Hyperspectral Image Processing and Analysis for Land Cover Characterization

Dr. Jay Pearlman and Dr. Melba Crawford
September 19, 2004
Course Overview

• Introduction to EO Remote Sensing
  – Photon end to end flow
  – Alternative sensor configurations
• Hyperion Instrument – design and trades
• Sensor Calibration
• Data Processing and Corrections
• Classification of Hyperspectral Data
  – Input space reduction
    • Feature selection and extraction
  – Output space reduction
    • Multiclassifier Systems
• A View Forward
EO Remote Sensing Process Overview

Solar Irradiance

Atm. Transmittance

Path Radiance

Surf. Reflectance

Data downlink

Measured signal

Ground Station

Calibration Parameters

Atmospheric Correction

Data Exploitation

Copyright of R. Green
Solar Irradiance
Reflection from a Metal Roof

San Francisco: January 17, 2001: Roof Top
HypGain_revA

Reflection of Solar Irradiance of a roof top, a reflective surface. Some atmospheric features are indicated.
Radiance To Reflectance Inversion

\[ L_t = F_0 \rho_a / \pi + \mu F_0 T_d \rho_s / \pi \]

- \( F_0 \) is the exo atmospheric irradiance
- \( \rho_a \) is the upward reflectance of the atmosphere
- \( T_d \) is the downward transmittance
- \( \rho_s \) is the reflectance of the surface
Alternative Sensor Configurations

- Multispectral vs. Hyperspectral
- Whiskbroom vs. Pushbroom
Hyperspectral and Multispectral Perspectives

Hyperspectral Imaging
Hundreds of bands

Spectral characteristic of scene

Multispectral Imaging
Few bands
Landsat Picks the Windows – But Misses a Lot

Crop and Soil Spectra
Landsat Bands (1)

Reflectance

Wavelength (nm)

L1 L2 L3 L4 Pan L5 L7

Legend:
- Rice
- Soy
- Soil
- Stubble
Hyperion surface composition map agrees with known geology of Mt. Fitton in South Australia

(1) Published Geologic Survey Map
(2) Hyperion three color image (visible) showing regions of interest
(3) Hyperion surface composition map using SWIR spectra above

Hyperion Maps Mt. Fitton Geology

Hyperion-based apparent reflectance compares with library reference spectra

(a) Hyperion Spectra
(b) Reference Spectra

Courtesy of CSIRO, Australia
Mt. Etna Sample Spectra

Hyperion Spectra of Mt Etna Scene July 13th 2001

- lava
- smoke
- vegetation

Wavelength (nm)

Radiance (W/m²-um-sr)

1234 nm
1639 nm
2226 nm
Hyperspectral Sensor Configurations

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Whiskbroom</th>
<th>Pushbroom</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detectors</td>
<td>Linear array</td>
<td>2-dimensional array</td>
</tr>
<tr>
<td>Relative pixel dwell time</td>
<td>Shorter</td>
<td>Longer</td>
</tr>
<tr>
<td>Spatial uniformity</td>
<td>Easier</td>
<td>Harder</td>
</tr>
<tr>
<td>Calibration</td>
<td>Easier</td>
<td>Harder</td>
</tr>
<tr>
<td>Application domain</td>
<td>Airborne</td>
<td>Space and Airborne</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>Similar</td>
<td>Similar</td>
</tr>
</tbody>
</table>
Hyperion Configuration

• Pushbroom configuration with common fore optics
  – 256 field-of-view locations, defines 7.7 km swath width, 30 meters each
  – 6925 frames of data, defines 185 km swath length, 30 meters each. (sample based on length of data), frame rate timed with spacecraft velocity

Hyperion Data Cube
X-swath width
Y-swath length
Z-spectra signature
The Hyperion Sensor
Hyperion is a push-broom imager

- 220 10nm bands covering 400nm - 2500nm
- 6% absolute rad. accuracy
- Swath width of 7.5 km
- IFOV of 42.4 μradian
- GSD of 30 m
- 12-bit image data
- Orbit is 705km alt (16 day repeat)
HYPERION TECHNOLOGIES

CALIBRATION
(spectral/pushbroom)

DATA RATES

SPECTROMETER
(curved grating)

Reflecting TELESCOPE

Pulse Tube CRYOCOOLER

DATA RATES
Hyperion Sensor Assembly

Baffle / CALIBRATION

Reflecting TELESCOPE

Signal processor

Spectrometer

Pulse Tube CRYOCOOLER
Hyperion Optical System

Three mirror anastigmat (TMA) foreoptic

Earth’s Surface

Spectrometer FPA

Entrance slit

Grating

Focal Plane Array
Hyperion Subassemblies

Hyperion Electronics Assembly (HEA)

Cryocooler Electronics Assembly (CEA)

Hyperion Sensor Assembly (HSA)
Hyperion Sensor

Hyperion Instrument Electronics

Aperture Cover

Cry-cooler Electronics (CEA)

S/C 28 VDC_1

1773 Command & Telemetry

RS-422

Hyperion Sensor

Hyperion Instrument Electronics

Thermally Isolated from S/C

Thermally Connected to S/C
### Hyperion Key Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Pre-launch</th>
<th>On-orbit</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSD (m)</td>
<td>29.88</td>
<td>30.38</td>
</tr>
<tr>
<td>Swath (km)</td>
<td>7.5</td>
<td>7.75</td>
</tr>
<tr>
<td>VNIR MTF @ 630nm</td>
<td>0.22-0.28</td>
<td>0.23-0.27</td>
</tr>
<tr>
<td>SWIR MTF @ 1650nm</td>
<td>0.25-0.27</td>
<td>0.28</td>
</tr>
<tr>
<td>VNIR X-trk Spec. Error</td>
<td>2.8nm@655nm</td>
<td>2.2nm</td>
</tr>
<tr>
<td>SWIR X-trk Spec. Error</td>
<td>0.6nm@1700nm</td>
<td>0.58</td>
</tr>
<tr>
<td>Abs. Radiometry (1Sigma)</td>
<td>&lt;6%</td>
<td>3.40%</td>
</tr>
<tr>
<td>VNIR SNR (550-700nm)</td>
<td>144-161</td>
<td>140-190</td>
</tr>
<tr>
<td>SWIR SNR (~1225nm)</td>
<td>110</td>
<td>96</td>
</tr>
<tr>
<td>SWIR SNR (~2125nm)</td>
<td>40</td>
<td>38</td>
</tr>
<tr>
<td>No. of Spectral Channels</td>
<td>220</td>
<td>200 (L1)</td>
</tr>
<tr>
<td>VNIR Bandwidth (nm)</td>
<td>10.19-10.21</td>
<td>**</td>
</tr>
<tr>
<td>SWIR Bandwidth (nm)</td>
<td>10.08-10.09</td>
<td>**</td>
</tr>
</tbody>
</table>

** not measured
Hyperion SNR

Radiometric performance model based on 60° Solar zenith angle, 30% albedo, standard scene

<table>
<thead>
<tr>
<th>Wavelength (nm)</th>
<th>Hyperion Measured SNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>550 nm</td>
<td>161</td>
</tr>
<tr>
<td>650 nm</td>
<td>144</td>
</tr>
<tr>
<td>700 nm</td>
<td>147</td>
</tr>
<tr>
<td>1025 nm</td>
<td>90</td>
</tr>
<tr>
<td>1225 nm</td>
<td>110</td>
</tr>
<tr>
<td>1575 nm</td>
<td>89</td>
</tr>
<tr>
<td>2125 nm</td>
<td>40</td>
</tr>
</tbody>
</table>
Spectrometer Overlays

Level 1 data: 438-926nm and 892-2406nm
Bands 9-57 and 75 - 225;
SWIR is West of VNIR and rotated CCW by one pixel

