

# Evaluation of algorithms for compressing hyperspectral data

Sid Cook<sup>\*a</sup>, Joseph Harsanyi<sup>b</sup>, Vance Faber<sup>c</sup>

<sup>a</sup>Lockheed Martin Space Systems, Box 179, Mail Stop B6020, Denver CO 80201

<sup>b</sup>Applied Signal & Image Technology, 7763 Old Telegraph Rd Suite 7, Seven, MD 21144

<sup>c</sup>Mapping Science Inc., 33740 N.E. 84<sup>th</sup> Place, Carnation, WA 98014

## ABSTRACT

With EO-1 Hyperion in orbit NASA is showing their continued commitment to hyperspectral imaging (HSI). As HSI sensor technology continues to mature, the ever-increasing amounts of sensor data generated will result in a need for more cost effective communication and data handling systems. Lockheed Martin, with considerable experience in spacecraft design and developing special purpose onboard processors, has teamed with Applied Signal & Image Technology (ASIT), who has an extensive heritage in HSI spectral compression and Mapping Science (MSI) for JPEG 2000 spatial compression expertise, to develop a real-time and intelligent onboard processing (OBP) system to reduce HSI sensor downlink requirements. Our goal is to reduce the downlink requirement by a factor  $> 100$ , while retaining the necessary spectral and spatial fidelity of the sensor data needed to satisfy the many science, military, and intelligence goals of these systems. Our compression algorithms leverage commercial-off-the-shelf (COTS) spectral and spatial exploitation algorithms. We are currently in the process of evaluating these compression algorithms using statistical analysis and NASA scientists. We are also developing special purpose processors for executing these algorithms onboard a spacecraft.

Keywords: Hyperspectral, JPEG2000, Compression, Principal Component Analysis, Field Programmable Gate Arrays (FPGAs)

## 1. INTRODUCTION

Our objective is to reduce spacecraft cost and increase the accessibility and utility of HSI data through appropriate use of onboard processing. Our strategy is to combine the expertise from ASIT, a leader in the HSI processing field and the developer of a library of HSI directed data compression and target detection algorithms, with spatial compression expertise of MSI and the LM remote sensing, spacecraft and OBP capabilities. Figure 1 presents the HSI processing chain from sensor to finished HSI output products. Our OBP compression algorithms are designed to run unsupervised using only the data statistics to determine the compression transformation. The calibrated sensor data are input to our directed data compression processing flow and compressed.

The compressed data are then transmitted to the ground where they are uncompressed, corrected to account for sensor calibration results thus converting raw to radiance data, and atmospheric absorption effects are removed from the spectra thus resulting in reflectance data. Material classification and identification can then be performed using standard techniques. This information can then be used to generate various image-based products; e.g., classification and abundance maps.

A top-level description of our directed data compression is presented in Figure 2. In step 1 dominant pixels are separated from anomalous pixels using the tunable anomaly detector. This initial segmentation is important since dominant and anomalous pixels will be compressed differently; hence the name "directed data compression". This separation occurs by first determining the second order statistics and estimating the number of eigenvectors associated with the dominant pixels. The pixels that do not belong to the dominant subspace are automatically thresholded and so those pixels that have

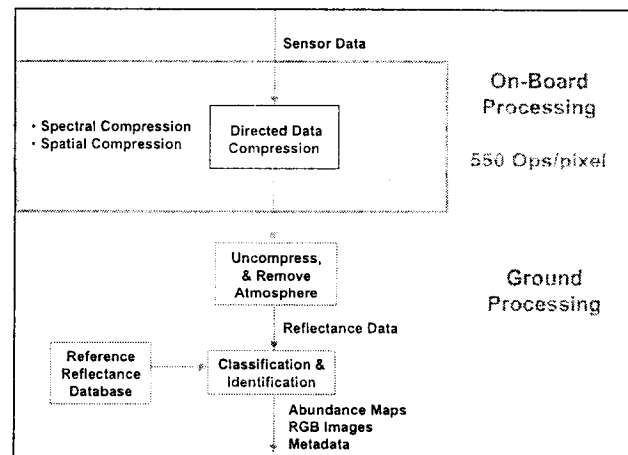


Figure 1: HSI Data Processing Chain

\* sidney.cook@lmco.com; phone 303-971-4609; fax 303-971-1627

inherently large amounts of information content are labeled anomalous. The anomalous pixels are passed along to the Spectral Uniqueness Monitor (SUM) where the subset of spectral signatures required to describe all anomalous pixel are determined. These spectral signatures are the anomalous subspace endmembers. The automatic subspace discrimination function is then used to assign each anomalous pixel to one of the SUM derived classes.

