MultiHU-TD: Multifeature Hyperspectral Unmixing Based on Tensor Decomposition

Online Seminar for Tensor Network Reading Group
Montreal Institute for Learning Algorithms (MILA), Quebec, Canada

Mohamad JOUNI

Co-authors: Mauro DALLA MURA, Lucas DRUMETZ, Pierre COMON

Gipsa-lab
Université Grenoble Alpes

November 11th, 2023

Université Grenoble Alpes | gipsa-lab Grenoble | images | parole | signal | automatique | laboratoire | Grenoble INP

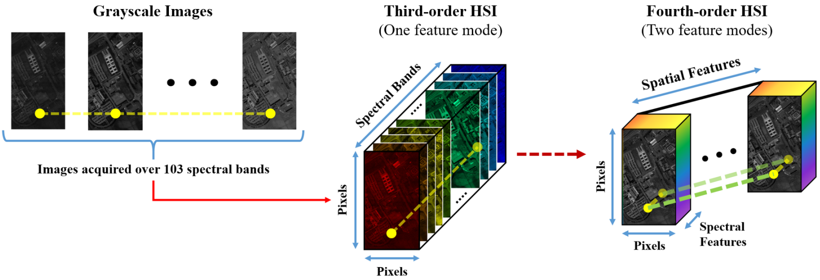
Table of Contents

Introduction

Introduction

Hyperspectral Images (HSI)

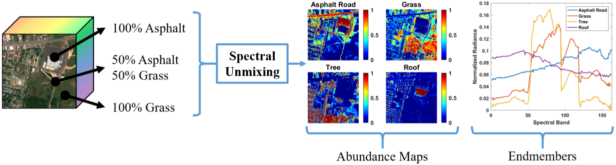
- A **Hyperspectral image (HSI)** consists of a **third-order** arrangement of grayscale images.

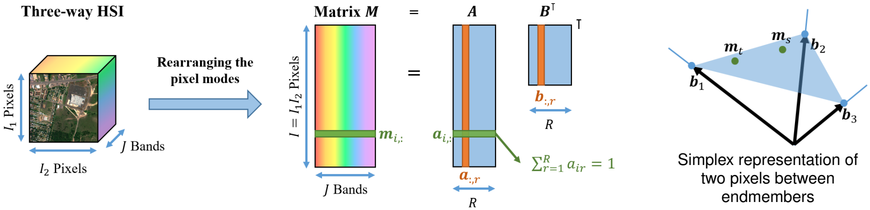


Classical (matrix-based) Hyperspectral Unmixing

Linear Mixing Model (LMM)

- Blind separation of the materials of a scene into **abundance maps** and **endmembers**
- **Linear Mixing Model:** $\boldsymbol{m}_{i,:}^{\mathsf{T}} = \sum_{r=1}^{R} a_{ir} \boldsymbol{b}_{:,r}^{\mathsf{T}} \iff \boldsymbol{M} = \boldsymbol{A}\boldsymbol{B}^{\mathsf{T}}$, with the constraints:
  - Nonnegativity: $\boldsymbol{A} \succeq 0, \;\; \boldsymbol{B} \succeq 0$
  - Abundance Sum-to-one Constraints (ASC): $\sum_{r=1}^{R} a_{ir} = 1$

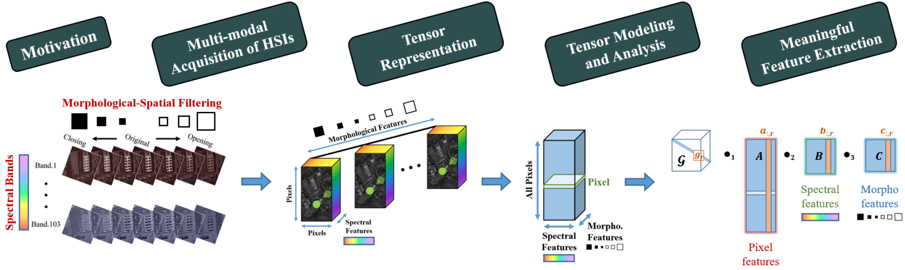




Motivation

Motivation

1. Enhancing **feature extraction** in HSI unmixing with additional features.
2. Conservation of the **multi-feature multi-modal arrangement** of data as **tensors**.
3. **Low-rank** representation of the data.
4. Incorporating the **Abundance Sum-to-one Constraint** (ASC) in tensor decomposition.
5. Flexibility of imposing constraints on each mode separately.
6. Lack of a **generalized and interpretable framework** for such applications.