<table>
<thead>
<tr>
<th>Band</th>
<th>Center(nm)</th>
<th>FWHM(nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>854.66</td>
<td>11.27</td>
</tr>
<tr>
<td>51</td>
<td>864.83</td>
<td>11.27</td>
</tr>
<tr>
<td>52</td>
<td>875</td>
<td>11.28</td>
</tr>
<tr>
<td>53</td>
<td>885.17</td>
<td>11.29</td>
</tr>
<tr>
<td>54</td>
<td>895.34</td>
<td>11.3</td>
</tr>
<tr>
<td>55</td>
<td>905.51</td>
<td>11.31</td>
</tr>
<tr>
<td>56</td>
<td>915.68</td>
<td>11.31</td>
</tr>
<tr>
<td>57</td>
<td>925.85</td>
<td>11.31</td>
</tr>
<tr>
<td>58</td>
<td>936.02</td>
<td>11.31</td>
</tr>
<tr>
<td>59</td>
<td>946.19</td>
<td>11.31</td>
</tr>
<tr>
<td>71</td>
<td>852</td>
<td>11.17</td>
</tr>
<tr>
<td>72</td>
<td>862.09</td>
<td>11.17</td>
</tr>
<tr>
<td>73</td>
<td>872.18</td>
<td>11.17</td>
</tr>
<tr>
<td>74</td>
<td>882.27</td>
<td>11.17</td>
</tr>
<tr>
<td>75</td>
<td>892.35</td>
<td>11.17</td>
</tr>
<tr>
<td>76</td>
<td>902.44</td>
<td>11.17</td>
</tr>
<tr>
<td>77</td>
<td>912.53</td>
<td>11.17</td>
</tr>
<tr>
<td>78</td>
<td>922.62</td>
<td>11.17</td>
</tr>
<tr>
<td>79</td>
<td>932.72</td>
<td>11.17</td>
</tr>
<tr>
<td>80</td>
<td>942.81</td>
<td>11.17</td>
</tr>
</tbody>
</table>
Polarization - VNIR

AVERAGED OVER FOV
In Units of Percent Polarization

Polarizer Angle [degrees]

Wavelength [nm]
Polarization - SWIR

AVERAGED OVER FOV
In Units of Percent Polarization

Spectral Channel

Polarizer angle [degrees]
• Ratio images for each intensity level were interpolated to create maps of echo variation across the entire FPA.

• Based on this map a technique for echo removal was implemented.

• The removal algorithm consists of subtracting a shifted and scaled version of each frame from itself. Scaling is a multiplication of the image by the echo variation map.
Profiles Before(red)/After(blue) Echo Removal
Hyperion Images Before / After Echo Removal
Anomalous Pixel Collage –
Possibilities including striping

<table>
<thead>
<tr>
<th>Band</th>
<th>Swath Pix</th>
<th>Band</th>
<th>Swath Pix</th>
<th>Band</th>
<th>Swath Pix</th>
<th>Band</th>
<th>Swath Pix</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>114</td>
<td>12</td>
<td>6</td>
<td>32</td>
<td>114</td>
<td>99</td>
<td>92</td>
</tr>
<tr>
<td>5</td>
<td>141</td>
<td>12</td>
<td>114</td>
<td>33</td>
<td>114</td>
<td>116</td>
<td>127</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>13</td>
<td>114</td>
<td>34</td>
<td>114</td>
<td>116</td>
<td>127</td>
</tr>
<tr>
<td>6</td>
<td>68</td>
<td>14</td>
<td>114</td>
<td>39</td>
<td>177</td>
<td>116</td>
<td>138</td>
</tr>
<tr>
<td>6</td>
<td>172</td>
<td>14</td>
<td>247</td>
<td>40</td>
<td>13</td>
<td>116</td>
<td>138</td>
</tr>
<tr>
<td>7</td>
<td>7</td>
<td>15</td>
<td>114</td>
<td>48</td>
<td>20</td>
<td>119</td>
<td>240</td>
</tr>
<tr>
<td>7</td>
<td>68</td>
<td>16</td>
<td>114</td>
<td>49</td>
<td>20</td>
<td>119</td>
<td>240</td>
</tr>
<tr>
<td>7</td>
<td>179</td>
<td>17</td>
<td>114</td>
<td>50</td>
<td>20</td>
<td>120</td>
<td>240</td>
</tr>
<tr>
<td>7</td>
<td>185</td>
<td>18</td>
<td>114</td>
<td>51</td>
<td>20</td>
<td>120</td>
<td>240</td>
</tr>
<tr>
<td>8</td>
<td>7</td>
<td>19</td>
<td>114</td>
<td>52</td>
<td>13</td>
<td>121</td>
<td>196</td>
</tr>
<tr>
<td>8</td>
<td>12</td>
<td>20</td>
<td>114</td>
<td>52</td>
<td>33</td>
<td>125</td>
<td>54</td>
</tr>
<tr>
<td>8</td>
<td>68</td>
<td>21</td>
<td>114</td>
<td>53</td>
<td>33</td>
<td>125</td>
<td>115</td>
</tr>
<tr>
<td>8</td>
<td>114</td>
<td>22</td>
<td>114</td>
<td>57</td>
<td>13</td>
<td>125</td>
<td>160</td>
</tr>
<tr>
<td>8</td>
<td>121</td>
<td>23</td>
<td>114</td>
<td>61</td>
<td>93</td>
<td>125</td>
<td>164</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>24</td>
<td>114</td>
<td>72</td>
<td>95</td>
<td>127</td>
<td>40</td>
</tr>
<tr>
<td>9</td>
<td>68</td>
<td>25</td>
<td>114</td>
<td>72</td>
<td>95</td>
<td>127</td>
<td>66</td>
</tr>
<tr>
<td>9</td>
<td>114</td>
<td>26</td>
<td>114</td>
<td>75</td>
<td>2</td>
<td>127</td>
<td>213</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>27</td>
<td>47</td>
<td>94</td>
<td>82</td>
<td>128</td>
<td>30</td>
</tr>
<tr>
<td>10</td>
<td>114</td>
<td>27</td>
<td>114</td>
<td>94</td>
<td>82</td>
<td>128</td>
<td>96</td>
</tr>
<tr>
<td>10</td>
<td>131</td>
<td>28</td>
<td>47</td>
<td>94</td>
<td>93</td>
<td>128</td>
<td>126</td>
</tr>
<tr>
<td>10</td>
<td>199</td>
<td>28</td>
<td>114</td>
<td>94</td>
<td>93</td>
<td>165</td>
<td>148</td>
</tr>
<tr>
<td>11</td>
<td>6</td>
<td>29</td>
<td>114</td>
<td>99</td>
<td>81</td>
<td>165</td>
<td>248</td>
</tr>
<tr>
<td>11</td>
<td>114</td>
<td>30</td>
<td>114</td>
<td>99</td>
<td>81</td>
<td>168</td>
<td>245</td>
</tr>
<tr>
<td>11</td>
<td>199</td>
<td>31</td>
<td>114</td>
<td>99</td>
<td>92</td>
<td>168</td>
<td>245</td>
</tr>
</tbody>
</table>

Bad Pix 3 - L1B1;  CSIRO – Jupp;  Dark Image – Pearlman;  PFC - Han
# Hyperion Nominal Data Modes