For Step 2, in Figure 2, the anomalous spectral basis signatures along with the dominant subspace eigenvectors are used together to formulate a compressing transformation in the spectral dimension. A spatial compression algorithm is then applied to further increase the overall compression ratio. Currently, the most promising algorithm uses JPEG2000 wavelet basis functions to provide the compression. Preliminary results show that additional 10X compression can be achieved spatially before degrading the data to the extent that science information is dramatically impacted. Thus with 10X compression for both spectrally and spatially a 100:1 compression results.

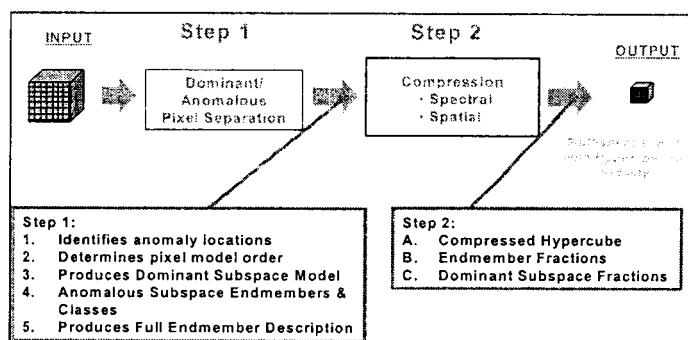


Figure 2: Directed Data Compression Approach

## 2. EVALUATION OF HSI DIRECTED DATA COMPRESSION RESULTS

We have used two methods for evaluating our compression results, analytical algorithms and scientists evaluations. The algorithms used for analytically evaluating the compression results are presented in Figure 3.

Figures 4 through 9 present the compression results of the Cuprite, Coleamberrally and Greater Victoris Watershed (GVWD) scenes acquired from the EO-1 satellite using the 242-band Hyperion sensor. Since 44 bands of the satellite did not function properly only 198 bands were calibrated and used for analysis. Even though little or no visual differences can be observed before an after compression there are some spectral differences. These spectral differences can be observed by comparing the spectral angles (angles, measured in degrees, determined by computing the spectral dot product) for corresponding pixels in the original and compressed data cubes. Small angles imply small differences in spectral data between the original and compressed data cubes. These computations were performed for every pixel in the scenes as observed in Figures 4, 6 and 8.

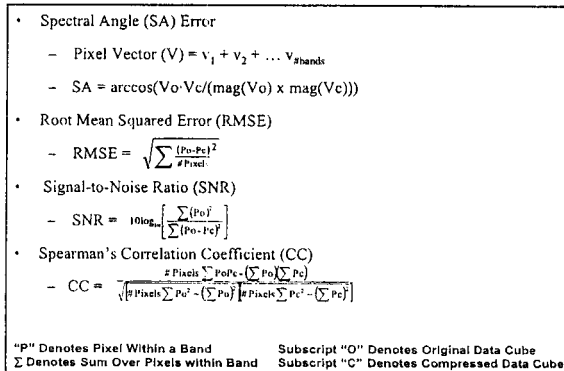


Figure 3: Compression Evaluation Algorithms

The root-mean-squared error (RMSE), signal-to-noise (SNR) and correlation coefficient (CC) analyses are presented in Figures 5, 7 and 9. The spikes shown in the RMSE curves appear to be caused by the vertical streaks that can be observed in the spectral angle error maps in Figures 5 and 7. These vertical strips are due to some of the sensor detectors malfunctioning in some of the sensor bands. For the GVWD scene most of the effects from the vertical streaks had been removed, thus the RMSE curves do not have spikes.