Proposal

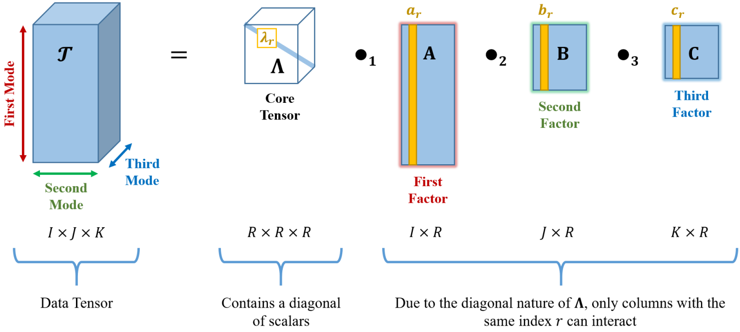
In this work...

1. We propose “MultiHU-TD”, an interpretable methodological framework for **low-rank** “Multi-feature Hyperspectral Unmixing based on Tensor Decomposition”.
2. “MultiHU-TD” is based on the **Canonical Polyadic decomposition** (CPD) and incorporates the **Abundance Sum-to-one Constraint** (ASC).
3. We provide mathematical, physical and graphical **interpretation** of the extracted features.
4. We provide **analogies** with the classical matrix-based spectral unmixing of HSIs.

Canonical Polyadic Decomposition (CPD)

Tucker Form

- CPD reveals the **tensor rank**, usually denoted by R, which is the minimum number of terms for the CPD to hold exact.
- For **feature extraction**, columns with different indices should not interact.



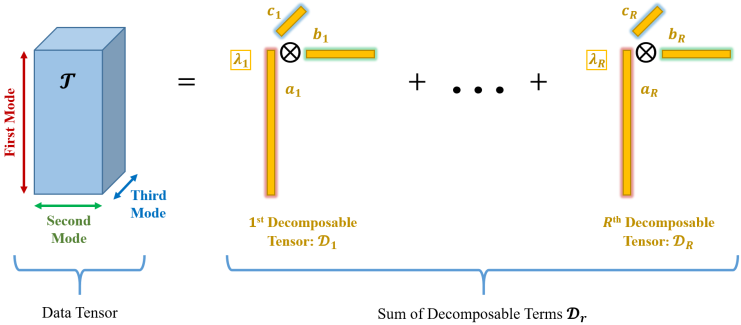
Canonical Polyadic Decomposition (CPD)

Sum of Rank-1 Terms

- The CPD of a rank-R tensor can be written as follows:

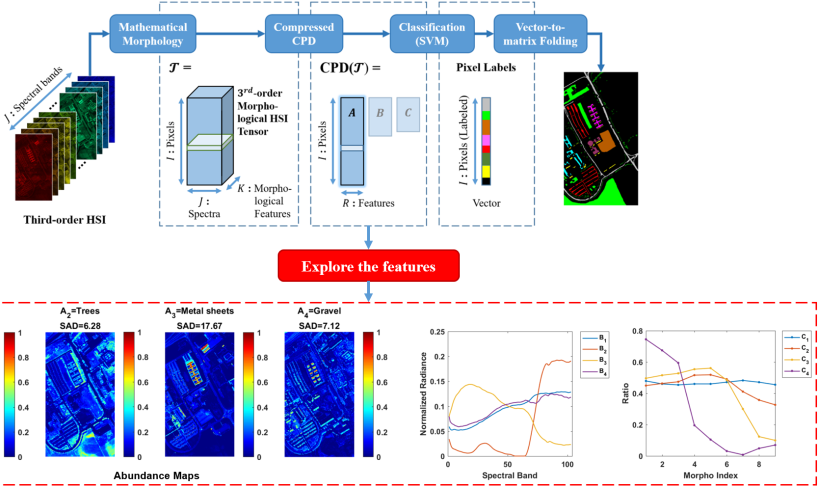
$$\mathcal{T} = \sum_{r=1}^{R} \lambda_{rrr} \left(\mathbf{a}_{:,r} \otimes \mathbf{b}_{:,r} \otimes \mathbf{c}_{:,r} \right) = \sum_{r=1}^{R} \lambda_{rrr} \left(\mathcal{D}_r \right) \quad (1)$$

- Then, $\mathcal{D}_r$ is an N-th order tensor which represents the composition of the physical properties defined by the vectors composing it, good for **material extraction**.