<table>
<thead>
<tr>
<th>Mode</th>
<th>Cover Position</th>
<th>Data collect</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standby</td>
<td>Closed</td>
<td>None</td>
<td>Default mode for active state</td>
</tr>
<tr>
<td>Dark calibration</td>
<td>Closed</td>
<td>Minimum 100 frames</td>
<td>Performed as close as possible to imaging, before and after</td>
</tr>
<tr>
<td>Lamp calibration</td>
<td>Closed</td>
<td>Minimum 100 frames</td>
<td>Performed after second dark calibration; two radiance levels</td>
</tr>
<tr>
<td>Solar calibration</td>
<td>Open 37 degrees</td>
<td>Minimum 1 second</td>
<td>Performed over North Pole only to keep cover out of ALI keep-out zone; yaw maneuvers required</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nominal 1 cube</td>
<td></td>
</tr>
<tr>
<td>Lunar calibration</td>
<td>Fully open (135°)</td>
<td>Minimum: 1 second</td>
<td>Performed on dark side of earth; off-track spacecraft pointing required</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nominal: 1 cube</td>
<td></td>
</tr>
<tr>
<td>Ground calibration</td>
<td>Fully open (135°)</td>
<td>Minimum 1 second</td>
<td>Ground target selected</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nominal 1 cube</td>
<td></td>
</tr>
<tr>
<td>Imaging</td>
<td>Fully open (135°)</td>
<td>Minimum 1 second</td>
<td>Nominal data collect is equivalent to Landsat scene, and takes 27 seconds.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Nominal 9 cubes</td>
<td></td>
</tr>
</tbody>
</table>
Performance Tradeoffs

- Spatial Resolution
- Spectral resolution
Coleambally Image Collection – 30m to 1km

Landsat

MODIS

SAC-C

250m

MODIS

1000m
## Comparison of Hyperspectral Instruments

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Lewis HSI</th>
<th>Hyperion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume (L x W x H, cm)</td>
<td>43x69x94</td>
<td>39x75x66</td>
</tr>
<tr>
<td>Weight (Kg)</td>
<td>39.5</td>
<td>49</td>
</tr>
<tr>
<td>Avg Power (W)</td>
<td>66</td>
<td>51</td>
</tr>
<tr>
<td>Peak Power (W)</td>
<td></td>
<td>126</td>
</tr>
<tr>
<td>Aperture (cm)</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>IFOV (mrad)</td>
<td><strong>0.057</strong></td>
<td><strong>0.043</strong></td>
</tr>
<tr>
<td>Crosstrack FOV (deg)</td>
<td>0.84</td>
<td>0.63</td>
</tr>
<tr>
<td>Wavelength Range (nm)</td>
<td>380 - 2450</td>
<td>400 - 2450</td>
</tr>
<tr>
<td>Spectral Resolution (nm)</td>
<td>5.1/6.45</td>
<td>10</td>
</tr>
<tr>
<td>No. Spectral Bands</td>
<td>384</td>
<td>220</td>
</tr>
<tr>
<td>Digitization</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>Frame Rate (Hz)</td>
<td>237</td>
<td>225</td>
</tr>
<tr>
<td>Typical SNR</td>
<td>100 - 200</td>
<td>65 - 130</td>
</tr>
<tr>
<td>Spectral Calibration (nm)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Radiometric Calibration</td>
<td>&lt;6%</td>
<td>&lt;6%</td>
</tr>
</tbody>
</table>
Hyperion Calibration
Ground Test and On-Orbit Validation

- System performance assessment strategy

- Present pre-flight and on-orbit measurement techniques
  - Absolute Radiometric Calibration
  - Spectral Calibration
  - Image Quality Characterization
Strategy

• Pre-Flight:
  – Establish fundamental characteristics of the instrument and assess requirement compliance
  – Establish instrument performance through the build, environmental test and spacecraft integration phases
  – Provide solid foundation for on-orbit comparison

• On-orbit:
  – Determine on-orbit performance and compare with pre-flight performance
  – Define data collects that can be used to assess identified performance parameters
  – Acquire and analyze data collections; and assess accuracy of technique
  – Compare on-orbit results with pre-flight
Special Targets for Characterization

- **Searchlights**
  - California

- **Planets**
  - Venus

- **Gas Flares**
  - Moomba

- **90 deg Yaw**
Desert Sites used for Vicarious Calibration

Lake Frome

RR Valley

Arizaro/Barreal Blanco
the good
the bad
and
the gotchas

The world of real life operations
Images of Lake Frome
The Salt Surface

Site 8
Site 10
Site 12
Site 18
Site 24
Site 25
Impact of Moisture on Salt Signature

Approximate lab analysis indicates changes of spectral signature based on amount of water added.
The Locusts That Did Not Make It Over

Calibration signature
Or
New culinary delicacy?
Arizaro Dry Lake bed
Ground Calibration
# Coleambally Collection History

<table>
<thead>
<tr>
<th>Date (GMT)</th>
<th>Julian Day (GMT)</th>
<th>Path/Row</th>
<th>X-track Position</th>
<th>In-track Position</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>23-Dec-00</td>
<td>358</td>
<td>93/84</td>
<td>100</td>
<td>3148</td>
<td>Cum (South); VNIR</td>
</tr>
<tr>
<td>1-Jan-01</td>
<td>001</td>
<td>92/84</td>
<td>82</td>
<td>2131</td>
<td>Clear</td>
</tr>
<tr>
<td>9-Jan-01</td>
<td>009</td>
<td>93/84</td>
<td>-</td>
<td>-</td>
<td>UARDRY</td>
</tr>
<tr>
<td>17-Jan-01</td>
<td>017</td>
<td>92/84</td>
<td>-</td>
<td>-</td>
<td>Cloudy</td>
</tr>
<tr>
<td>25-Jan-01</td>
<td>025</td>
<td>93/84</td>
<td>95**</td>
<td>2481**</td>
<td>High Clouds (North)</td>
</tr>
<tr>
<td>2-Feb-01</td>
<td>034</td>
<td>92/84</td>
<td>36</td>
<td>2599</td>
<td>Clear</td>
</tr>
<tr>
<td>9-Feb-01</td>
<td>041</td>
<td>93/84</td>
<td>-</td>
<td>-</td>
<td>Cloudy</td>
</tr>
<tr>
<td>18-Feb-01</td>
<td>049</td>
<td>92/84</td>
<td>41</td>
<td>3782</td>
<td>Shadows (East)</td>
</tr>
<tr>
<td>25-Feb-01</td>
<td>056</td>
<td>93/84</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>6-Mar-01</td>
<td>065</td>
<td>92/84</td>
<td>95</td>
<td>3977</td>
<td>Clear</td>
</tr>
<tr>
<td>13-Mar-01</td>
<td>072</td>
<td>93/84</td>
<td>104</td>
<td>2926</td>
<td>Clear</td>
</tr>
<tr>
<td>22-Mar-01</td>
<td>082</td>
<td>92/84</td>
<td>-</td>
<td>-</td>
<td>Cloudy</td>
</tr>
</tbody>
</table>
ARIZARO, Argentina

High altitude (~12,000ft) dry salt lakebed.

Surface: extremely rough, locally uniform, and bright.

Some years no rain at all!

It rained the day we arrived.

