Benchmarking of the HyspIRI VSWIR Compression, Level 1 and Level 2 Algorithms

Sarah Lundeen, Robert O. Green, Bo-Cai Gao, Didier
Overview

- Challenge and Objective
- HyspIRI VSWIR data rates and volumes
- On-board VSWIR compression
- Ground processing Level 1 data
- Ground processing Level 2 data
- Summary and Conclusions
Challenge and Objective

- HyspIRI is a high data rate and volume mission at ~ 1 terabit per orbit and 14 orbits per day

- This data rate lead to challenges of handling the data on the satellite, downlinking the data, and processing the data on the ground

- HyspIRI has mission baseline solutions at all elements of data handling

- The objective of the testing and benchmarking is to demonstration the viability of the solutions using HyspIRI like data

- This work validates the cost estimates and reduces the risk of the HyspIRI
HyspIRI Mission Architecture

**Orbit**
626 km Altitude
10:30 AM LMT
Descending Node

**Calibration**
- Daily Solar View (VSWIR)
- Monthly Lunar View (VSWIR + TIR)
- Black Body and Deep Space Views (TIR)
- Vicarious US and International

**SSR**
1 Tb for Simple Store and Forward

**VSWI**
- Revisit
19 Day Revisit

**TIR**
- 5 Day Revisit

**IPM**
- 20 Mbps Direct Broadcast

**Downlink**
800 Mbps X Band To Northern Stations (Svalbard and Alaska)

**Fiber Link**
Data Transmitted to Processing Center within 2 weeks of downlink

**Global Coverage**
60 m Resolution: Land + Shallow Water (Benthic & AV)
HyspIRI Downlink Data Volume

Total downlinked data volume for the 3 year mission: 5024 Tbits

Avg (Tb) Min (Tb) Max (Tb)

<table>
<thead>
<tr>
<th>Rate</th>
<th>On-board Compression</th>
</tr>
</thead>
<tbody>
<tr>
<td>VSWIR_land</td>
<td>804.1 Mb/s</td>
</tr>
<tr>
<td>VSWIR_shallow</td>
<td>865.9 Mb/s</td>
</tr>
<tr>
<td>VSWIR_ocean</td>
<td>3.9 Mb/s</td>
</tr>
<tr>
<td>TIR_land</td>
<td>130.2 Mb/s</td>
</tr>
<tr>
<td>TIR_shallow</td>
<td>130.2 Mb/s</td>
</tr>
<tr>
<td>TIR_ocean</td>
<td>0.6 Mb/s</td>
</tr>
</tbody>
</table>

Per Day 4.64 3.59 5.29
Per Orbit 0.31 0.00 0.81

Baseline selected to minimize system level cost and risk
On-board storage capacity
- 1 Tb
- 0.31 Tb/orbit
WorldView-1 and -2 have 2.2 Tb SSR
- WorldView1: 0.33 Tb/orbit
  - Different downlink strategy requires larger SSR than HyspIRI
- WorldView2: 0.52 Tb/orbit
30% margin added to calculated required SSR size
Fast Lossless Compression Algorithm

- **Objective:** State-of-the-art lossless compression, with low complexity (i.e., fast)
- **Approach:** Predictive compression that adapts to the data via the sign algorithm (a variation of the least mean square (LMS) algorithm) (see boxes below)
- **Compared** to Transformed-based compression techniques (such as DCT, Wavelet transform), this approach:
  - requires fewer arithmetic operations and less memory, simplifies data handling, and is more straightforward to implement (in software, DSP, or hardware)
  - yields significantly faster lossless compression
  - But provides only lossless (and potentially near-lossless) compression

### Predictive Compression

- Encodes samples one-at-a-time, typically in raster scan order
- Estimates sample value probability distribution from previously encoded samples. These estimates are used to efficiently encode the sample value.
- The difference between an estimated sample value in the actual sample value is encoded in the compressed bitstream.