Related Works

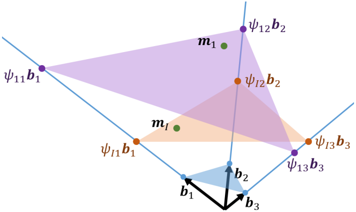
From Multi-feature HSI Supervised Classification... [Jouni et al. 2020]



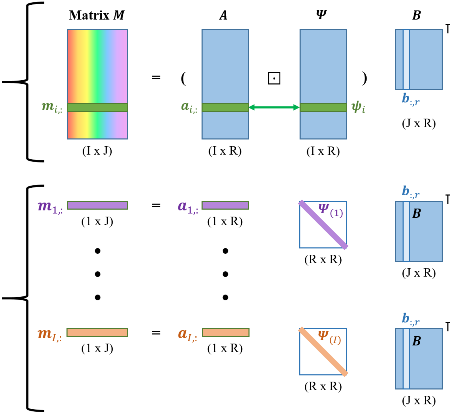
From Matrix-based Spectral Unmixing...

Extended LMM [Drumetz et al. 2016]

- Extended LMM (ELMM) deals with pixel-wise **spectral variability**. For instance:
  $\boldsymbol{m}_i = \sum_{r=1}^{R} a_{ir} \boldsymbol{f}_i(\boldsymbol{b}_r) = \sum_{r=1}^{R} a_{ir} \boldsymbol{b}_r^{(i)} = \sum_{r=1}^{R} a_{ir} \psi_{ir} \boldsymbol{b}_r$
- **We are interested in ELMM since it shares a lot of similarities with CPD.**



ELMM: Simplex representation of two pixels between **scaled versions** of the endmembers



MultiHU-TD

AO-ADMM-ASC with Nonnegativity and Sparsity

Cost Function

- In **CPD**, we impose **nonnegativity** on all factor matrices, and **sparsity** and **ASC** on the abundances. The optimization problem is to solve the cost function:

$$\begin{aligned} &\operatorname*{argmin}_{\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}} \|\mathcal{T} - \boldsymbol{\Lambda} \underset{1}{\bullet} \boldsymbol{A} \underset{2}{\bullet} \boldsymbol{B} \underset{3}{\bullet} \boldsymbol{C}\|_F^2 + \alpha\|\boldsymbol{A}\|_1 \\ &\text{s.t. } \boldsymbol{A} \succeq 0,\ \boldsymbol{B} \succeq 0,\ \boldsymbol{C} \succeq 0,\ \sum_{r=1}^{R} a_{i,r} = 1 \,|_{\forall i \in \{1,\ldots,I\}} \end{aligned} \quad (2)$$

- For the **ASC** solution, set the $(J+1)$-th lateral slice and row vector of $\mathcal{T}$ and $\boldsymbol{B}$ as follows:
  - $\boldsymbol{T}_{:,J+1,K} = \delta \mathbf{1}_I$, i.e., $t_{i,J+1,K} = \delta \quad \forall i \in \{1,\ldots,I\}$
  - $b_{J+1,r} = \delta c_{K,r}^{-1} \quad \forall r \in \{1,\ldots,R\}$,

  which ensures that $\sum_{r=1}^{R} a_{i,r} = 1 \quad \forall i \in \{1,\ldots,I\}$.
- Using AO-ADMM [Huang et al. 2016], the problem boils down to an alternating optimization of **ADMM subproblems** with respect to the factor matrices. For instance, with respect to $\boldsymbol{A}$:

$$\boldsymbol{A} = \operatorname*{argmin}_{\boldsymbol{A}} \frac{1}{2}\|\tilde{\boldsymbol{T}}_{(1)} - \tilde{\boldsymbol{W}}_{(A)}\boldsymbol{A}^{\mathsf{T}}\|_F^2 + \alpha\|\boldsymbol{A}\|_1 \quad \text{s.t. } \boldsymbol{A} \succeq 0 \quad (3)$$

where $\tilde{\boldsymbol{W}}_{(A)} = \tilde{\boldsymbol{B}} \odot \boldsymbol{C}$ represents the Khatri-Rao product [Comon. 2014].