The good news is that conditions were good for the experiment on the 7th of February 2001.
Calibration & Validation Can Be Fun
Let’s Address Calibration
System Performance Verification Strategy

Performance Verification (end-to-end)

- Geometric
  - GSD
  - VNIR/SWIR Spatial Coregistration
- Spectral
  - MTF
  - Center Wavelength
- Absolute Radiometry
  - Applying Calibration
  - Defining Calibration
    - Solar Cal.
    - Dark Removal
      - Artifact Removal
  - Repeatability
  - Vicarious Calibration
Absolute Radiometric Calibration
## Key Factors Impacting Calibration

<table>
<thead>
<tr>
<th></th>
<th>Absolute Knowledge</th>
<th>Intermediate Properties</th>
<th>Spacecraft Pointing</th>
<th>Strengths</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Solar Calibration</strong></td>
<td>Models avail to community VNIR more accurate then SWIR</td>
<td>Diffuse reflectance of Hyperion cover</td>
<td>Critical to modeling intermediate properties</td>
<td>Uniform across field-of-view Constant</td>
</tr>
<tr>
<td><strong>Lake Frome (vicarious)</strong></td>
<td>Based on ground truth measurements</td>
<td>Atmospheric effects must be modeled</td>
<td>Depends on surface</td>
<td>User oriented effort</td>
</tr>
<tr>
<td><strong>Lunar Calibration</strong></td>
<td>Based on Lunar models</td>
<td>none</td>
<td>Spacecraft scans moon. Relative moon, sun, sat angle</td>
<td>No intermediate properties. Constant</td>
</tr>
</tbody>
</table>
On-Orbit Radiometric Calibration

• Solar Calibration
  – Absolute Comparison: VNIR within 2%, SWIR 5-8% low; SWIR has larger uncertainty due to solar model and BRDF model of cover surface
  – Used to correct for pixel-to-pixel corrections
  – Included in repeatability assessment 0.6% for VNIR, 1.6% for SWIR
  – Used to define noise level as a function of signal level to determine SNR

• Lunar Calibration
  – Used to eliminate artifacts of diffuser

• Vicarious Calibration and Cross Calibration
  – Calibration accuracy in 4-10% range
Solar Calibration Trending

History of VNIR Response to Solar Calibration

History of SWIR Response to Solar Calibration
Lunar Calibration Trending

Intensity Dependent on Moon-Sun-Spacecraft Relative Angles

VNIR and SWIR track intensity changes
The target spot is viewed by pixel 256 during frame 1, and then by pixel 255 during frame 2, and so on…
Analysis for Yaw Data

Level 1R  
Transposed 45 Deg  
Flatfielded (using 2001197 coefficients)  

Hyperion Band 20 (VNIR)

Differences between the sets of correction coefficients for the entire vnir focal plane are within ±2.5%
Analysis for Yaw Data

Transposed 45 Deg  Flatfielded  Flatfielded
(using 2001197 coefficients)

Hyperion Band 91 (SWIR)
Spectral Calibration
Pre-Flight Spectral Calibration

Monochromometer stepped spectrally in 2nm increments

Monochromometer profiles used to define center wavelength and bandwidth at discrete locations

Spectral response modeled as a gaussian with a center wavelength and full width half max
Pre-Flight Spectral Calibration

- **Center wavelength and bandwidth**
  - Measured at discrete locations 20 VNIR locations, 25 SWIR locations.
  - Used to define the center wavelength and bandwidth for every VNIR and SWIR pixel, 256 field-of-view locations and 242 spectral bands.

- **Dispersion (nm/pixel)**
  - Spacing of spectral channels, Hyperion dispersion (~10nm/pixel) closely matches the bandwidth (10 nm)

- **Cross-track spectral difference**
  - Maximum wavelength difference across field-of-view for a single spectral channel,
  - VNIR = 2.6-3.6 nm
  - SWIR = 0.40-0.97 nm.
Pre-Flight Calibration Algorithm Development

Process:
1. Image contains reference spectra uniform across the field of view. (pre-flight: doped Spectralon)
2. High resolution reference spectra convolved with sensor spectral response function
3. Resulting reference spectra aligned with Hyperion measured spectra to determine spectral calibration.
On-Orbit Spectral Calibration

- Early solar cal, sun is rising through earth’s limb during collect.
- View sun, off cover, through atmosphere
- Extension of Pre-flight algorithm development
SWIR

Identification of Spectral Features

Process:
1.) Create Pseudo-Hyperion Spectra from reference: Modtran-3 for atmosphere, and Cary 5 & FTS measurements for diffuse reflectance of the cover
2.) Correlate Spectral Features: band number units of Hyperion max/min correlated with reference wavelength of max/min
3.) Calculate Band to Wavelength map: apply low order polynomial to fit the data over the entire SWIR regime

VNIR

- Spectral calibration based on two lines:
  - one solar line (520 nm) and an oxygen line (762.5 nm)
Spectral-Spatial-Uniformity

**Depiction**
- Grids are the detectors
- Spots are the IFOV centers
- Colors are the wavelengths

```
Cross Track Sample
```

```
Requirement
```

```
Failure by Frown(aka smile)
```

```
Failure by Twist
```

```
Failure by Spectral-IFOV-Shift
```
Characteristics of Spectral Calibration

**VNIR Spectral Variation Across the Field of View**
- VNIR Band 10: 447.892 nm
- VNIR Band 30: 651.278 nm
- VNIR Band 55: 905.511 nm

Wavelength Relative to FOV 128 (nm)

- 2.6 to 3.6 nm

**SWIR Spectral Variation Across the Field of View**
- SWIR Band 75: 892.35 nm
- SWIR Band 150: 1648.96 nm
- SWIR Band 225: 2405.63 nm

Wavelength Relative to FOV 128 (nm)

- .40 to .97 nm
• Why is this important if it is only a few nm?
  – Some spectral effects such as movement of the chlorophyll edge
due to stress is of the same order
  – Reproducibility of results due to variations in pointing
The Oxygen A Bands at 1 cm\(^{-1}\) & 1 nm Steps

Hyperion Smile Issue
Modtran O2-A "Band"

[Graph showing the transmittance (Modtran) against wavelength (nm) with peaks and valleys indicating the bands.]

68
Different Answers – Continuum?

Smile at O2-A Band
Rob, Barbara & FRome

![Graph showing differences in samples for Rob, Barbara, and FRome.](image-url)
Oxygen Line from Lake Frome Image
Image Quality
Image Quality

- Modulation Transfer Function
  - Pre-flight used knife edge and slit to measure Cross-track direction, Along-track was Cross-Track*2/pi
  - On-orbit used Ice Shelf & Lunar Limb (knife edge) and bridge (slit) to measure Cross-Track and Along-Track directly.

- Co-registration of VNIR and SWIR
  - Pre-flight used test bed to project a slit with a broad spectrum at multiple locations
  - On-orbit used combination of edges (Lunar, Ross), point sources (clouds, flares), ground control points

- Ground Sample Distance
  - Pre-flight measured IFOV using test bed
  - On-orbit triangulated marked features in well mapped scene
MTF Approach

• Calculate cross-track and in-track MTF using a step response and impulse response example

• Results of on-orbit analysis give good agreement with the pre-launch laboratory measurements
VNIR Imaging of “Starburst Pattern”

Pattern used to validate optics performance and level 1 processing

Level 0 data  Level 1 data
Dec 24, 2000. Bridge is the Mid-bay bridge near Destin, Florida. Bridge width (13.02 m) acquired and utilized in the MTF processing. Bridge angle small, every 5th line used to develop high resolution bridge image. MTF result at Nyquist is between 0.39 to 0.42; pre-flight measurement was 0.42.
Vertical and Horizontal MTF can be calculated from diagonal edge
MTF Calculation Process Summary

- MTF measurement on-orbit requires collection of scenes
  - step response: Ross Ice Shelf, Lunar Collect
  - impulse response: Bridge (Eglin, Cape Canaveral)

- Data Processing steps (simplified):
  1.) Define Edge Spread Function (ESF): Interlace adjacent lines from an object that is at a slight angle to the spacecraft motion
     - allows over sampling (sub-pixel) sampling

  2.) Calculate Line Spread Function (LSF):
     - edge technique: calculate an error function curve-fit to the ESF to derive the LSF, OR take the band-limited derivative of the ESF with a Tukey window.
     - slit technique: LSF is obtained from interlacing adjacent lines and de-convolving the profile with the slit width (in the frequency domain)

  3.) MTF is the Fourier Transform of the LSF
What’s Next

• We have looked at instrument characteristics and calibration

• Now what do we do to “correct” the data?
  – Radiometric
  – Geometric
  – Atmospheric
  – Uniformity (striping)
  – Other nuances

• How well can all this work?
Hyperion Data Flow

- **L0 Proc.**
  - L0 Science data
  - Ancillary data in engineering units
  - METADATA
  - Final product: level 1 data, metadata file attached

**Science Data:** Level 0 or Level 1 (radiometrically corrected) data products with VNIR and SWIR data frames combined. Includes solar, lunar calibrations, earth images, dark and light calibrations.

