Band Selection from Statistical Wavelet Models

Siwei Feng, Yuki Itoh, Mario Parente, Marco F. Duarte
Motivation

- Hyperspectral remote sensors
  - Larger data volume than multispectral remote sensors
  - Problems in data computation, storage, and transmission (*Du et al. 2009*)
  - Potential (significant) redundancy in adjacent bands (*Jimenez et al. 1998*)

- Feature extraction (PCA, LDA, etc)
  - Potential high computational cost
  - Changes the original data representation

- Feature selection (Band Selection)
  - Select a subset of features (spectral bands in this work) (*Kononenko 1994, Huang et al. 2005, Guo et al. 2006, Yang et al. 2011*)
  - Avoid high computation cost
  - Performance measured by classification results on testing samples

Image from: https://en.wikipedia.org/wiki/Hyperspectral_imaging
Overall Band Selection Scheme Description

- Measure and rank the priority of bands in a training dataset (band ranking)
  - Feature design (training samples -> multiscale discrete labels)
    - Use discrete multiscale labels to encode the discriminating spectral information in training samples and abandon the uninformative parts
  - Feature integration (samples -> class)
    - Integrate the obtained labels for all samples in a certain class to get the discriminating spectral information at different bands for that class
  - Band priority criterion
    - Set a criterion by using the integrated labels for each class to measure and rank the priority of each method

- Select the top ranked bands from the training set
  - Pre-defined number of selected bands

- Use the selected bands to classify unseen samples
Band Priority (Feature Design)

1. Wavelet Transform (Undecimated)

Training Data

Wavelet Coefficient Matrices (Haar Wavelet)
Band Priority (Feature Design)

- NHMC Model Parameters (*Feng et al. 2014*)
  - Capture both *compression* and *persistence* properties of wavelet coefficients
  - Estimated through EM algorithm

Gaussian Mixture Model
Band Priority (Feature Design)

Wavelet Coefficient Matrices (Haar Wavelet)

(2). Model Training

NHMC Model

Sample State Label Matrices

(3). Label Computing (Viterbi Algorithm)

# Sample

# Scale

# Band

# Sample

# Scale

# Band
Band Priority (Feature Integration)

Sample State Label Matrices

(4). Sample State Label Matrices Combining in Terms of Each Class

Class State Label Matrices
Band Priority (Band Priority Criterion)

Class State Label Matrices

(5). Class-wise Correlation for Each Band

Final Selected Band Subset

(6). Band Ranking via Average Correlation Coefficient Value
Testing

Sample under Testing \( \rightarrow \) Applying Selected Bands \( \rightarrow \) Selected Bands \( \rightarrow \) Classification \( \rightarrow \) Predicted Class Label

(7). Applying Selected Bands

(8). Classification
Image Description

- **Whole Indian Pine**
  - Whole version of 92AV3C (145 x 145 pixels, 16 classes)
  - 58 classes (39 classes are used)
  - 2166 x 614 pixels (39000 pixels are used)
  - Wavelength range 0.4–2.5 micrometers (totally 220 bands). Keep 200 bands by removing bands corresponding to water absorption: [104-108], [150-163], 220
Experiment Setup

- Five-fold cross validation
  - 20% for training
  - 80% for testing

- Overall classification rate
  - Average the overall classification rates from five-fold cross validation testing experiments

- Classifier
  - Support vector machine (SVM) (LibSVM Chang et al. 2007)
  - Radial basis function (RBF) kernel
State of the Art

- Minimum Estimated Abundance Covariance (MEAC) (Yang et al. 2011)
- Mutual Information (MI) (Guo et al. 2006)
- Feature Weighting (FW) (Huang et al. 2005)
Classification Results

<table>
<thead>
<tr>
<th>Method</th>
<th>Approximate</th>
<th>Lossless</th>
</tr>
</thead>
<tbody>
<tr>
<td>NHMC (2 Gaussian Components)</td>
<td>120 (99.21%)</td>
<td>140 (100.09%)</td>
</tr>
<tr>
<td>MEAC</td>
<td>170 (99.12%)</td>
<td>200 (100%)</td>
</tr>
<tr>
<td>MI</td>
<td>170 (99.56%)</td>
<td>200 (100%)</td>
</tr>
<tr>
<td>FW</td>
<td>150 (99.02%)</td>
<td>190 (100.02%)</td>
</tr>
<tr>
<td>Relief-F</td>
<td>180 (99.35%)</td>
<td>190 (101.85%)</td>
</tr>
</tbody>
</table>
Future Work

- We will focus on the fusion of band selection and spatial information in hyperspectral classification problems.
- The extension to unsupervised band selection will also be considered.


