Hyperspectral Data Processing and Analysis using ENVI / Python

K. V. Kale
Professor,
Department of Computer Science and IT,
Dr. Babasaheb Ambedkar Marathwada University, Aurangabad.

Network Program of Imaging Spectroscopy and Applications
NISA – 2016
(Dec 20-24, 2016)

Design and Development of Hyperspectral Data Analysis Tool and Algorithm for End Member Identification

Principle Investigator
Professor K. V. Kale
kvkale91@gmail.com

Co-Principle Investigator
Dr. Ramesh R. Manza
(Associate Professor)
manzaramesh@gmail.com

Department of Computer Science and IT,
Dr. Babasaheb Ambedkar Marathwada University, Aurangabad (MS).
Contents

- Introduction
- Hyperspectral Data
- ENVI for Hyperspectral Data
- Python for Hyperspectral Data
- References

Introduction

- Background
  - Remote Sensing
  - Extensive Information due to synoptic view, map like format, and repetitive coverage area.
  - Provides data for weather prediction, agricultural forecasting, resource exploration, land cover mapping and environmental monitoring, to name a few.
  - Has significance advances in data acquisition, storage and management capabilities.
Introduction

➢ Background

• Continuous acquisition leads to huge database.

• Increased data volume needs automated analysis techniques.

• Techniques must be objective, reproducible and feasible to implement within available resources.

Introduction

➢ Spectral Imaging

• Multispectral Imaging

• No doubt! Multispectral are innovative but due to relatively few spectral bands, their spectral resolution is insufficient for many precise earth surface studies.

• Hyperspectral Imaging

• Hundreds of narrow contiguous wavelength intervals.

• Improvement in Spectral and Spatial Resolution.
Introduction

Fig. 1 Hyperspectral Image Cube

Challenges for Interpretation

- The Need for Calibration (Atmospheric Correction)
- Data Volume
- Data Redundancy
- Dimensionality Problem
Hyperspectral Data

**AVIRIS**

- AVIRIS: Airborne Visible/Infrared Imaging Spectrometer.
- Collects data for characterization of the Earth's surface and atmosphere from geometrically coherent spectroradiometric measurements.
- AVIRIS Application Areas: oceanography, environmental science, snow hydrology, geology, volcanology, soil and land management, atmospheric and aerosol studies, agriculture, and limnology.
- Application Under Development: Assessment and monitoring of toxic waste, oil spills, and land/air/water pollution.

**AVIRIS Sensor Specification**

- **Scanner type:** nadir-viewing, whiskbroom
- **Image width (swath):** 11 km (high altitude), 1.9 km (low altitude)
- **Typical image length:** 10 - 100 km
- **Spatial response:** 1.0 mrad, corresponding to a "pixel" 20m x 20m (high altitude) or 4m x 4m (low altitude) on the ground
- **Spectral response:** visible to near-infrared (400 to 2500 nm), with 224 contiguous channels, approximately 10 nm wide
- **Data quantization:** 12 bits
- Data capacity: 10 gigabytes, corresponding to about 850 km of ground track data, per flight
Hyperspectral Data

- **Hyperion**
  - Earth observation for improved Earth surface characterization.
  - Continuous acquisition into narrow spectral bands.
  - Complex land eco-systems can be imaged and accurately classified.
  - Hyperion Applications: geology, forestry, agriculture, and environmental management.
  - Detailed classification of land assets through the Hyperion will enable more accurate remote mineral exploration, better predictions of crop yield and assessments, and better containment mapping.

Professor K. V. Kale

Hyperspectral Data

- **Hyperion Sensor Specification**
  - **Spectral Channels:** 220 spectral bands.
  - **Spectral Coverage:** 0.4 to 2.5 μm
  - **Spatial Response:** Pixel size 30 m x 30 m resolution.
  - **Data Quantization:** 16 bit for level 1 product.
  - **Image Swath:** 7.5 km by 100 km land area per image

Professor K. V. Kale
ENVI for Hyperspectral Analysis

ENVI: Introduction

• ENVI: Environment for Visualizing Images

• Commercial software for spectral image analysis and visualization.

• Suitable for multispectral, Imaging and Non-Imaging Hyperspectral, LiDAR, and SAR datasets etc.

• Who use ENVI?:

  Scientists, researchers, image analysts, and GIS professionals and people from defense & intelligence, urban planning, mining, geology, space science, and earth science etc.
ENVI for Hyperspectral Analysis

ENVI: Functionality

• Directly access intuitive imagery analysis and processing tools

• Quickly visualize and analyze any source of remotely sensed data – including LiDAR, radar, optical, thermal and 3D etc.

