Unsupervised Feature Extraction for Hyperspectral Images Using Combined Low Rank Representation and Locally Linear Embedding

Mengdi Wang\textsuperscript{1}    Jing Yu\textsuperscript{2}    Lijuan Niu\textsuperscript{3}    Weidong Sun\textsuperscript{1}

\textsuperscript{1}Tsinghua University, China
wangmd12@mails.tsinghua.edu.cn; wdsun@tsinghua.edu.cn

\textsuperscript{2}Beijing University of Technology, China
jing.yu@bjut.edu.cn

\textsuperscript{3}Chinese Academy of Medical Sciences, China

ICASSP 2017
Outline

1. Introduction
2. The Proposed Method
3. Experiments and Discussions
4. Conclusion
1 Introduction

2 The Proposed Method

3 Experiments and Discussions

4 Conclusion
What is hyperspectral images?

- Hyperspectral images (HSIs)
  - captured by the remote sensing platforms
  - contain hundreds of bands across the spectral dimension

- can provide not only spatial but also spectral information of the land-covers in a scene
Applications and Problems of HSIs

- Applications of HSIs
  - agriculture
  - environment
  - monitoring
  - food safety
  - medicine
  - mineralogy
  - etc.

- Problems
  - hundreds of bands
  - curse of dimensionality
Applications and Problems of HSIs

- Applications of HSIs
  - agriculture
  - environment
  - monitoring
  - food safety
  - medicine
  - mineralogy
  - etc.

- Problems
  - hundreds of bands
  - curse of dimensionality

Feature extraction and dimension reduction for HSIs
Related Works

- **Widely used methods**
  - PCA, ICA, MNF
  - LLE, ISOMAP, Laplacian Eigenmap

- **HSI-specified methods based on the endmember mixing nature**
  - VCA (vertex component analysis)
  - MVC-NMF (minimum volume constrained nonnegative matrix factorization)

- **Recently works**
  - OTVCA (orthogonal total variation component analysis), *TGRS’16*
  - IR (Intrinsic Representation), *TGRS’16*
Structure of the spectral space in HSIs

- The spectral space in HSIs can be divided into several subspaces according to the land-covers $\{S_c\}_{c=1}^{C}$, and $S_{c_1} \cap S_{c_2} (c_1 \neq c_2) = \emptyset$

- The spectral space $S$ can be represented by $S = \bigcup_{c=1}^{C} S_c$

- The spectral vectors in each class share high similarity, thus $S_c$ should be low-rank.

- The spectral space in HSIs is a union of multiple low-rank subspaces.
Structure of the spectral space in HSIs

- The spectral space in HSIs can be divided into several subspaces according to the land-covers $\{S_c\}_{c=1}^C$, and $S_{c_1} \cap S_{c_2} (c_1 \neq c_2) = \emptyset$
- The spectral space $S$ can be represented by $S = \bigcup_{c=1}^C S_c$
- The spectral vectors in each class share high similarity, thus $S_c$ should be low-rank.
- The spectral space in HSIs is a union of multiple low-rank subspaces.

An informative data representation when used for FE should:

1. preserve the subspace-inherent structures
2. minimize the inter-subspace components
Framework of LRR

- Assume $X \in \bigcup_{c=1}^{C} S_c$ and $X = [X_1, X_2, \ldots, X_C]$, $X_c \in S_c$
- If there is a structured dictionary $A = [A_1, A_2, \ldots, A_C]$, $A_c \in S_c$
- Then if $X$ is modelled as,

  $$\min_{Z} \text{rank}(Z) \quad \text{s.t.} \quad X = AZ$$

  Rank constraint on $Z$ will lead to

  $$Z^* = \begin{bmatrix} Z_1^* & 0 & 0 & 0 \\ 0 & Z_2^* & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & Z_C^* \end{bmatrix},$$

- Dictionary selection: $A = X$
Unsupervised FE using LRR

FE model:

\[
\min_{Z,E} \text{rank}(Z) + \lambda \|E\|_{2,0} \quad \text{s.t. } X = XZ + E
\]

- E is constituted by the vectors that has the inter-subspace components
- Number of such vectors should be small
- Thus the column-sparse constraint $\ell_{2,0}$ norm is used.

Convex model,

\[
\min_{Z,E} \|Z\|_* + \lambda \|E\|_1 \quad \text{s.t. } X = XZ + E
\]

Solved using the inexact augmented Lagrange multiplier (IALM) method.
Spatial constraint using LLE

Introduce the spatial similarity in the FE procedure based on locally linear embedding (LLE)

1. Select the neighbors
2. Construct the topology structure within the neighborhood in the original feature space
3. Preserve this topology structure in the extracted feature space
Procedure of LLE

1. Select the neighbors

2. Construct the topology structure using the quadratic fit,

\[
\{ W_{ij} \} = \arg \min_{W_{ij}} \| X_i - \sum_j W_{ij} X_j^{(i)} \|_F^2
\]

3. Preserve this topology in the extracted feature space

\[
L = \sum_i \| Y_i - \sum_j W_{ij} Y_j^{(i)} \|_F^2 = \text{Tr} \left( Y (I - W)^T (I - W) Y^T \right)
\]

\([W]_{ij}\) being \(W_{ij}\) if \(X_j\) is a neighbour of \(X_i\) and 0 if not
Combine LRR and LLE for unsupervised FE