Interpretability: Tensor-based ELMM

CPD and ELMM analogies

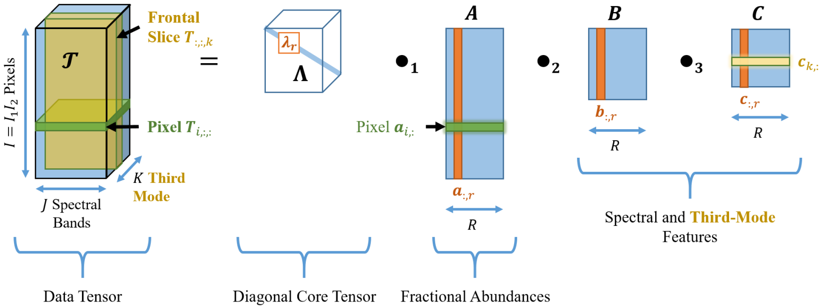
- CPD describes **frontal slice-based** spectral variability:

$$\boldsymbol{T}_{:,:,k} = \boldsymbol{A}\,\mathrm{Diag}\{\boldsymbol{c}_{k,:}\}\boldsymbol{B}^{\mathsf{T}} = \boldsymbol{A}\,\boldsymbol{\Psi}_{(k)}\,\boldsymbol{B}^{\mathsf{T}} = \boldsymbol{A}\,\tilde{f}_k(\boldsymbol{B})^{\mathsf{T}} \iff$$

$$\boldsymbol{t}_{i,:,k} = \sum_{r=1}^{R} a_{ir} c_{kr} \boldsymbol{b}_r = \sum_{r=1}^{R} a_{ir}\, f_k(\boldsymbol{b}_r) = \sum_{r=1}^{R} a_{ir} \boldsymbol{b}_r^{(k)}$$

- ELMM describes **pixel-based** spectral variability:

$$\boldsymbol{m}_i = \sum_{r=1}^{R} a_{ir} \psi_{ir} \boldsymbol{b}_r = \sum_{r=1}^{R} a_{ir}\, f_i(\boldsymbol{b}_r) = \sum_{r=1}^{R} a_{ir} \boldsymbol{b}_r^{(i)}$$



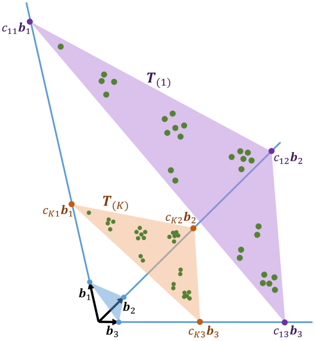
Interpretability: Tensor-based ASC

MultiHU-TD: Generalized Interpretation

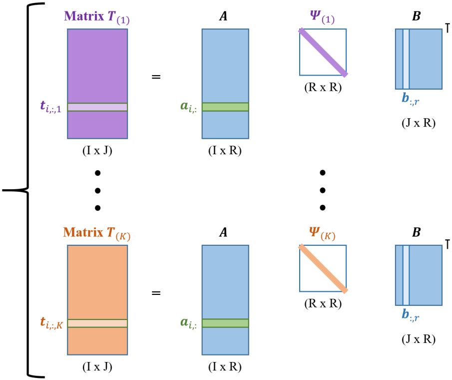
- The frontal slices have **common factors *A* and *B***.
- Each frontal slice can be represented as a simplex. The vertices are formed at the columns of $\boldsymbol{B}^{(k)}$.
- ***A*** and ***B*** are independent of the third-mode differences in the hyperspectral scene along the slices:

$$\boldsymbol{M}^{(\mathrm{CPD})} = \boldsymbol{A}\,\boldsymbol{B}^{\mathsf{T}} \tag{4}$$

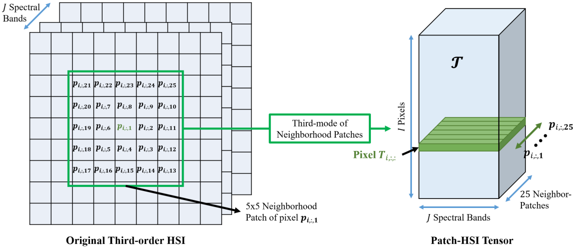
- The rows of ***C*** encode the third-mode variabilities, as ***A*** and ***B*** factorize the abundance and spectral features.