• Completely automate geospatial imagery workflows using the ArcGIS platform and other GIS technology.

• Easily extract meaningful information from imagery – regardless of your technical ability with remote sensing technologies.

ENVI for Hyperspectral Analysis

ENVI: Functionality

• Read and Analyze different data formats

• Fuse multiple data modalities

• Exploit information from different sensor types

• Easily process large datasets
ENVI for Hyperspectral Analysis

ENVI: Supported Raster File Formats

- Band Interleaved by Line (BIL)
- Band Interleaved by Pixel (BIP)
- Band Sequential (BSQ)

ENVI for Hyperspectral Analysis

ENVI: Supported Raster File Formats

- ENVI supported data: Panchromatic, Multispectral, Imaging and Non-Imaging Hyperspectral, radar, thermal, HDF5, Full Motion Video, Net CDF-4, and LiDAR.

- ENVI supported sensors:

<table>
<thead>
<tr>
<th>ADS80</th>
<th>ASTER</th>
<th>AVIRIS</th>
<th>AVHRR</th>
<th>GOES-R</th>
<th>Himawari-8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Landsat 8</td>
<td>NPP VIIRS</td>
<td>Spectroradiometer Data</td>
<td>QuickBird</td>
<td>RADARSAT</td>
<td>SkySat-1</td>
</tr>
<tr>
<td>SPOT</td>
<td>TMS</td>
<td>USGS DEM</td>
<td>Geiger-Mode</td>
<td>LiDAR</td>
<td>radar</td>
</tr>
</tbody>
</table>
ENVI for Hyperspectral Analysis

- ENVI supported sensors

ENVI: File Management Tool

- This tool is used to handle Data from different sensors of Different format.
- Also Provides Export and Import Facility.
ENVI for Hyperspectral Analysis

ENVI: Basic Tool

- Resize Data
- Layer Stacking
- Format Conversion
- Mosaicking
- Masking
- Preprocessing

ENVI for Hyperspectral Analysis

ENVI: Spectral Tool

- Spectral Libraries
- MNF
- PPI
- N-D Visualizer
- Mapping Methods

Spectral Tool Continued
ENVI for Hyperspectral Analysis

ENVI: Spectral Tools

- Profiles and Plots
  - Extract pixels from an image that can be compared to spectral libraries or other pixels

- Spectral Library Viewer
  - Visualize data from ENVI standard spectral library (SLI) files, THOR Metadata Rich Spectral Library (MRSL) output, and Analytical Spectral Devices (ASD) spectrometer output.
  - Provides laboratory spectra from: NASA Jet Propulsion Laboratory (JPL), Johns Hopkins University (JHU), U. S. Geological Survey (USGS), ASTER.

ENVI for Hyperspectral Analysis

ENVI: Dimensionality Reduction

Minimum Noise Fraction

- To determine the inherent dimensionality of image data.
- To segregate noise in the data
ENVI for Hyperspectral Analysis

ENVI Spectral Tool: Dimensionality Reduction

• Minimum Noise Fraction (MNF)
  
  **MNF Rotation** transforms to determine the inherent dimensionality of image data, to segregate noise in the data, and to reduce the computational requirements for subsequent processing.

  MNF linear transformation consists separate principal components analysis rotations:

  ➢ *First Rotation*

  Used for noise whitening.

ENVI for Hyperspectral Analysis

ENVI Spectral Tool: Dimensionality Reduction

• Minimum Noise Fraction (MNF)

  ➢ *Second Rotation*

  Dimensionality of the data is determined by examining the final eigenvalues and the associated images.

  You can divide the data space into two parts: one part associated with large eigenvalues and coherent eigenimages, and a complementary part with near-unity eigenvalues and noise-dominated images.

  Using only the coherent portions separates the noise from the data, thus improving spectral processing results.
ENVI for Hyperspectral Analysis

ENVI: Spectral Tools

- **Pixel Purity Index (PPI)**
  - To find pure pixel.
  - PPI is computed by repeatedly projecting n-D scatter plots on a random unit vector.
  - PPI runs on MNF transformed result excluding the noise bands.
  - In Pixel Purity Image pixel value corresponds to the number of times that pixel was recorded as extreme.