The sign algorithm and the LMS algorithm are members of a family of low complexity adaptive linear filtering techniques.

- Used extensively in signal processing applications
- Used for compression of audio data
- Not previously well studied for image or hyperspectral data compression
FL Compression Testing with AVIRIS-Classic
Tests using 19 uncalibrated AVIRIS data sets:
- original sample size: 12 bits/sample
- data size: (614 x 512 pixels x 224 bands) (680 x 512 pixels x 224 bands)

Methods:
JPEG-LS: is most efficient for 2D; USES uses chip; ICER-3D SOA (MER rovers)
FL Compression Algorithm Features

• **Performance:** good compression effectiveness

• **Robust:** requires no training data or other specific information about the nature of the spectral bands for a fixed instrument dynamic range

• **Simple:** well-suited for implementation on FPGA hardware and easily parallelizable

• **Low computational complexity.** required operations per sample are:
  – 6 integer multiplications
  – 25 integer addition, subtraction, or bit shift operations
  – Golomb coding operations

• **Modest memory requirement:** enough to hold one spatial-spectral slice of the data (e.g., ≤300 Kbytes for AVIRIS data with 224 bands and 680 samples/line)

• **Instrument:** well-suited to push broom instruments
Early Test of AVIRIS-NG Like Data

- This test shows compression of > 5X
- HyspIRI baseline is 3X
Real-time Data Compression for HyspIRI

- Developed an FPGA implementation of the Fast Lossless (FL), a state-of-the-art lossless HSI compression algorithm providing compression performance up to 4:1.
- Implemented on a commercial Virtex 5 (equivalent to V5 Rad-hard device). Compresses one sample every clock cycle, a speed of 40 MSample/sec with total power of 700 mW.
- FL compression implementation is currently being tested in National Instruments PXI environment which includes a PXIe-7966R board with Xilinx Virtex-5 SX95T and two 256MBytes DRAMs. The test system is connected to the airborne AVIRIS-NG HSI instrument and will be compressing HSI data in real-time on the plane.

Carnegie Airborne Observatory:
A sample image
640(width) x 512(length) x 427(bands), 13 bits per sample
Compression rate: 2.366 bits/sample

### Virtex-5 Device Utilization Summary (CBE)

<table>
<thead>
<tr>
<th>Logic Utilization</th>
<th>Used</th>
<th>Available</th>
<th>Utilization</th>
</tr>
</thead>
<tbody>
<tr>
<td># Slice Registers</td>
<td>15715</td>
<td>58880</td>
<td>26.7%</td>
</tr>
<tr>
<td># Slice LUTs</td>
<td>24155</td>
<td>58880</td>
<td>41%</td>
</tr>
<tr>
<td># Block RAM</td>
<td>76</td>
<td>244</td>
<td>31.1%</td>
</tr>
<tr>
<td># DSP48s</td>
<td>6</td>
<td>640</td>
<td>0.9%</td>
</tr>
</tbody>
</table>
HyspIRI VSWIR Level 1 and 2 Algorithm Benchmarking
Level 0-1 AVIRIS-CL/AVIRIS-NG Data Flow

AVIRIS Instrument Ground Data System

- Raw Flight Science Data
- Navigation Data

Level 0 Processing
- Raw Science Data Store
- Raw Image Data Cubes (DN)
- Quicklook Program
- Quicklook Products
- Quicklook Products Store
- L0 Products Store
- Performance Evaluation and Data Quality Check Program
- PEP & DQC Products
- PEP & DQC Store
- Level 0 Products Store
- Radiometric Calibration Program
- Orthorectification Program
- Ortho Products
- Ortho Products Store
- Flight Locator Program
- Flight Locator Products
- Flight Locator Store
- Level 1 Processing
- Flight Locator Tool Web Server
- NAS Backup Storage

Manual Process
Automated Process
Product Generation Executive (PGE)
Data Products Store
Level 1 and Level 2 and The Signal