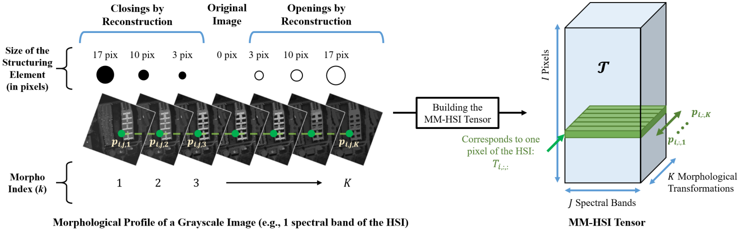
MultiHU-TD: There are K simplexes (as many as the frontal slices).



Patches and Mathematical Morphology

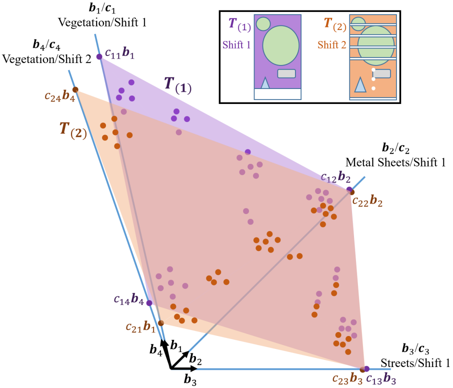


An illustration of constructing a 5 × 5 Patch-HSI tensors based on [Veganzones et al. 2016]

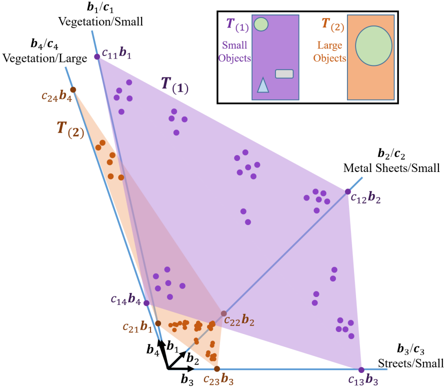


An illustration of constructing a MM-HSI tensors based on [Jouni et al. 2020]

Geometric Interpretation of the Decomposed Factors



Patches



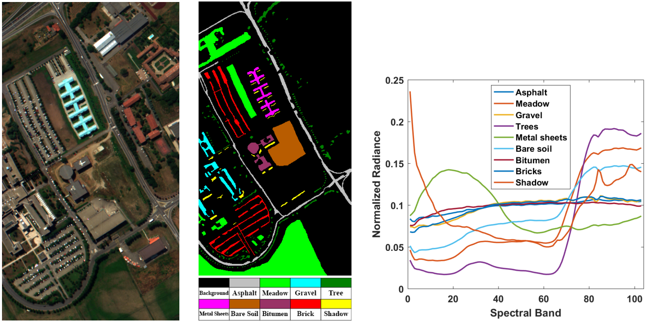
Mathematical Morphology

Experimental Discussion

Real HSI Data Sets[1]

HSI Data Sets

- We use two real HSIs for testing, composed of a landscape of buildings, streets and vegetation of different materials and sizes.
  - Pavia University (on the left), with dimensions 610 × 340 × 103.
  - Urban (on the right), with dimensions 307 × 307 × 162.



Figures: Pavia HSI in false colors and its spatial groundtruth (GT)

[1] We also test this framework using a synthetic HSI. The details are in the paper appendix.

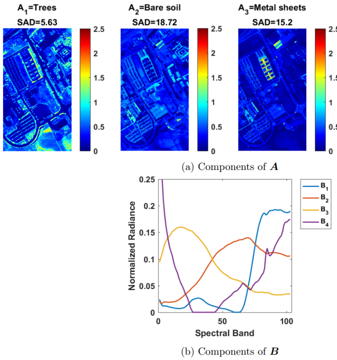
LMM (NMF) vs MultiHU-TD (Rank-4 AO-ADMM-ASC)

Qualitative Evaluation

- LMM does not highlight all the features.
- MultiHU-TD with MM better highlights the features (shadow, asphalt, etc).