**Metadata:** Data about the science data. Information to support higher level processing, e.g., pre-flight characterization data.

**Ancillary Data:** Supporting data derived from spacecraft telemetry during image collection.
Level 1 Data Processing Flow

**Step 0**
*Flag pixels >4095:*
- Pre- & post-image darks
- Image
- Output to log file (MD15)

**Step 1**
*Smear correction*
- Apply smear correction to SWIR data
- Both dark and image files
- Log file generated (MD9)

**Step 2**
*Echo removal*
- Remove echo from smear corrected files
- Both dark and image files
- Log file generated (MD8)

**Step 3**
*Average dark files*
- Average frames of echo and smear corrected pre- and post-image dark files.
- Provide averaged pre-image (MD3A) and post-image (MD3B) dark files
- Compute average value of corrected pre- and post-image dark files. Log file generated reporting average values (MD5A for pre-image and 5B for post-image)

**Step 4**
*Background removal:*
- Subtract interpolated dark file from echo and smear corrected image file

**Step 5**
*Apply calibration:*
- Multiply dark subtracted image by lab gain file (MD7) to obtain radiometrically corrected (Level 1) data
- Multiply VNIR radiance by 40
- Multiply SWIR radiance by 80
- Log file generated (MD2)

**Step 6**
*Fixstripes:*
- Repair known bad pixels and output level 1_A file in signed integer format (.L1_A)
- Log file generated (MD10)

**Step 7**
*QA .L1_A image:*
- Display image in ENVI
- Complete QA form during evaluation (MD11)
- Add center wavelengths to ENVI header (.hdr) file (MD12)
Hyperion Processing: One Step Further

User Receive ‘L1A’ → Untar tape → Perform Quick Checks and Subset the Data → L1A subset

- VNIR bands 1:70 divide by 40.0
- SWIR bands 71:242 divide by 80.0

Return to Absolute Radiance, float

Subset data file, absolute radiance in float, SWIR aligned to VNIR, in a single file

Recombine Data files ‘L1P’

- SWIR aligned to VNIR, absolute radiance, float
- Align VNIR to SWIR Using swir_to_vnir.pts
  Or -1 cross track
  And 0-1 along track 0-1
Data Processing Levels

Level 1A: delivered data
Level 1P: Radiometric and VNIR/SWIR overlay
Level 2G: Geometric and VNIR/SWIR merge
Data Evolution - Processing L1A -> L2G

Level 1A: delivered data
Level 1P: Radiometric and VNIR/SWIR overlay
Level 2G: Geometric and VNIR/SWIR merge

VNIR/SWIR merge: Bands 9-55 & 77-220
Geo-Correction: L1A → L2G
A time series of polynomials allowing for geo-correction were developed by simultaneously fitting Ground Control Points (GCPs) collected in multiple images (date and sensor) using the MOSMOD module of microBRIAN. These polynomials allow the images to be resampled to geographic coordinates.

Fifty GCPs were collected between the base Aerial photography, an ETM image acquired 02-January-2001 (ETMjan), Hyperion VNIR, acquired 02-January-2001, 03-February-2001, 07-March-2001 (HypVNIRjan, HypVNIRfeb, HypVNIRmar), and Hyperion SWIR for those same dates (HypSWIRjan, HypSWIRfeb, HypSWIRmar).

MOSMOD both optimizes the images to geographic coordinates and the images internal registration.
High resolution (2m) digital aerial photographs acquired January 2001 used for creating positionally accurate field and rice bay GIS datasets.

Over 466km of linear road network and 129 well-defined points digitised with a Differential Global Positioning System (DGPS). These datasets can be used for geo-referencing Hyperion time series.
Coleambally Geo - Correction

• Six Hyperion ‘images’ (VNIR and SWIR for 02 Jan; 03 Feb; and 07 Mar 2001); GCPs collected in each Hyperion ‘image’ and 2m air photos

• Prior to fitting, outlier GCPs were identified and redefined; a linear polynomial was selected and a transitive chain method (MOSMOD) was used

Pixel Size
Avg X = 30.77m; Avg RMS error = 2.52 m
Avg Y = 30.49m; Avg RMS error = 6.87 m

Registration
Cross-track 12.9m (SD 0.6m)
Along-track 11.6m (SD 2.1m)

Presented at IGARSS 01
Preliminary Results / Geo-correction

Cross Track Differences Calculated Using Coleambally

- January Cross Track Difference
- February Cross Track Difference
- Linear (March Cross Track Difference)
- March Cross Track Difference
- Linear (January Cross Track Difference)
- Linear (February Cross Track Difference)

Field of view or Pixel Number

delta

-0.75 to -1.25
Preliminary Results / Geo-correction

Along Track Differences Calculated Using Coleambally

- January Along Track Difference
- March Along Track Difference
- February Along Track Difference

Linear (January Along Track Difference) - Linear (February Along Track Difference) - Linear (March Along Track Difference)
Aligning the VNIR and SWIR

- VNIR and SWIR are independent spectrometers and focal plane arrays

- For the user to align the SWIR to the VNIR apply
  - constant cross track shift of –1 pixel
  - linear along track shift of 0 pixels at field-of-view 0 and 1 pixel at field of view 256.
Reflectance Processing
Hyperion Band 223 (2385 nm)

1. De-striping
2. Atmospheric Correction (FLAASH-based)
3. MNF Smoothing

Input Radiance Image

Output Reflectance Image
Should Atmospheric Correction be Performed?

• Hyperspectral remote sensing has effects of atmosphere as well as earth materials

• The result is easily recognized and its worst effects can be avoided

• Should you atmospherically correct?
  – Some applications do not need it (classification, MNF, CVA, exploration)
  – Some can benefit from it (standardization)
  – Some need it (modeling)

• The data are there to use – do not be put off nor wait until all the “problems” have been solved!
A Simple Atmospheric Model

\[ L_t = \frac{1}{\pi} E_g T \frac{\rho_t}{1 - s \rho_t} + L_p \]

**Modelling:** \[ \rho_t \rightarrow L_t \]

**Atmospheric Correction:** \[ L_t \rightarrow \rho_t \]
Not so frightening after all?! (DLBJ)
Atmospheric Correction Codes

Hatch, FLAASH, ACORN
Atmospheric Correction Codes

Hatch, FLAASH, ACORN
FLAASH – ASD Comparison

Soil1 Means

Stubble1 Means

Soil

Stubble
Atmospheric Avoidance

“Stable” set of bands can be considered if application permits

<table>
<thead>
<tr>
<th>Bands</th>
<th>Wavelength (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10-57</td>
<td>448-926</td>
</tr>
<tr>
<td>81-97</td>
<td>953-1114</td>
</tr>
<tr>
<td>101-119</td>
<td>1155-1336</td>
</tr>
<tr>
<td>134-164</td>
<td>1488-1790</td>
</tr>
<tr>
<td>182-221</td>
<td>1972-2365</td>
</tr>
</tbody>
</table>
Yerington, NV: pixel 62, 35

HATCH
ATREM
Cirrus Cloud Detection Over Mojave Desert

Visible Image

Image from 1380 nm
Accounting For Water Vapor

Measured ➔ Modeled

Water Vapor Parameter map

15.92 precipitable mm

0.00 mm PW
0.10 mm PW
1.00 mm PW
10.0 mm PW
50.0 mm PW

Radiance (µW/cm²/nm/sr)

Wavelength (nm)
Yerington, NV: Water Vapor Images

ATREM

HATCH

0.68 cm  1.45 cm
0.60 cm  1.10 cm
Stripes in the Image

- Stripes may not impact analyses

- Three spatial frequency scales occur:
  - High
  - Medium
  - Low

- Is there a temporal effect?
Approaches for De-Striping

1. Set a column mean (intensity) equal to the global mean.

2. Set the column mean and standard deviation equal to the global mean and standard deviation.

3. Use a locally-based mean and standard deviation in spatial dimensions.

4. Use a locally-based mean and standard deviation in spatial and spectral dimensions.
Effects of ‘Global’ and ‘Local’ De-Striping

- Typical spatial scales are pixel, surface features and swath

- ‘Global’ removes streaks and low-frequency effects (but ‘Global’ can bias the results) – it depends on the surface cover characteristics along the column

- ‘Local’ removes spikes – especially in the VNIR but leaves in the low frequency effects such as the smile radiance effect.