Upwelling Spectral Radiance

Imaging Spectrometer Measured Data

Calibrated Imaging Spectrometer Radiance

Surface Spectral Reflectance
Atmospheric Correction

The measured radiance at the satellite level can be expressed as:
\[ L_{\text{obs}} = L_a + L_{\text{sun}} \times t \times \rho \]  

(1)

- \( L_a \): path radiance;
- \( \rho \): surface reflectance;
- \( L_{\text{sun}} \): solar radiance above the atmosphere;
- \( t \): 2-way transmittance for the Sun-surface-sensor path

Define the satellite apparent reflectance as
\[ \rho_{\text{obs}}^* = \pi \frac{L_{\text{obs}}}{(\mu_0 E_0)} \]  

(2)

\[ \rho_{\text{obs}}^* = T_g \left[ \rho_a + t \rho / (1 - \rho s) \right] \]  

(3)

By inverting Eq. (3) for \( \rho \), we get:
\[ \rho = \left( \frac{\rho_{\text{obs}}^*}{T_g} - \rho_a^* \right) / \left[ t + s \left( \frac{\rho_{\text{obs}}^*}{T_g} - \rho_a^* \right) \right] \]  

(4)

AVIRIS-NG Orthorectification Test

Raw Data

Successful Orthorectification

AVIRIS-NG Orthorectified

NASA
**Server**
- OS: 64-bit Windows Server 2008
- Processor: Dual Intel Xeon X5560 2.80GHz Quad-Core Processors
- Memory (RAM): 24.0 GB ECC DDR3

**Storage**
- 4 RAID arrays populated with 24, 2TB 7200 RPM 64 MB Cache Enterprise Class SATA II HDD’s
- 4Gb Fibre Channel to SAS/SAS(SATA) Controller RAID Unit
- Total of 156 TB usable storage

**Backup Storage**
- iNAS 36 bay populated with (36) 4TB 7200 RPM 64 MB SATA II HDDs – Total 116 TB usable storage
<table>
<thead>
<tr>
<th>Data Set Type</th>
<th>Windows Server 2008</th>
<th>Beowulf Cluster – Single Node</th>
<th>Beowulf Cluster – 26 Nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVIRIS-Classic</td>
<td>- Rad Cal: 353 GB in 761 min ➔ 7.9 MB/s</td>
<td>- Rad Cal: 3.6 GB in 11 min ➔ 5.6 MB/s</td>
<td>- Rad Cal: 90.1 GB in 11.5 min ➔ 133.7 MB/s</td>
</tr>
<tr>
<td></td>
<td>- Ortho: 35.7 GB in 63 min ➔ 9.6 MB/s</td>
<td></td>
<td>- Ortho: ~390 MB/s</td>
</tr>
<tr>
<td></td>
<td>- ATREM: 732.6 GB in 1375 min ➔ 9.1 MB/s</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Traditional Beowulf Cluster - 27-node, 336 Processor Cores

**Head Node**
- OS: LinuCentOS 6.0
- Processor: (2) Intel Westmere E5650 2.66GHz, Hexa-Core
  Processors – Total of 12 cores
- Memory (RAM): (6) 8GB ECC DDR3 – Total 48GB
- Local Storage: (2) 1TB Mirrored HDD’s – Total 1 TB usable storage

**Compute Nodes**
- OS: LinuCentOS 6.0
- Processor: (2) Intel Westmere E5650 2.66GHz, Hexa-Core
  Processors – Total of 12 cores
- Memory (RAM): (6) 8GB ECC DDR3 – Total 48GB
- Local Storage: (2) 1TB striped disks – Total 2 TB usable storage

**Storage Node**
- OS: LinuCentOS 6.0
- (16) 3TB enterprise server disks configured under RAID-6 array – Total 39TB usable storage after configuration
- Processor: (2) Quad-core Intel Westmere E5620 2.40GHz – Total 8 cores
- Memory (RAM): (6) 4GB ECC DDR3 – Total 24GB

