COMMON MODE PATTERNS FOR SUPERVISED TENSOR SUBSPACE LEARNING

International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2019)

K.Makantasis[1], A.Doulamis[2], N.Doulamis[2], A.Voulodimos[3]

[1]School of Production Engineering and Management, Technical University of Crete, Chania, Greece
[2]School of Rural and Surveying Engineering, National Technical University of Athens, Athens, Greece
[4]Department of Informatics and Computer Engineering, University of West Attica, Athens, Greece

Outline

High-order Data

Supervised Tensor Subspace Learning

Problem Formulation

CMP Algorithm

Results

Outline

High-order Data

Supervised Tensor Subspace Learning

Problem Formulation

CMP Algorithm

Results

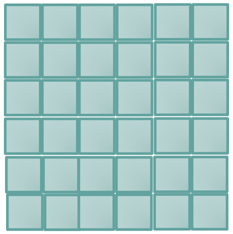
High-order data

The order of the data

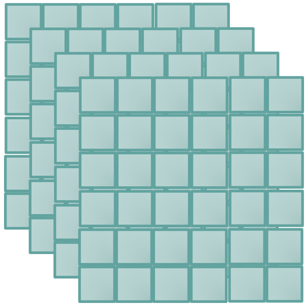
Zero order → scalars

First-order → vectors

Second-order → matrices



Third-order → 3d cubes

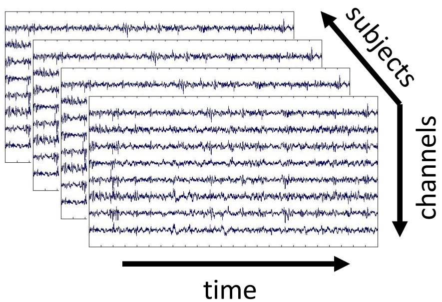


High-order data

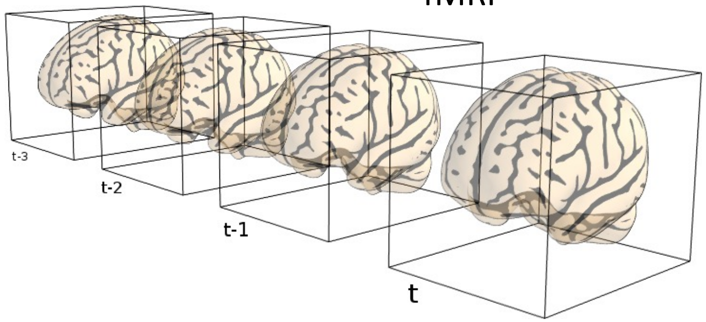
Advances in sensing technologies

Generation of large amounts of high-order data

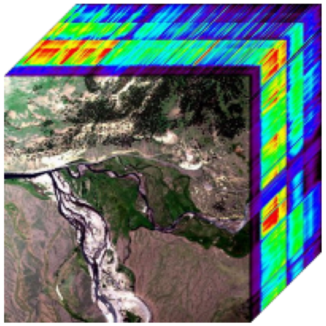
EEG



fMRI



Hyperspectral



Outline

High-order Data

Supervised Tensor Subspace Learning

Problem Formulation

CMP Algorithm

Results

Supervised Tensor Subspace Learning

- Our algorithm is motivated by Common Spatial Patterns
  - A modification of Karhunen-Loeve expansion
- Extract features that increase inter-class separability
  - A modification of Multilinear Principal Component Analysis

Supervised Tensor Subspace Learning

- Our algorithm is motivated by Common Spatial Patterns
  - A modification of Karhunen-Loeve expansion
- Extract features that increase inter-class separability
  - A modification of Multilinear Principal Component Analysis

Our method can be used for **Supervised Dimensionality Reduction** of Tensor objects

- Project data to a lower dimension space
- The explained variance of each class is maximized

Supervised Tensor Subspace Learning

To be more rigorous….

- We assume that the **explained variance** of a set of tensor objects $\{\mathcal{A}_m\}_{m=1}^{M} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ is given by

$$\Psi_{\mathcal{A}} = \frac{1}{M} \sum_m \|\mathcal{A}_m - \bar{\mathcal{A}}\|_{\mathrm{F}}^2$$

Supervised Tensor Subspace Learning

To be more rigorous....

- The projection of $\mathcal{A}_m$ to another tensor $\mathcal{S}_m \in \mathbb{R}^{P_1 \times \cdots \times P_N}$, with $P_n < I_n$ is given by

$$\mathcal{S}_m = \mathcal{A}_m \times_1 U_1^T \times_2 \cdots \times_N U_N^T$$

with $U_n \in \mathbb{R}^{I_n \times P_n}$.

Supervised Tensor Subspace Learning

To be more rigorous....