- VNIR and SWIR noise are different. Should be treated separately.
Striping

VNIR

SWIR
De-streaking Approach

\((m_{ij}, s_{ij})\) are column averages for sample i band j

\((\bar{m}_{ij}, \bar{s}_{ij})\) are reference values for them

\[ x_{ijk} \rightarrow \alpha_{ij} x_{ijk} + \beta_{ij} \quad \text{sample i band j line k} \]

\[ \alpha_{ij} = \frac{\bar{s}_{ij}}{s_{ij}} \]
\[ \beta_{ij} = \bar{m}_{ij} - \alpha_{ij} m_{ij} \]

Selection of reference values determines whether the de-striping is local or global
Global De-striping MNF1 and 15

Original Data

De-Streaked Data

MNF 1 Radiance

MNF 15 Radiance

MNF 1 Radiance

MNF 15 Radiance
Effects of ‘Global’ and ‘Local’ De-Striping

Change Analysis

• Global De-streaked minus Original Radiance
• Local De-streaked minus Original Radiance

• Global De-streaking removes ‘smile’ effect and streaks
• But also alters mid frequency spatial effects in the data

• Local de-streaking removes stripes but leaves the ‘smile’ effect in
Other Thoughts
Other Effects - Directional Illumination

Midday

Early Morning
Through the Glass Darkly – or “no free lunch”

- Hyperspectral data is fighting for photons
- Small pixels & many narrow bands lead to lower SNR
- The technology is good but there is a limit

- Landsat FWHM ranges from 70 in VNIR to 300-400 in the SWIR

- Hyperion Bands have an FWHM and step of 10 nm
## Sensor Characteristics

<table>
<thead>
<tr>
<th>SAC-C</th>
<th>ETM+</th>
<th>ALI</th>
<th>MODIS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Band 1</td>
<td>490.5</td>
<td>19.3</td>
<td>Band 1</td>
</tr>
<tr>
<td>Band 2</td>
<td>548.8</td>
<td>19.2</td>
<td>Band 2</td>
</tr>
<tr>
<td>Band 3</td>
<td>656.6</td>
<td>58.0</td>
<td>Band 3</td>
</tr>
<tr>
<td>Band 4</td>
<td>812.9</td>
<td>34.2</td>
<td>Band 4</td>
</tr>
<tr>
<td>Band 5</td>
<td>1598.5</td>
<td>106.6</td>
<td>Band 5</td>
</tr>
<tr>
<td>Band 6</td>
<td>2208.1</td>
<td>251.3</td>
<td>Band 6</td>
</tr>
<tr>
<td>Pan</td>
<td>719.9</td>
<td>319.5</td>
<td>Pan</td>
</tr>
</tbody>
</table>

**Table Notes:**
- SAC-C
- ETM+ (bands 1-5, 812.9 nm)
- ALI (bands 1-5, 812.9 nm)
- MODIS (bands 1-5, 812.9 nm)
SAC-C VNIR Bands vs ALI, ETM+

ALI

ETM+

Red: SAC-C
Blue: Other Instrument
# Hyperspectral Imaging
## Applications & Benefits

<table>
<thead>
<tr>
<th>Application</th>
<th>Existing Satellite Capabilities (SPOT, LandSat)</th>
<th>Hyperion Capability</th>
<th>Perceived Benefits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mining/Geology</td>
<td>Land cover classification</td>
<td>Detailed mineral mapping</td>
<td>Accurate remote mineral exploration</td>
</tr>
<tr>
<td>Forestry</td>
<td>Land cover classification</td>
<td>Species ID</td>
<td>Forest health/infestations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Detail stand mapping</td>
<td>Forest productivity/yield analysis</td>
</tr>
<tr>
<td>Agricultural</td>
<td>Land cover classification, Limited crop mapping, Soil mapping</td>
<td>Foliar chemistry</td>
<td>Forest inventory/harvest planning</td>
</tr>
<tr>
<td>Environmental Management</td>
<td>Resource meeting, Land use monitoring</td>
<td>Tree stress</td>
<td>Yield prediction/commodities</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Crop differentiation</td>
<td>crop health/vigor</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Crop stress</td>
<td>Contaminant Mapping</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Chemical/mineral mapping &amp; analysis</td>
<td>Vegetation Stress</td>
</tr>
</tbody>
</table>
Hyperspectral Image Provides Forestry Detail

LandSat Analysis

Hyperspectral Analysis

Legend

No Data
Hardwood
Softwood
Grass / Fields

Legend

No Data
Open field
Red Maple
Red Oak
Mixed Hardwood
Hardwood/Conifer Mix
White Pine

Hemlock/Hardwood Mix
Mixed Conifer
Norway Spruce
Red Pine
Spruce Swamp
Hardwood Bog

Analysis by Mary Martin University of New Hampshire
Hyperspectral Image Provides Geological Data

GEOTHERMAL AREA
(no specific mineral information)

CALCITE
(gold bearing quartz)

MULTISPECTRAL ANALYSIS

HYPERSONICAL ANALYSIS

Analysis courtesy AIG Limited Liability Company
Analysis of Red Edge Shift- sample process
Now, Let’s Classify the Data!
Pixel-Based Supervised Classification Methods for Hyperspectral Data
Important Issues for Classification of Hyperspectral Data

• Impact of noise and calibration
  – Lower SNR than corresponding multispectral data
  – Striping in data acquired by pushbroom sensors
  – Data susceptible to atmospheric artifacts and sensor anomalies

• Large input space
  – Enormous number of parameters to estimate
  – Sparse data in hyperspectral domain
  – Adjacent bands are often highly correlated
  – Quantity of training data typically limited

• Potentially large output space
  – Possible overlapping spectral signatures
Approaches to Resolve Issues

• High dimensional input spaces
  – Simplify or reduce dimension of input space via extraction
    • Subset feature selection
    • Projection of original features on a lower dimensional space
  – Regularize covariance matrix of input space
  – Utilize unlabelled samples via semi-supervised learning
  – Develop ensembles of classifiers

• High dimensional output spaces
  – Output space decomposition

• Utilization of nonparametric classifiers
Dealing with High Dimensional Input Spaces
Feature Selection/Extraction

• **Goals**
  – Reduce no. observations required to train the classifier
  – Reduce computational complexity
  – Improve classification accuracy

• **Desired Properties of Extractors**
  – Class dependent
  – Exploit band ordering
  – Transformations should maximize discrimination among classes
• Selection vs extraction
  – Selection of subset of original features
    • Less redundancy, but potentially some information loss
    • Better interpretability (domain knowledge)
    • Possibly improved generalization
  – Extraction
    • Typically computationally superior to selection
    • No loss of information as all features retained
    • Resultant features unrelated to physical phenomena
Approaches to Feature Extraction