**Backup Storage**
- iNAS 36 bay populated with (36) 4TB 7200 RPM 64 MB SATA II HDDs – Total 116 TB usable storage
## Data Set Type

<table>
<thead>
<tr>
<th>Data Set Type</th>
<th>Windows Server 2008</th>
<th>Beowulf Cluster – Single Node</th>
<th>Beowulf Cluster – 26 Nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVIRIS-Next Generation</td>
<td>- Rad Cal: 13 GB in 96 min ➔ 2.3 MB/s</td>
<td>- Rad Cal: 110.83 GB in 179 min ➔ 10.6 MB/s</td>
<td>- Rad Cal: ~276 MB/s</td>
</tr>
<tr>
<td></td>
<td>- Ortho: 31.9 GB in 185 min ➔ 2.9 MB/s</td>
<td>- Ortho: 100.8 GB in 106.5 min ➔ 16.2 MB/s</td>
<td>- Ortho: ~420 MB/s</td>
</tr>
<tr>
<td></td>
<td>- ATREM: 33.6 GB in 79 min ➔ 7.3 MB/s</td>
<td>- ATREM: 31.9 GB in 62 min ➔ 8.8 MB/s</td>
<td>- ATREM: ~228 MB/s</td>
</tr>
</tbody>
</table>
HyspIRI VSWIR daily uncompressed (at16 bits) data volumes:
- Minimum: 849 GB
- Mean: 1229 GB
- Maximum: 1436 GB

<table>
<thead>
<tr>
<th>Processing Algorithm</th>
<th>Minimum Daily Rate</th>
<th>Mean Daily Rate</th>
<th>Maximum Daily Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radiometric Calibration</td>
<td>53 min</td>
<td>76 min</td>
<td>88 min</td>
</tr>
<tr>
<td>Orthorectification</td>
<td>34 min</td>
<td>50 min</td>
<td>58 min</td>
</tr>
<tr>
<td>ATREM</td>
<td>64 min</td>
<td>92 min</td>
<td>108 min</td>
</tr>
<tr>
<td>Total</td>
<td>151 min/2.5 hrs</td>
<td>218 min/3.6 hrs</td>
<td>254 min/4.2 hrs</td>
</tr>
</tbody>
</table>
Summary and Conclusion

- HyspIRI VSWIR data rates and volumes are high by current standards and could represent a risk to implementation and cost estimates.

- The HyspIRI Mission concept team has been testing algorithms and implementation approaches to demonstrate the validity of the HyspIRI Mission concept. The on-board VSWIR FL compression algorithm is now being tested in the flight-like FPGA implementation.

- HyspIRI analog ground Level 1 and Level 2 data processing algorithms are being tested with AVIRIS-Classic and AVIRIS-Next Generation data sets. Scaling these benchmarking results shows that the HyspIRI mission concept is viable for on-board compression and ground processing of Level 1 and Level 2.

- There may be options for improved implementation margins and reduced costs in the HyspIRI mission concept in these areas.
HyspIRI Decadal Survey Mission

Key Science and Science Applications

**Climate**: Ecosystem biochemistry, condition & feedback; spectral albedo; carbon/dust on snow/Ice; biomass burning; evapotranspiration

**Ecosystems**: Global plant functional-type, physiological condition, and biochemistry including agricultural lands.

**Fires**: Fuel status, fire occurrence, severity, emissions, and patterns of recovery globally.

**Coral reef and coastal habitats**: Global composition and status.

**Volcanoes**: Eruptions, emissions, regional and global impact.

**Geology and resources**: Global distributions of surface mineral resources and improved understanding of geology and related hazards.

---

**Mission Urgency:**

The HyspIRI science and application objectives are important today and uniquely addressed by the combined imaging spectroscopy, thermal infrared measurements, and IPM direct broadcast.

