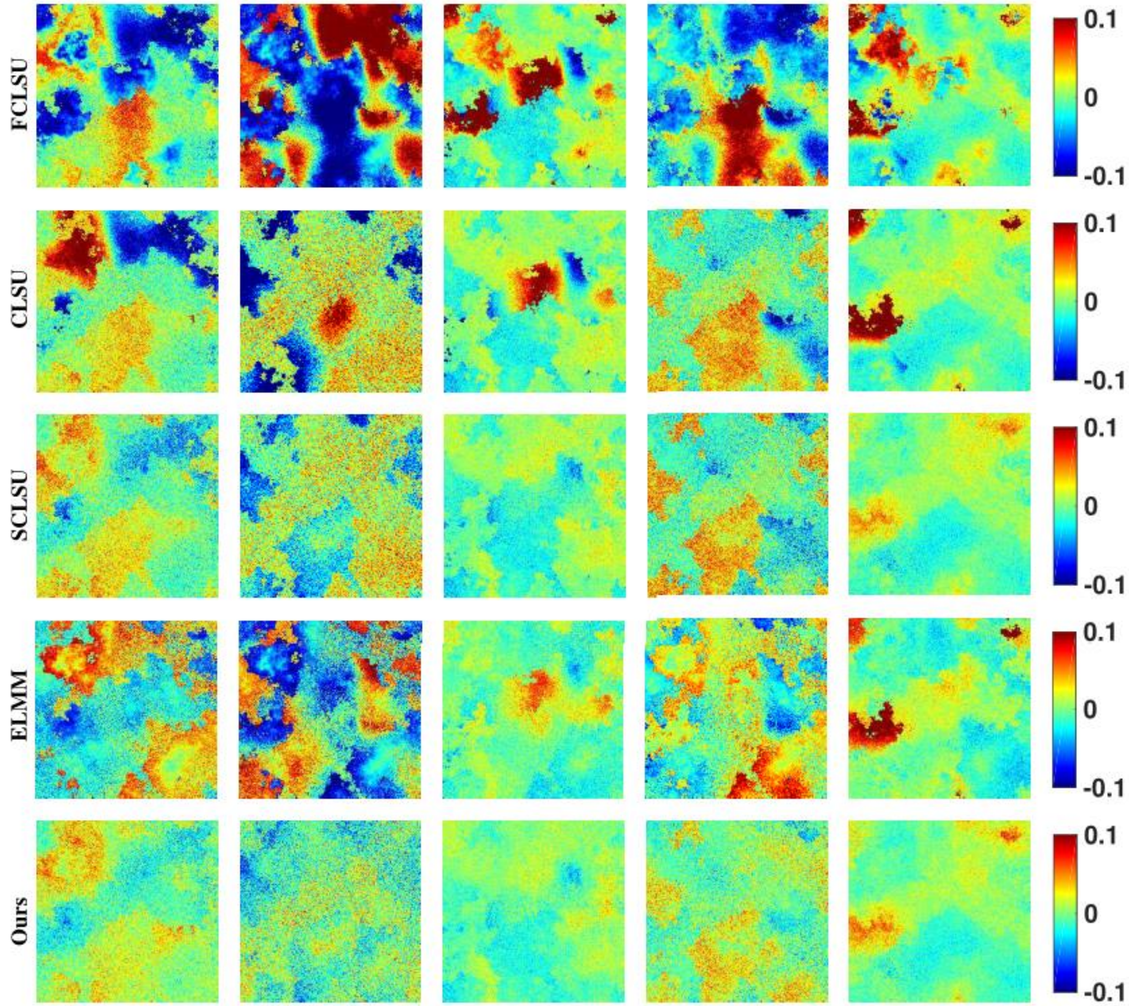
Abundance Difference Map for Simulated Data



A false color simulated data



Endmembers used for generating the simulated data



The abundance difference maps obtained between estimated abundance map and true map (each column corresponds to one endmember extracted by VCA)

Quantity Experiments (Simulated Data)

$$xRMSE = \frac{1}{N}\sum_{k=1}^{N}\sqrt{\frac{1}{P}\sum_{p=1}^{P}\left(\mathbf{x}_{pk} - \hat{\mathbf{x}}_{pk}\right)^2}$$

$$yRMSE = \frac{1}{N}\sum_{k=1}^{N}\sqrt{\frac{1}{L}\sum_{l=1}^{L}\left(\mathbf{y}_{pk} - \hat{\mathbf{y}}_{pk}\right)^2}$$

$$ySAM = \frac{1}{N}\sum_{k=1}^{N}\arccos\left(\frac{\mathbf{y}_k \bullet \hat{\mathbf{y}}_k}{\left\|\mathbf{y}_k\right\| \times \left\|\hat{\mathbf{y}}_k\right\|}\right)$$