---

ENVI for Hyperspectral Analysis

ENVI: Spectral Tools

- **N-D Visualizer**
  - **n-D Visualizer** to locate, identify, and cluster the purest pixels in a dataset in n-dimensional space.
  - The n-D Visualizer was designed to help you visualize the shape of a data cloud that results from plotting image data in spectral space (with image bands as plot axes).
ENVI for Hyperspectral Analysis

ENVI: Spectral Tools

- N-D Visualizer
  - You typically use the n-D Visualizer with spatially subsetted Minimum Noise Fraction (MNF) data that use only the purest pixels determined from the Pixel Purity Index (PPI).
  - n-D Visualizer allows interactively rotate data in n-D space, select groups of pixels into classes, and collapse classes to make additional class selections easier.
  - Makes it easy to export the selected classes to ROIs and use them as input into classification, Linear Spectral Unmixing, or Matched Filtering techniques.
Unsupervised Classification

ISODATA Classification

ENVI for Hyperspectral Analysis

ENVI: Unsupervised Classification

ENVI 4.4

File Basic Tools Classification Transform Filter Spectral Map Vector Topographic

Supervised

Unsupervised IsoData K-Means

Decision Tree

Endmember Collection

Create Class Image from ROIs

Post Classification

Professor K. V. Kale

ENVI for Hyperspectral Analysis

ENVI Spectral Tools: Spectral Angle Mapper (SAM)

- SAM is a physically-based spectral classification that uses an $n$-D angle to match pixels to reference spectra.

- The algorithm determines the spectral similarity between two spectra by calculating the angle between the spectra and treating them as vectors in a space with dimensionality equal to the number of bands.

- SAM compares the angle between the endmember spectrum vector and each pixel vector in $n$-D space.

Professor K. V. Kale
ENVI for Hyperspectral Analysis

ENVI Spectral Tools: Spectral Angle Mapper (SAM)

- SAM is a physically-based spectral classification that uses an $n$-D angle to match pixels to reference spectra.

- The algorithm determines the spectral similarity between two spectra by calculating the angle between the spectra and treating them as vectors in a space with dimensionality equal to the number of bands.

- SAM compares the angle between the endmember spectrum vector and each pixel vector in $n$-D space.

- Smaller angles represent closer matches to the reference spectrum. SAM classification assumes reflectance data.

ENVI for Hyperspectral Analysis

ENVI: Spectral Tools

- **Linear Spectral Unmixing**

  - Linear Spectral Unmixing to determine the relative abundance of materials that are depicted in multispectral or hyperspectral imagery based on the materials’ spectral characteristics.

  - The reflectance at each pixel of the image is assumed to be a linear combination of the reflectance of each material present within the pixel.

  - Spectral unmixing results are highly dependent on the input endmembers.

  - Changing the endmember changes the results.
Python for Hyperspectral Data Analysis

Python: Introduction

- Interpreted, object-oriented, high-level programming language with dynamic semantics.
- Uses an elegant syntax, making the programs you write easier to read
- Large standard library that supports many common programming tasks such as Image Processing, Machine Learning, Networking applications.
- Runs anywhere, including Mac OS X, Windows, Linux, and UNIX.
- It doesn't cost anything to download or use Python, or to include it in your application. Python can also be freely modified and re-distributed, because while the language is copyrighted it's available under an open source license.
Python for Hyperspectral Analysis

Python 2.x vs. 3.x

• “Should I use Python 2 or Python 3 for my development activity?”

• One sentence difference: “Python 2.x is legacy and Python 3.x is the present and future of the language.

• Python 2.x and Python 3.x share many similar capabilities but they should not be thought of as entirely interchangeable.

• Considerable differences in code syntax and handling.

• Python 2.x or Python 3 depends on third party libraries you rely on.

Python for Hyperspectral Analysis

Python Spatial Packages

<table>
<thead>
<tr>
<th>Data Handling</th>
<th>Data Analysis</th>
<th>Plotting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shapely</td>
<td>NumPy</td>
<td>Matplotlib</td>
</tr>
<tr>
<td>GDAL / OGR</td>
<td>SciPy</td>
<td>Prettyplotlib</td>
</tr>
<tr>
<td>pyQGIS</td>
<td>Shapely</td>
<td>Descartes</td>
</tr>
<tr>
<td>Pyshp</td>
<td>Pandas</td>
<td>Cartopy</td>
</tr>
<tr>
<td>pyproj</td>
<td>GeoPandas</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PySAL</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Rasterio</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SciKit-Image</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SciKit-Learn</td>
<td></td>
</tr>
</tbody>
</table>