- The projection of $\mathcal{A}_m$ to another tensor $\mathcal{S}_m \in \mathbb{R}^{P_1 \times \cdots \times P_N}$, with $P_n < I_n$ is given by
$$\mathcal{S}_m = \mathcal{A}_m \times_1 U_1^T \times_2 \cdots \times_N U_N^T$$
with $U_n \in \mathbb{R}^{I_n \times P_n}$.
- Using a matrix representation (mode-*n* matricization) the same projection is given by
$$S_{(n),m} = U_n^T \cdot A_{(n),m} U_{\Phi_n}$$
with
$$U_{\Phi_n} = U_{n+1} \otimes \cdots \otimes U_N \otimes U_1 \otimes \cdots \otimes U_{n-1}$$

Outline

High-order Data

Supervised Tensor Subspace Learning

Problem Formulation

CMP Algorithm

Results

Problem Formulation

Now we are in a position to define the problem that we are trying to address…

Problem: *Estimate set* $\{\widetilde{U}_n\}_{n=1}^{N}$ of projection matrices that satisfy

$$\{\widetilde{U}_n\}_{n=1}^{N} = \arg \max_{\mathrm{U}_1,\cdots,U_N} \Psi_{\mathcal{S}^{(i)}}$$

for both classes, i.e., $i = 1,2$*, such that the projected samples that belong to different classes will not share common important features.*

Problem Formulation

Now we are in a position to define the problem that we are trying

to address…

Problem: *Estimate set* $\left\{\widetilde{U}_n\right\}_{n=1}^{N}$ of projection matrices that satisfy

$$\left\{\widetilde{U}_n\right\}_{n=1}^{N} = \arg \max_{\mathrm{U}_1,\cdots,U_N} \Psi_{\mathcal{S}^{(i)}}$$

for both classes, i.e., $i = 1{,}2$, *such that the projected samples that belong to different classes will not share common important features.*

For solving the problem, first, we have to *sphere* the data. After sphering the data…

Outline

High-order Data

Supervised Tensor Subspace Learning

Problem Formulation

CMP Algorithm

Results

CMP Algorithm

… we can proceed with the CMP algorithm.

Theorem: *Let* $\left\{\widetilde{U}_n\right\}_{n=1}^{N}$ *be the solution to our Problem. Then, given all matrices* $\widetilde{U}_1, \cdots, \widetilde{U}_{n-1}, \widetilde{U}_{n+1}, \cdots \widetilde{U}_N$ *the matrix* $\widetilde{U}_n$ *consists of the* $P_n/2$ *eigenvectors corresponding to the largest eigenvalues of* $\Phi_n^{(1)}$ *and the* $P_n/2$ *eigenvectors corresponding to the largest eigenvalues of* $\Phi_n^{(2)}$*.*

$$\Phi_n^{(i)} = \frac{1}{M_i} \sum_{m=i}^{M_i} \left(A_{(n),m}^{(i)} - \bar{A}_{(n)} \right) \cdot \widetilde{U}_{\Phi_n} \cdot \widetilde{U}_{\Phi_n}^T \cdot \left(A_{(n),m}^{(i)} - \bar{A}_{(n)} \right)^T$$

CMP Algorithm

Based on the previous Theorem the CMP algorithm is as follows

1. Sphere the data
2. Initialize $\widetilde{U}_n = I_n$
3. **Repeat**

   for $n = 1, \cdots, N$

   set the columns of $\widetilde{U}_n$ equal to the eigenvectors of $\Phi_n^{(i)}$

   Until *termination criteria are met*
4. For each U_n keep the $P_n/2$ eigenvectors with the largest eigenvalues and $P_n/2$ eigenvectors with the smallest eigenvalues.

Outline

High-order Data

Supervised Tensor Subspace Learning

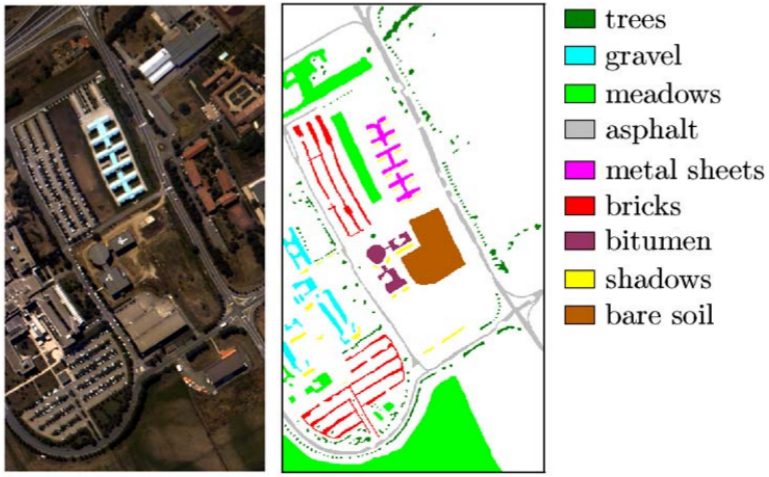
Problem Formulation

CMP Algorithm

Results

Results

- We evaluate our method on a Pavia university hyperspectral dataset.
- The objective is to classify depicted materials into 2 classes
  - Manmade objects (*asphalt, metal sheets, bricks and bitumen*)
  - Non manmade objects



The Pavia University dataset

Results

- Split the image into overlapping windows of size $5 \times 5 \times c$
  - c corresponds to the number of bands
- We assume that the label of a patch should be the same with the label of its central pixel
  - This assumption is valid for the vast majority of the pixels
- Reduce the spectral dimension of the patches
  - MPCA with 26 and 10 spectral features
  - CMP with 26, 13, 10, and 5 spectral features
- For the classification we used CNN, Rank-1 FNN and Rank-1 Tensor regression

Results

Overall classification accuracy results (%)

	CNN	Rank-1 FNN	Rank-1 TR
MPCA - 26	85.05	86.80	77.56
CMP - 26	**90.41**	**91.25**	**77.96**
CMP - 13	88.57	88.67	76.90
MPCA - 10	83.49	84.39	**77.59**
CMP - 10	**88.23**	**88.31**	77.52
CMP - 5	86.76	86.27	76.08

- CMP is more efficient than MPCA
  - Exploits labels' information during dimensionality reduction
  - Almost for all employed classifiers

Thanks

for

your attention