• Principal Component Transform [Anderson (1984)]
  – Orthogonal linear combinations of the \( p \) original bands with maximum variance. At pixel \( x \),

\[
Z(x) = \{Z_i(x), i = 1,...p\}^t
\]

– Formulated as an optimization problem

\[
\text{Max } a_i^t \Sigma a_i
\]
subject to

\[
a_i^t a_i = 1, \quad i = 1,....p
\]

\[
E\left\{a_i^t Z(x) Z(x)^t a_j \right\} = 0 \quad \forall i \neq j, \quad j = 1,...i
\]

– Solution is a subset of successively ordered eigenvalues of the covariance matrix \( \Sigma \); \( a_i \) are corresponding eigenvectors
– Characteristics of Principal Component Transform
  • Developed to maintain variance of data in smaller set of orthogonal inputs
  • Order of components not necessarily related to data quality
  • Components have no physical relationship to the targets
  • PCT data are image content dependent (poor generalization)
  • Does not exploit the adjacency of correlated bands
  • Maximum variability has no inherent relationship to criterion for classification
  • Sensitive to outliers
Principal Components of AVIRIS Data
• Segmented Principal Component Transformations [Jia and Richards (1999)]
  – Detect boundaries in correlation matrix, compute PCT’s and select subset. Further prune using Bhattacharya distance
  – Characteristics of SPCT
    • Restricted to two-class problems
    • Utilizes adjacent band correlation structure
    • Not related to discrimination measures
    • Sensitive to outliers
• Maximum Noise Fraction Transform [Green, Berman, Switzer, and Craig (1988)]
  – Uncorrelated linear combinations of bands that maximize “noise fraction” for each band.

\[
Z(x) = \{Z_i(x), i = 1, \ldots, p\}^t = S(x) + N(x)
\]

where

\[
\Sigma_z = \Sigma_s + \Sigma_n, \quad \Sigma_z, \Sigma_n \text{ known}
\]

– Define “noise fraction” at each pixel \(x\)

\[
NF_i(x) = \frac{\text{Var}\{N_i(x)\}}{\text{Var}\{Z_i(x)\}}, \quad i = 1, \ldots, p
\]
– Formulated as an optimization problem similar to PCT, except that orthogonal linear combinations are sought to maximize the Noise Fraction

\[
\text{Max } \left[ NF_i (x) = a_i' \Sigma_N a_i / a_i' \Sigma Z a_i \right]
\]

subject to

\[
a_i' \Sigma Z a_i = 1 \quad i = 1, \ldots p
\]

\[
E \left\{ a_i' Z(x) Z(x)' a_j \right\} = 0 \quad \forall i \neq j, \ j = 1, \ldots i
\]

– Optimal weights: \( a_i^* = \text{eigenvectors of } \Sigma_N \Sigma^{-1} \)
- Characteristics of MNF Transform
  - Developed to filter noisy bands, then transform data back to original domain
  - Ordering related to quality of transformed data
  - Invariant under data rescaling
  - Does not exploit the adjacency of correlated bands or class specific information
  - Resultant bands of transformed data *may* be related subjectively to phenomena, but are not related to discrimination between classes
MNF Bands for AVIRIS Data
• Fisher’s Linear Discriminant [Anderson (1984)]
  – Seeks projection matrix $A^*$ to maximize separation between classes $\Omega = \{\omega_i, i = 1,..C\}$

$$\text{Max } \frac{|A'BA|}{|A'WA|}$$

where

$$B = \sum_{\omega \in \Omega} |X_{\omega}| (\mu_{\omega} - \bar{\mu})(\mu_{\omega} - \bar{\mu})' \quad W = \sum_{\omega \in \Omega} \sum_{x \in X_{\omega}} (x - \mu_{\omega})(x - \mu_{\omega})'$$

$A^*$ = generalized eigenvectors of $W^{-1}B$ (dim $\leq C - 1$)
Characteristics of Fisher’s Linear Discriminant

- Requires a trained classifier
- Poor discrimination well between classes with similar means
- Classes with “different” variance dominate the combination
- Ignores band ordering
- Problematic for hyperspectral data with small training sample due to required matrix inversion
• Decision Boundary Feature Extraction [Lee and Landgrebe (1997)]
  – Computes a decision boundary that determines the projection of the hyperspectral space for a two-class problem
    • C-class problem solved using wtd sum of 2-class decision boundary feature matrices
  – Requires trained classifier
  – Ignores band ordering

• Projection Pursuit [Jimenez and Landgrebe (1999)]
  – Compute projection for two classes based on Bhattacharya distance
  – Subsets constrained to be smaller groups of adjacent bands
  – One set of features selected for all classes
  – C-class problem uses sum of discriminatory power for pairs (problematic for large no. of classes)
• **Best Bases Band Aggregation [Kumar et al. (2002); Morgan et al. (2004)]**
  – Adjacent bands merged based on correlation measure (or product of correlation and Fisher discriminant)
  – Implemented in Top-Down and Bottom-Up approaches
  – Characteristics
    • Exploits band adjacent correlations, maintains unique individual bands
    • Features are class specific
    • Affects signal to noise ratio in a non-uniform way
    • Mitigates impact of small training sample problem
• Polyline Feature Extraction [Henneguelle et al. (2003)]
  – Approximate mean spectrum via piece-wise polynomials (usually linear).
  – Select break points via top-down or bottom-up methods and determine parameter values (Linear: slope, midpoint of segment)
  – Parameters of piece-wise functions become input features to classifier
 Characteristics

• Exploits band adjacency
• Features are class dependent
• Computationally efficient
• Provides approximation to derivatives
• Features robust for generalization
• Sensitive to stopping criterion (no. of bands vs error in approximation)
• **Discrete Fourier Transform**
  – Describes band sequence via trigonometric functions
  – Subset of modes selected to represent features
  – Characteristics
    • Multiscale representation of spectral signature
    • Difficult to determine optimal modes to represent original problem
    • Computation unrelated to class discrimination
    • Modes may contain domain knowledge
• Discrete Wavelet Transform [Bruce et al. (2001, 2002); Kaewpijit et al. (2003)]
  – Multiscale method that accounts for both frequency and position
  – Describes band sequence via discrete wavelet transform
  – Implemented in conjunction with a selection method
  – Characteristics
    • Provides robust features
    • No. of inputs must be a power of 2
      – Accomplished via padding or interpolation
    • Criterion unrelated to class discrimination
Approaches to Feature Selection

• Selection techniques can be implemented individually, or in conjunction with extraction methods

• Methods that guarantee optimal solutions
  – Exhaustive search
  – Branch & bound (only if objective function is monotonic)

• Heuristic methods
  – Characteristics
    • Typically produce sub-optimal solutions
    • Developed to reduce computational complexity
    • Traditionally implemented via sequential forward selection or backward elimination
Heuristic Feature Selection Methods

• Sequential Floating Search Methods [Pudil et al. (1994); Serpico and Bruzzone (2001)]
  – Evaluate a number of features for elimination/forward selection for each feature (group) added/eliminated

• Tabu Search [(Zhang and Sun (2002); Korycinski et al. (2003)]
  – Developed for solution of combinatorial optimization problems
  – Non-greedy heuristic that adaptively and reactively adjusts search space, flexible in number of features evaluated at each iteration

• Various simulated annealing and genetic algorithms [Siedlecki and Sklansky (1988, 1989)]
Regularization of Covariance Matrix