Professor K. V. Kale
Python for Hyperspectral Analysis

Python: Dependencies

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Dependency</th>
<th>Version</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NumPy</td>
<td>1.11.2</td>
</tr>
<tr>
<td>2</td>
<td>GDAL</td>
<td>2.1.2</td>
</tr>
<tr>
<td>3</td>
<td>Matplotlib</td>
<td>1.5.3</td>
</tr>
<tr>
<td>4</td>
<td>Spectral</td>
<td>0.18.0</td>
</tr>
<tr>
<td>5</td>
<td>Tkinter</td>
<td>8.6</td>
</tr>
<tr>
<td>6</td>
<td>Xlsxwriter</td>
<td>0.9.4</td>
</tr>
<tr>
<td>7</td>
<td>PIL</td>
<td>1.1.7</td>
</tr>
<tr>
<td>8</td>
<td>SciKit-Learn</td>
<td>1.18.1</td>
</tr>
</tbody>
</table>

NumPy

- NumPy is the fundamental package for scientific computing in Python.
- It is a Python library that provides a multidimensional array object.
- Offers various derived objects such as masked arrays and matrices, an assortment of routines for fast operations on arrays, including mathematical, logical, shape manipulation, sorting, selecting, I/O, discrete Fourier transforms, basic linear algebra, basic statistical operations, random simulation and much more.
Python for Hyperspectral Analysis

Python: Dependencies

❖ NumPy

• At the core of the NumPy package, is the ND array object. This encapsulates n-dimensional arrays of homogeneous data types. The more important attributes of an ndarray object are:

  • Dimensions
  • Shape
  • Size
  • Data Type

❖ GDAL

• Translator library for raster and vector geospatial data formats.

• Released under an X/MIT style Open Source license by the Open Source Geospatial Foundation.

• GDAL/OGR is considered a major free software project for its "extensive capabilities of data exchange".

• GDAL: Supports the reading/writing/translation of numerous raster formats

• OGR: Supports reading/writing/translation of numerous vector data formats.
Python for Hyperspectral Analysis

Python: Dependencies

- GDAL Data Model
  - Dataset
    A dataset (represented by the `GDALDataset` class) is an assembly of related raster bands and some information common to them all.

    Dataset has a concept of the raster size (in pixels and lines) that applies to all the bands.

    The dataset is also responsible for the georeferencing transform and coordinate system definition of all bands.

    The dataset itself can also have associated metadata, a list of name/value pairs in string form.

- Coordinate System
  Dataset coordinate systems are represented as OpenGIS Well Known Text strings. This can contain:
  - An overall coordinate system name.
  - A geographic coordinate system name.
  - A datum identifier.
  - An ellipsoid name, semi-major axis, and inverse flattening.
  - A prime meridian name and offset from Greenwich.
Python for Hyperspectral Analysis

Python: Dependencies

- GDAL Data Model
  - Coordinate System:
    - A projection method type (i.e. Transverse Mercator).
    - A list of projection parameters (i.e. central_meridian).
    - A units name, and conversion factor to meters or radians.
    - Names and ordering for the axes.
    - Codes for most of the above in terms of predefined coordinate systems from authorities such as EPSG.

- SpectralPython
  - Pure Python module for processing hyperspectral image data.
  - Free, open source software distributed under the GNU General Public License.
  - It depends on several other freely available python modules. (i.e. NumPy)
  - Spectral Python Supported Raster File Formats:
    1. ENVI Headers
    2. AVIRIS
    3. ERDAS
Python for Hyperspectral Analysis

Python: Dependencies

- Matplotlib
  - Python 2D plotting library.
  - Open Source
  - Matplotlib tries to make easy things easy and hard things possible.
  - Used to generate plots, histograms, power spectra, bar charts, error charts, scatterplots etc.
  - Provides Matlab like plotting interface.

Python for Hyperspectral Analysis

Python: Dependencies

- Matplotlib
  - Conceptually divided into three parts:
    1. Pylab interface
       Set of functions to create plots quite similar to MATLAB figure generating code.
    2. Matplotlib Frontend
       Set of classes that do the heavy lifting, creating and managing figures, text, lines, plots and so on.
       This is an abstract interface that knows nothing about output.
    3. Matplotlib Backend
       The backends are device-dependent drawing devices (I.e GTK, Wx, Tkinter)
Python: Dependencies

- **Tkinter**
  - Standard GUI library for Python.
  - Provides controls / widgets for GUI application:

