Analysis of High-Dimensional Hyperspectral Imagery

Edward H. Bosch, John Greer, Robert Rand
NGA Basic and Applied Research
Remote Sensing Team
February 25, 2009
Compressive Sensing Workshop
Duke University, North Carolina
Hyperspectral Imagery (HSI)

• What is HSI?
• What can be done with this data?
  – Detection, Identification, Classification
• Issues arising from this type of data
  – Large dataset
  – Correlated narrow bands
  – Noise
  – How to extract relevant information?
Hyperspectral Image: for any point \((x,y)\) in the image its spectral signature \(S(x,y)\) is treated mathematically as the vector

\[
S(x,y) = [s_1 \ s_2 \ s_3 \ s_4 \ s_5 \ s_6 \ \ldots \ s_k]^T
\]

HyMap Scene A.P. Hill, VA
With permission from US Army TEC

Hyperspectral Image Cube:
400-2500 nm

Visible Region of the spectrum:
400-700 nm
• Spectral and spatial resolution of hyperspectral sensors provide image analysts with an enhanced capability to exploit the resulting imagery for a variety of remote sensing applications such as:
  • Monitoring crops, vegetation stress and chemical/oil spills, mining, change in mineral composition of soil, disaster relief, land/water management, detection of man-made materials, distinguish different objects with very similar spectral signatures.
Spectral signatures from different classes may be similar but there is enough subtle information to discriminate one class from the other.
• However,
  • Modeling high-dimensional databases may require large amounts of computer memory, processing time
  • Adjacent bands tend to be highly correlated, contain redundant information
  • Some bands are noisy
Correlation Matrix Of a Hyperspectral Cube
Noisy band
(sensor noise, atmospheric attenuation, …)

Noise-free band

HyMap Scene A.P. Hill, VA
With permission from US Army TEC
• What can be done to effectively handle so much data?
  • Dimension reduction
  • Extract features of the data
• There are many dimension reduction techniques. Some preserve physical information (e.g. reflectance/radiance values) others do not.
  • Spectral dimension
    • genetic algorithms (GA)
    • neural networks (NN)
    • wavelet transform
    • Principal Component Analysis (PCA)
    • matrix factorization techniques that include the Singular Value Decomposition
  • Spatial dimension
    • Geometry diffusion, local linear embedding, isomap
  • Spectral and spatial
Problem – Dimension Reduction

Solution - Solve for $\Phi: \mathbb{R}^n \rightarrow \mathbb{R}^m$ generating features $\Phi s(x,y) \in \mathbb{R}^m$, where $m<<n$. $\Phi$ maps the data to a lower dimensional space.

Application – Better separate objects to be discerned

HyMap Scene A.P. Hill, VA
With permission from US Army TEC
Hyperspectral image dimension reduction

- Reflectance $s(x,y) \in \mathbb{R}^n$
- Map $\Phi: \mathbb{R}^n \to \mathbb{R}^m$
- Features $\Phi s(x,y)$
HYDICE Scene near Fort Hood, TX
Color Composite of IR Bands

HYDICE Scene Copperas Cove, TX
With permission from US Army TEC
Derivative Difference Squared Classification with Original Hyperspectral Cube – 116 bands

HYDICE Scene Copperas Cove, TX
With permission from US Army TEC
Derivative Difference Squared Classification with Wavelet-Smoothed Cube – 61 bands

HYDICE Scene Copperas Cove, TX
With permission from US Army TEC
Normalized Difference Vegetation Index (NDVI) for LANDSAT multispectral imagery (7 bands) for each pixel in the image is computed by

\[
NDVI = \frac{z_7 - z_5}{z_7 + z_5} = \frac{z_7}{\sqrt{2}} - \frac{z_5}{\sqrt{2}}
\]

At any pixel we can generalize the above by applying the ratio throughout the components of the spectral signature \( z \) yielding what we call the Generalized Difference Feature Index

\[
GDFI_2(i, t) = \frac{h_0 z_i + h_1 z_{i+t}}{g_0 z_i + g_1 z_{i+t}}
\]

\[
h_0 = \frac{1}{\sqrt{2}} \quad \text{and} \quad h_1 = -\frac{1}{\sqrt{2}}
\]

\[
g_0 = g_1 = \frac{1}{\sqrt{2}}
\]

*Haar High and Low Pass Filters*
Wavelet-based feature indices

\[
NDVI = \frac{z_7 - z_5}{z_7/z_5} = \frac{z_7/\sqrt{2} - z_5/\sqrt{2}}{z_7/z_5}
\]

\[
GDFI_{2}(i,t) = \frac{h_0 z_i + h_1 z_{i+t}}{g_0 z_i + g_1 z_{i+t}}
\]

\[
GDFI_{2n}(i,t) = \frac{h_0 z_i + h_1 z_{i+t} + \ldots + h_{2n-1} z_{i+(2n)t}}{g_0 z_i + g_1 z_{i+t} + \ldots + g_{2n-1} z_{i+(2n)t}}
\]

\[
h_0 \ast 0^k + h_1 \ast 1^k + \ldots + h_{2n-1} \ast (2n-1)^k = 0, \quad 0 \leq k \leq n - 1
\]

\[
g_0 + g_1 + \ldots + g_{2n-1} = \sqrt{2}
\]
Wavelet-based Generalized Feature Indices
Trees, Shrubs & Bushes

HYDICE Scene Copperas Cove, TX
With permission from US Army TEC
Wavelet-based Generalized Feature Indices
Buildings

HYDICE Scene Copperas Cove, TX
With permission from US Army TEC