- **Methods to regularize estimated covariance matrix**
  - Pseudo-inverse [Fukanaga (1990); Skurichina and Duin (1996); Raudys and Duin (1998)]
    - Poor performance when ratio of amount of training data to input dimension is small
    - Exhibits “peaking” effect at $|X| = d/2$
  - Shrinking and pooling covariance matrices [Tadjudin and Landgrebe (1999)]
    - Reduce variance of estimates, but introduce bias
• Semi-supervised learning
  – Augment existing training data with unlabeled samples [Jeon and Landgrebe (1999); Tadjudin and Landgrebe (2000); Jackson and Landgrebe (2001)]
    • Classify unlabeled observations and update estimates, usually via EM algorithm
    • Alternatively perturb sample data
  – Effective for generalization and knowledge re-use
  – Sensitive to original training samples, impacted by outliers
Dealing with Large Output Spaces
Multiclassifier Systems

**Multiclassifier Systems**
- Ensembles of base classifiers that are learned individually to produce an aggregated predictor
- Characteristics
  - Component classifiers often weak
  - Decision boundaries for individual classifiers typically simple
  - Generalization typically superior, if classifiers are diverse

**Goals**
- Improve classification accuracy and generalization
- Reduce complexity and improve interpretation
- Enhance transferability of classifiers to new problems or beyond spatial extent of training data
- Mitigate the impact of sensor artifacts
- Improved utilization of small quantity of training data

**Challenges**
- Achieving highly diverse, but relevant classifiers
- Maintaining interpretability
Multiclassifiers from Input Space Decomposition

• Selecting sub-samples of original data, creating classifiers, and developing a classifier for each sample. Combine results from individual classifiers
  – Perform poorly for extremely small sample sizes as ensemble methods cannot overcome lack of diversity
• Bagging [Brieman (1996)]
  – Create multiple samples with replacement; develop individual classifiers
• Arcing
  – Adaptively reweighting and combining (e.g. boosting) [Freund and Schapire (1996); Dietterich (2000)]
  – Simple random sub-sampling [Ho (1998); Skurichina and Duin (2002)] (See previous slides for details)
Random subspace selection methods [Ho (1998); Skurichina and Duin (2002)]

- Utilized in conjunction with multiclassifier systems
  - Randomly sample input features for each classifier
  - Combine outputs of multiple classifiers via voting or maximizing aposteriori probabilities

- Characteristics
  - Provides diversity and robust classifiers with good generalization
  - Reduces band redundancy, but ignores band adjacency, unless implemented with band aggregation (which reduces diversity)
  - Mitigates impact of small sample size problems and stabilizes parameter estimates
  - Computationally expensive
  - Does not provide domain knowledge on feature characteristics
  - Mitigates the impact of anomalous sensor behavior (striping)
Multiclassifier Systems: Output Space Decomposition

• **C one-vs-rest problems** [Anand et al. (1995)]
  – Classify individual classes vs the mixture of the “rest”
  – Can result unbalanced priors for Bayesian methods if sample sizes for some classes extremely small
  – Does not exploit class affinity, so convergence can be very slow

• **Pairwise classifiers** [Crawford et al.(1999); Kumar et al. (2001); Furnkranz (2002)]
  – Selects best set of features for class pairs, performs classification, then combines results for C-class problem via voting or maximum *aposteriori* rule
  – Can incorporate pairwise feature extraction
  – Yields \( \binom{C}{2} \) classifiers
  – Potential coupling of outputs
Error correcting output codes [Dietterich and Bakiri (1995); Rajan and Ghosh (2004)]

- C-class problem encoded as $\bar{C} \gg C$ binary classifiers which are combined by voting
- Each class assigned according to $\bar{C} \times C$ code book
- Observations labeled as class with code closest to code formed by outputs of $\bar{C}$ classifiers

Characteristics
- Implemented with user-selected classifiers
- Feature selection/extraction tunable to pairwise classifiers
- Output dependent on classifier separation
- Reduces impact of small sample sizes
- Does not exploit natural class affinities

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
• **Support vector machines** [Vapnik (1995); Hsu and Lin (2002)]
  – Maximize the margin between two class samples instead of accuracies
  – Characteristics
    • Formulated as an optimization problem for a binary classifier
    • Nonparametric
    • Potentially high dimensional decision boundary
    • Classification accuracies dependent on typically slow tuning
    • Tends not to overtrain, which leads to high classification accuracies and good generalization
Hierarchical Decomposition Methods

- Characteristics
  - Adopt a sequential “divide and conquer” approach involving decomposition based on input or output spaces
  - Maintain natural groupings with useful domain knowledge

- Hierarchical Decision Trees
  - Tree constructed sequentially based on series of binary questions and splitting rule which maximizes decrease in impurity of parent and child nodes
    - CART: Best split based on linear combinations of input features; Gini impurity index, post pruning [Brieman et al. (1984)]
    - C4.5: Impurity index based on entropy [Quinlan (1986)]
– Binary Hierarchical Classifier [Kumar et al. (2002); Morgan et al. (2004)]
  • Output space decomposition approach
    – $C$ leaf nodes, $C-1$ internal nodes
  • Top down tree constructed via deterministic simulated annealing
  • Feature extraction/selection tunable to each internal node
  • Classifier specific to each internal node (Fisher discriminant, SVM)
  • Mitigates small sample problems
  • Exploits natural class affinities
Let’s look at an example
Study Site: Okavango Delta, Botswana
Classification scheme consists of 14 landcover classes selected using IKONOS imagery, aerial photography, and field campaigns.
Hyperion Hyperspectral Data Inputs

- Hyperion hyperspectral data
  - Acquired by Earth Observer-1 Satellite May 31, 2001
  - Converted to reflectances, destriped, georeferenced

<table>
<thead>
<tr>
<th>Class #</th>
<th>Class Name</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>water</td>
<td>270</td>
</tr>
<tr>
<td>2</td>
<td>hippo grass</td>
<td>101</td>
</tr>
<tr>
<td>3</td>
<td>floodplain grasses1</td>
<td>251</td>
</tr>
<tr>
<td>4</td>
<td>floodplain grasses2</td>
<td>215</td>
</tr>
<tr>
<td>5</td>
<td>reeds</td>
<td>269</td>
</tr>
<tr>
<td>6</td>
<td>riparian</td>
<td>269</td>
</tr>
<tr>
<td>7</td>
<td>firescar</td>
<td>259</td>
</tr>
<tr>
<td>8</td>
<td>island interior</td>
<td>203</td>
</tr>
<tr>
<td>9</td>
<td>acacia woodlands</td>
<td>314</td>
</tr>
<tr>
<td>10</td>
<td>acacia shrublands</td>
<td>248</td>
</tr>
<tr>
<td>11</td>
<td>acacia grasslands</td>
<td>305</td>
</tr>
<tr>
<td>12</td>
<td>short mopane</td>
<td>181</td>
</tr>
<tr>
<td>13</td>
<td>mixed mopane</td>
<td>268</td>
</tr>
<tr>
<td>14</td>
<td>exposed soils</td>
<td>95</td>
</tr>
</tbody>
</table>
Design of Experiment for Random Forests

• Create 10 independent stratified samples of training/test data
  – Select 75%:25% of training data for training and test
  – Subsample from 75% to achieve 50, 15% sample size; test remains constant
  – Select spatially removed test sample

• Create BHC tree
  – Select bagged sample
  – Randomly sample input features
    • \((D/5)\) or \(\min(D/5, 20)\)
  – Develop tree

• Grow BHC Random Forest

• Compare to BB-BHC, RS-BHC, BB RS BHC, RF CART
Adequate Destriping is Critical for Classification

Classification Results with Poor Destriping
Classification Results: BHC Algorithm
Okavango Hyperion

(a) Subset of Hyperion data (Bands 51, 149, 31),

(b) Classified image of Hyperion data
Classification Accuracies, Independent Test Set

Classification Accuracy, Ind Test Set (%)

Accuracy

Training Sample Fraction

Classification Accuracies, Independent Test Set

Std. Deviation, Classification Accuracy

Training Sample Fraction

0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8
Next Steps - an interactive discussion