<table>
<thead>
<tr>
<th>Button</th>
<th>Entry</th>
<th>Listbox</th>
<th>Message</th>
<th>Text</th>
<th>Paned Window</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canvas</td>
<td>Frame</td>
<td>Menu</td>
<td>Radiobutton</td>
<td>Toplevel</td>
<td>Label Frame</td>
</tr>
<tr>
<td>Checkbutton</td>
<td>Label</td>
<td>Menubutton</td>
<td>Scrollbar</td>
<td>Spinbox</td>
<td>tkMessageBox</td>
</tr>
</tbody>
</table>

- **Xlsxwriter**
  - Xlsxwriter is a Python module that can be used to write text, numbers, formulas and hyperlinks to multiple worksheets in an Excel 2007+ XLSX file. It supports features such as:
    - 100% compatible Excel XLSX files.
    - Full formatting
    - Charts.
    - Auto filters.
Python for Hyperspectral Analysis

Python: Dependencies

- **Xlsxwriter**
  - It supports features such as:
    - Data validation and drop down lists.
    - Conditional formatting.
    - Worksheet PNG/JPEG images.
    - Rich multi-format strings.
    - Textboxes.
    - Memory optimization mode for writing large files.

Python for Hyperspectral Analysis

**Python Dependencies : Xlsxwriter**

- **Advantages**
  - It supports more Excel features than any of the alternative modules.
  - It has a high degree of fidelity with files produced by Excel.
  - In most cases the files produced are 100% equivalent to files produced by Excel.
  - It has extensive documentation, example files and tests.
  - It is fast and can be configured to use very little memory even for very large output files.

- **Disadvantage**
  - It cannot read or modify existing Excel XLSX files.
Python for Hyperspectral Analysis

### Python Dependencies: Python Imaging Library (PIL)

- Adds image processing capabilities to your Python interpreter.
- The library contains basic image processing functionality, including point operations, filtering with a set of built-in convolution kernels, and color space conversions.
- The library also supports image resizing, rotation and arbitrary affine transforms.
- Method to pull some statistics out of an image. This can be used for automatic contrast enhancement, and for global statistical analysis.

### Python Dependencies: SciKit-Learn

- Scikit-learn is a Python module integrating classic machine learning algorithms.
- Simple and efficient tools for data mining and data analysis.
- Accessible to everybody, and reusable in various contexts.
- Built on NumPy, SciPy, and Matplotlib.
- Open source, commercially usable - BSD license
- Problems it tackles range from building a prediction function linking different observations, to classifying observations, or learning the structure in an unlabeled dataset.
KVK - Hyperspectral Image Analysis Tool

- Tool Functionalities
  - Image Display
  - Preprocessing
  - End Member Extraction
  - Classification
KVK-Hyperspectral Data Analysis Tool

- Tool Functionalities
  - Image Display
    - Monochrome
    - RGB
    - Multispectral
    - Hyperspectral

KVK-Hyperspectral Data Analysis Tool

- Tool Functionalities
  - Number of Bands
    - Display Single Band
    - Display RGB
KVK-Hyperspectral Data Analysis Tool

- Tool Functionalities
  - Gray Channel Display
    - Band 10
  - RGB Display
    - Band 190 + 210 + 100
KVK-Hyperspectral Data Analysis Tool

- Tool Functionalities

  RGB View : 50 + 60 + 210

  RGB View : 50 + 60 + 210

KVK-Hyperspectral Data Analysis Tool

- Tool Functionalities

  ✓ Preprocessing
KVK-Hyperspectral Data Analysis Tool

Tool Functionalities

- Dimensionality Reduction
  - PCA

KVK-Hyperspectral Data Analysis Tool

Tool Functionalities

- PCA
  - Original Image
  - Principal Components
KVK-Hyperspectral Data Analysis Tool

**Tool Functionalities**

- **End Member Extraction**
  - PPI

---

KVK-Hyperspectral Data Analysis Tool

**Tool Functionalities**

- **PPI**
  - Input Image
  - Iterations
  - Threshold
  - Centered
  - Start Position
  - Display
KVK-Hyperspectral Data Analysis Tool

✓ PPI Result

![Hyperspectral Image](image1)

![PPI (200 iterations)](image2)

---

KVK-Hyperspectral Data Analysis Tool

- Tool Functionalities
  - Classification
    - Supervised
    - Unsupervised
KVK-Hyperspectral Data Analysis Tool

- Tool Functionalities
  - Unsupervised
    - KMeans

KVK-Hyperspectral Data Analysis Tool

KMeans Clustering Result
References


Professor K. V. Kale

Thank You…

Professor K. V. Kale