- LRR framework

\[ \min_{Z,E} \|Z\|_* + \lambda \|E\|_1 \quad \text{s.t.} \quad X = XZ + E \]

- \( Z_i \) is actually the transform of \( X_i \) in the self-representation domain, therefore \( Z_i \) should preserve the same neighborhood topology structure as \( X_i \),

\[ \text{Tr} \left( Z (I - W)^T (I - W) Z^T \right) \]

- The combined LRR and LLE for unsupervised FE is,

\[ \min_{Z,E} \|Z\|_* + \lambda \|E\|_1 + \frac{\beta}{2} \text{Tr} \left( Z (I - W)^T (I - W) Z^T \right) \quad \text{s.t.} \quad X = XZ + E \]
Combine LRR and LLE for unsupervised FE

The combined LRR and LLE for unsupervised FE is,

$$\min_{Z,E} \|Z\|_* + \lambda \|E\|_1 + \frac{\beta}{2} \text{Tr} \left( Z(I - W)^T (I - W) Z^T \right)$$

s.t. \( X = XZ + E \)

- The structural extracted features are \( \hat{X} = XZ^* \)
- The dimension remains unchanged, so the PCA is adopted to reduce the dimension.
1 Introduction

2 The Proposed Method

3 Experiments and Discussions

4 Conclusion
Experiments set-up

- **Evaluation way**
  - The following classification task is used as evaluation way
  - Support vector machine (SVM) with the radial basis function (RBF).

- **Datasets**
  - AVIRIS data: *Indian Pines*, $145 \times 145 \times 200$
  - ROSIS data: *Pavia University*, $610 \times 340 \times 103$

- **Compared methods**
  - PCA, ICA
  - MVC-NMF (TGRS’07)
  - IR (TGRS’16)
  - LLE
  - LRR
Indian Pines

(a) False color

(b) Ground truth

- Alfalfa
- Corn-mintill
- Grass-pasture
- Grass-P.-M.
- Oats
- Soybean-mintill
- Soybean-clean
- Stone-S.-T.
- Corn-notill
- Corn
- Grass-trees
- Hay-windrowed
- Soybean-notill
- Wheat
- Woods
- Buildings-G.-T.-D.
Pavia University

(a) False color
(b) ground truth

- Asphalt
- Meadows
- Gravel
- Trees
- Painted metal sheets
- Bare Soil
- Bitumen
- Bricks
- Shadows
## Classification results

<table>
<thead>
<tr>
<th>Method</th>
<th>Indian Pines</th>
<th></th>
<th>Pavia University</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reduce dimension: 20</td>
<td>Training set: 10%</td>
<td>Reduce dimension: 15</td>
<td>Training set: 1%</td>
</tr>
<tr>
<td>OA</td>
<td>AA</td>
<td>kappa</td>
<td>OA</td>
<td>AA</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><strong>original</strong></td>
<td>82.76</td>
<td>80.76</td>
<td>0.8034</td>
<td>88.31</td>
</tr>
<tr>
<td><strong>PCA</strong></td>
<td>79.95</td>
<td>79.87</td>
<td>0.7712</td>
<td>74.47</td>
</tr>
<tr>
<td><strong>ICA</strong></td>
<td>74.27</td>
<td>70.71</td>
<td>0.7057</td>
<td>83.27</td>
</tr>
<tr>
<td><strong>MVC-NMF</strong></td>
<td>74.04</td>
<td>71.12</td>
<td>0.7023</td>
<td>82.96</td>
</tr>
<tr>
<td><strong>LLE</strong></td>
<td>79.76</td>
<td>77.47</td>
<td>0.7694</td>
<td>87.77</td>
</tr>
<tr>
<td><strong>LRR</strong></td>
<td>82.34</td>
<td>78.47</td>
<td>0.7984</td>
<td>91.22</td>
</tr>
<tr>
<td><strong>IR</strong></td>
<td>88.5</td>
<td>88.1</td>
<td>0.869</td>
<td>93.1</td>
</tr>
<tr>
<td><strong>LRR_LLE</strong></td>
<td><strong>94.13</strong></td>
<td><strong>93.30</strong></td>
<td><strong>0.9330</strong></td>
<td><strong>95.03</strong></td>
</tr>
</tbody>
</table>
Classification results w.r.t. feature dimension

Indian Pines
Number of training sets is fixed.

(a) OA

(b) AA
Classification results w.r.t. number of training samples

Indian Pines
Reduced dimension is fixed.

![Graph (a) OA](image)

![Graph (b) AA](image)

(a) OA

(b) AA
We proposed a novel unsupervised feature extraction method using combined LRR and LLE:

- LRR is capable to structurally represent the union spectral space of multiple low-rank subspaces, therefore can help preserve the subspace-inherit components;
- LLE is a nonlinear dimension reduction method, help to preserve the locally geometric manifold in the spatial domain;
- The combination model can simultaneously employ the spectral correlation and the locally spatial correlation information during the FE procedure.

Experiments with a following classification task using SVM show that the proposed method LRR_LLE outperforms the state-of-art methods when used for unsupervised FE in HSIs.
Thank you

wangmd12@mails.tsinghua.edu.cn
jing.yu@bjut.edu.cn
wdsun@tsinghua.edu.cn