	Trees	Bare Soil	Metal Sheets	Shadow
LMM	5.63	18.72	15.2	25.49
MultiHU-TD	6.28	7.95	17.67	7.12*

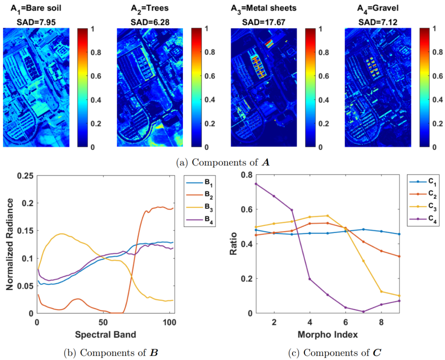
Table: Spectral Angular Distance (SAD), in degrees



(a) Components of $\boldsymbol{A}$

(b) Components of $\boldsymbol{B}$

Pavia. NMF results with ASC



(a) Components of $\boldsymbol{A}$

(b) Components of $\boldsymbol{B}$

(c) Components of $\boldsymbol{C}$

MM-HSI NCPD results for $R = 4$

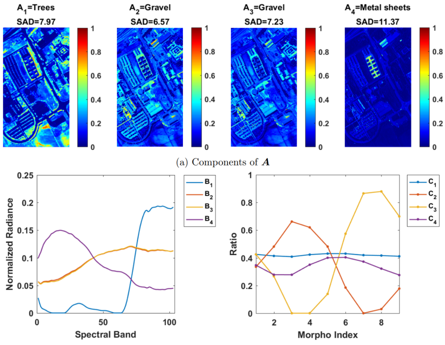
Patches vs MM Features with $R = 4$

Qualitative Evaluation

- Patch MultiHU-TD “replicates” some features.
- MM MultiHU-TD highlights features based on morphological properties (scale, brightness).

	Trees	Bare Soil	Metal Sheets	Gravel*
Patches	7.97	6.57	11.37	6.57*
MM	6.28	7.95	17.67	7.12*

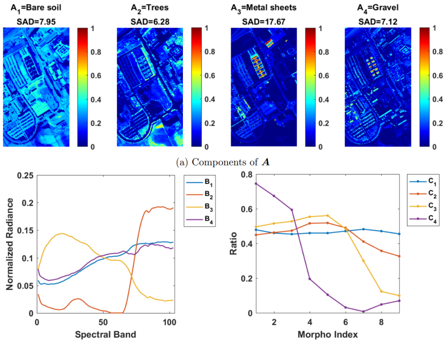
Table: Spectral Angular Distance (SAD), in degrees



(a) Components of $\boldsymbol{A}$

(b) Components of $\boldsymbol{B}$ (c) Components of $\boldsymbol{C}$

Patch-HSI NCPD results for $R = 4$

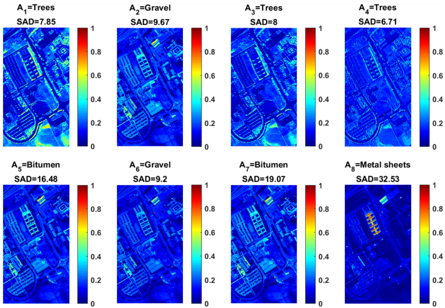


(a) Components of $\boldsymbol{A}$

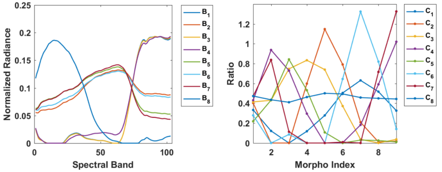
(b) Components of $\boldsymbol{B}$ (c) Components of $\boldsymbol{C}$