---

**Mission Concept Status:**

Preliminary Draft Program Level 1 Requirements: Stable

**Payload**: Imaging Spectrometer, Thermal Infrared Imager, and IPM-Direct Broadcast subset

**Spacecraft**: Small

**Payload**: JPL/GSFC

**Launch Vehicle**: ~1000 kg class

**Launch date**: TBC (partner opportunities)

**Mission**: Class C 3-5 years

**Trajectory or Orbit**: LEO, Sun sync.

**S/C & Instrument Mass**: 561 kg (30% margin)

**S/C & Instrument Power**: 650W (66% margin)

The HyspIRI mission concept is mature and stable with excellent heritage, low risk and modest cost.

---

**Measurement:**

**Imaging Spectrometer (VSWIR)**
- 380 to 2500 nm in 10nm bands
- 60 m spatial sampling
- 19 days revisit
- Global land and shallow water

**Thermal Infrared (TIR)**
- 8 bands between 4-12 μm
- 60 m spatial sampling
- 5 days revisit
- Global land and shallow water

**IPM-Direct Broadcast**

---

**Ecosystems**

- Spruce/Fir
- White Pine
- Hemlock
- Beech
- Sugar Maple
- Red Maple
- Other Mixed HW

---

**Snow & Ice**

**Fires**

**Evapotranspiration**

**Volcanoes**

**Coastal Habitats**
Backup
### HyspIRI VSWIR Level 1 and 2 Algorithm Prototyping - Radiometric Calibration, Orthorectification, and ATREM Processing Speed

<table>
<thead>
<tr>
<th>Data Set Type</th>
<th>Windows Server 2008</th>
<th>Beowulf Cluster – Single Node</th>
<th>Beowulf Cluster – 26 Nodes</th>
</tr>
</thead>
</table>
| AVIRIS-Classic | - Rad Cal: 353 GB in 761 min ➔ 7.9 MB/s  
- Ortho: 35.7 GB in 63 min ➔ 9.6 MB/s  
- ATREM: 732.6 GB in 1375 min ➔ 9.1 MB/s | - Rad Cal: 3.6 GB in 11 min ➔ 5.6 MB/s  
- Ortho: Unable to test  
- ATREM: Unable to test | - Rad Cal: 90.1 GB in 11.5 min ➔ 133.7 MB/s  
- Ortho: ~390 MB/s |
| AVIRIS-Next Generation | - Rad Cal: 13 GB in 96 min ➔ 2.3 MB/s  
- Ortho: 31.9 GB in 185 min ➔ 2.9 MB/s  
- ATREM: 33.6 GB in 79 min ➔ 7.3 MB/s | - Rad Cal: 110.83 GB in 179 min ➔ 10.6 MB/s  
- Ortho: 100.8 GB in 106.5 min ➔ 16.2 MB/s  
- ATREM: 31.9 GB in 62 min ➔ 8.8 MB/s | - Rad Cal: ~276 MB/s  
- Ortho: ~420 MB/s  
- ATREM: ~228 MB/s |
Estimated Processing Time of Uncompressed HyspIRI VSWIR using 26-node Beowulf Cluster

<table>
<thead>
<tr>
<th>Processing Algorithm</th>
<th>Minimum Daily Rate</th>
<th>Mean Daily Rate</th>
<th>Maximum Daily Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radiometric Calibration</td>
<td>53 min</td>
<td>76 min</td>
<td>88 min</td>
</tr>
<tr>
<td>Orthorectification</td>
<td>34 min</td>
<td>50 min</td>
<td>58 min</td>
</tr>
<tr>
<td>ATREM</td>
<td>64 min</td>
<td>92 min</td>
<td>108 min</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>151 min/2.5 hrs</strong></td>
<td><strong>218 min/3.6 hrs</strong></td>
<td><strong>254 min/4.2 hrs</strong></td>
</tr>
</tbody>
</table>

HyspIRI VSWIR daily uncompressed (unconverted to 16 bit) data volumes:
- Minimum: 849 GB
- Mean: 1229 GB
- Maximum: 1436 GB