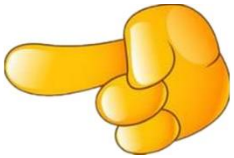
N : the number of pixel

xRMSE : Abundance Root Mean Square Error

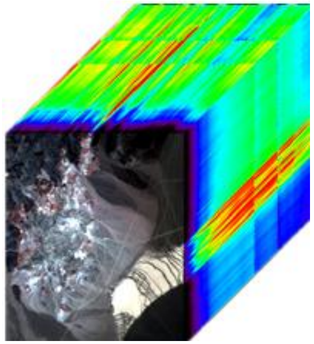
yRMSE : Reconstruction Root Mean Square Error

ySAM : Average Spectral Angle Mapper

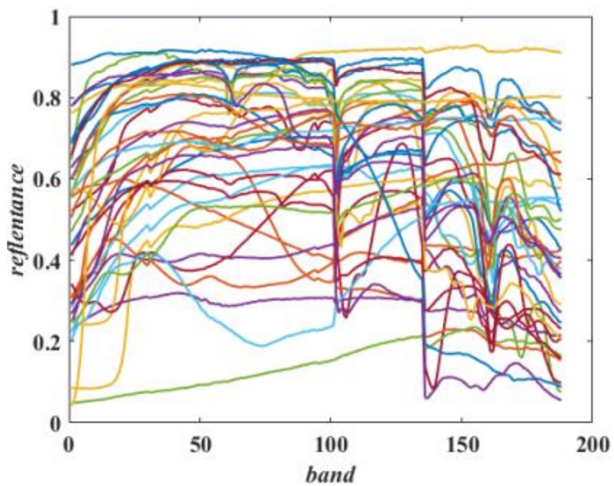
Algorithm	FCLSU	CLSU	CLSU+Sparse	SCLSU	SCLSU+Sparse	ELMM	Ours
xRMSE	0.0524	0.0380	0.0379	0.0251	0.0248	0.0337	0.0206
yRMSE	0.0151	0.0123	0.0127	0.0123	0.0127	0.0088	0.000006
ySAM	1.9600	1.7713	1.7715	1.7713	1.7715	1.2998	0.0007



Abundance Map for Real Data

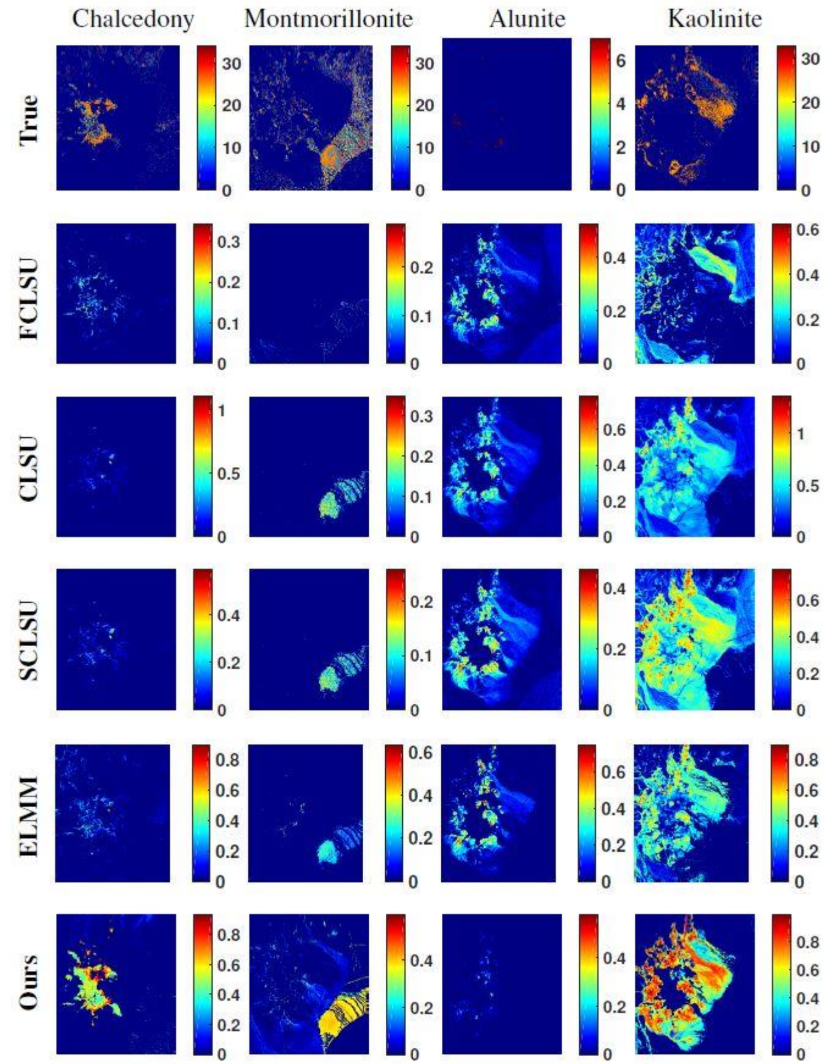


Cuprite Data



Selected Endmembers from Spectral Library *

* Due to the complexity of spectral variability in Cuprite Data, here we get help from spectral library to construct endmember dictionary.



Roadmap

- Background & Motivation
- The Proposed Spectral Mixing Model
- Experiments
- **Conclusion & Future Work**

Conclusion and Future work

Conclusion

1. A novel spectral mixing model is proposed to simultaneously consider the scaling factors (main) and other spectral variabilities.
2. A data-driven dictionary learning method is explored with the low-coherent regularization to design the spectral variability dictionary.
3. An alternating iterative optimization strategy is applied to solve the proposed model using ADMM-based framework.

Future work

1. Spatial regularization should be able to further improve the performance of spectral unmixing.
2. Distributed strategy could be introduced to promote a large-scale spectral unmixing.

Thank you for your attention!