MM-HSI NCPD results for $R = 4$

Patches vs MM Features with $R = 8$

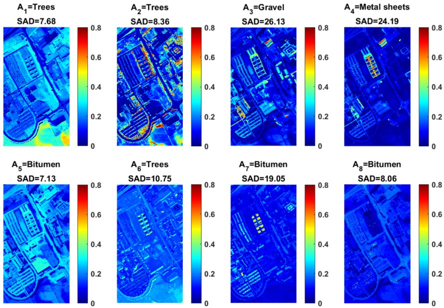


(a) Components of $\boldsymbol{A}$

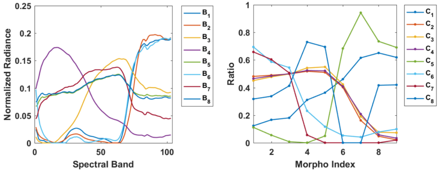


(b) Components of $\boldsymbol{B}$ (c) Components of $\boldsymbol{C}$

Patch-HSI NCPD results for $R = 8$



(a) Components of $\boldsymbol{A}$



(b) Components of $\boldsymbol{B}$ (c) Components of $\boldsymbol{C}$

NCPD results with ASC imposed

Conclusion

General Conclusions

Conclusions

- We proposed **MultiHU-TD**, a generalized and interpretable framework for multi-feature HSIs using tensor decomposition.
- We incorporated **Abundance Sum-to-one Constraints** in a tensor-based multi-feature blind source separation problem.
- MultiHU-TD **conserves the low-rankness** of the data by rearranging the modes of pixels. Incorporating spatial information as features **conserves the relevant neighborhood information** without losing the multi-modal data structure.
- We discussed the methodological and applicative aspects of multi-feature unmixing by:
  - Establishing **mathematical**, **graphical**, and **geometrical** analogies between matrix- and tensor-based source separation, and between the physical LMM and CPD.
  - Stressing the importance of using **physically meaningful features** (such as MM)

Perspectives

Perspectives

- Subspace representations of the factor matrices
- Exploring other types of tensor decomposition such as Block Term Decomposition
- Exploring other types of multi-feature representations of HSIs
- Semi-supervised classification
- Extend the study to multivariate function representations as tensors

Where to go from here?

- Subspace Learning of latent spaces
- Machine Learning for tensor analysis
- Tensor Networks for big multi-modal HSI representations

Thanks!

Thank you for your attention!

Example: Tensor vs Matrix Analysis

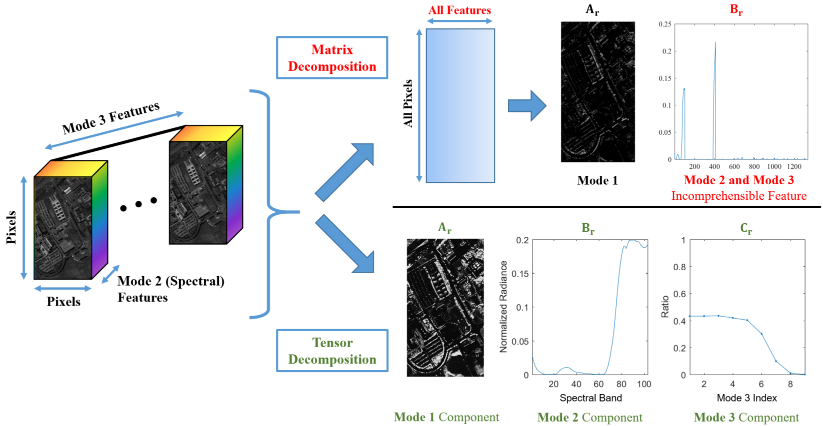
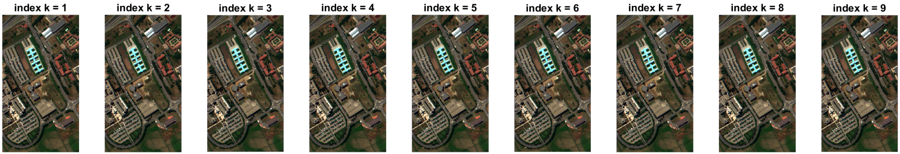


Figure: Example of the difference between matrix and tensor techniques given the same data set and physical component

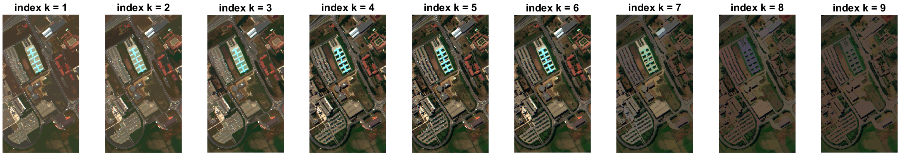
Example: Patch-HSI and MM-HSI Tensors

Physical Interpretation

- **Patch-HSI Tensors:** The frontal slices are slightly-shifted versions of each others.
- **MM-HSI Tensors:** The frontal slices represent physical spatial features.



Pavia HSI: (3×3)-Patch



Pavia HSI: Mathematical Morphology [Jouni et al. 2020]