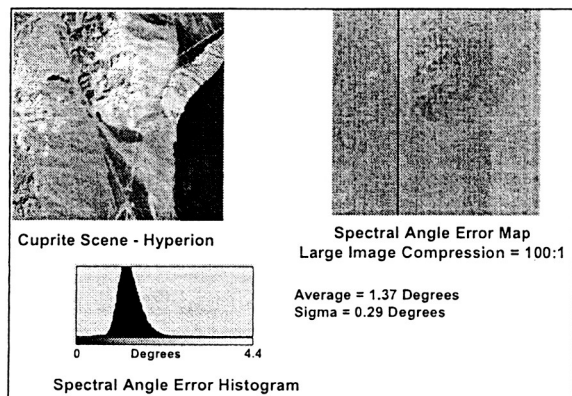


Figure 4: Spectral Angle Error – 198 Calibrated Bands

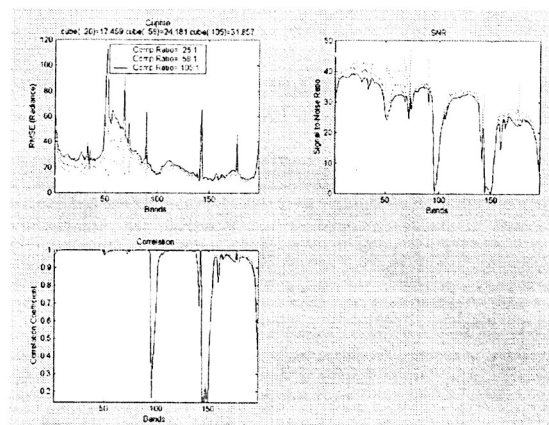


Figure 5: Cuprite Error Analysis

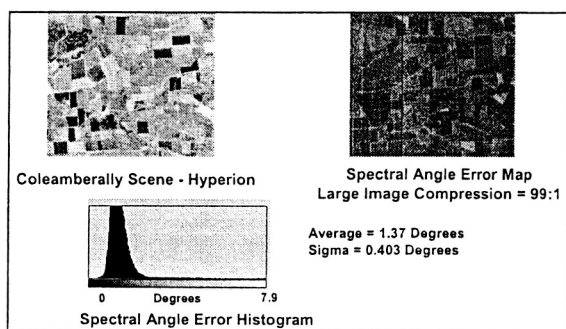


Figure 6: Spectral Angle Error – 198 Bands

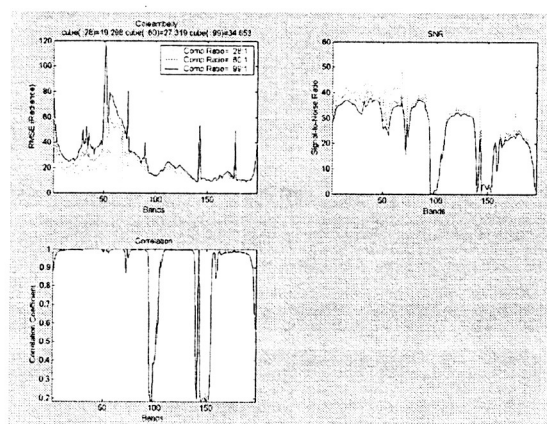


Figure 7: Coleambally Error Analysis

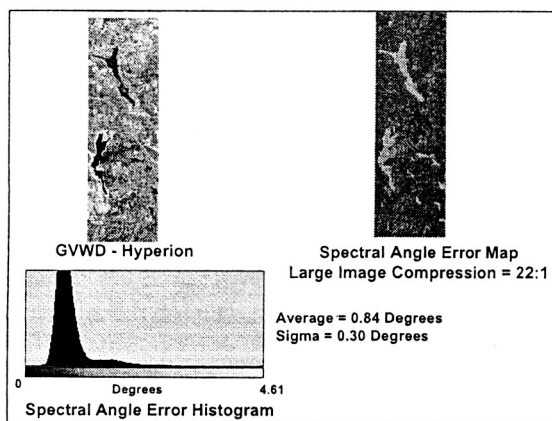


Figure 8: Spectral Angle Error – 198 Bands

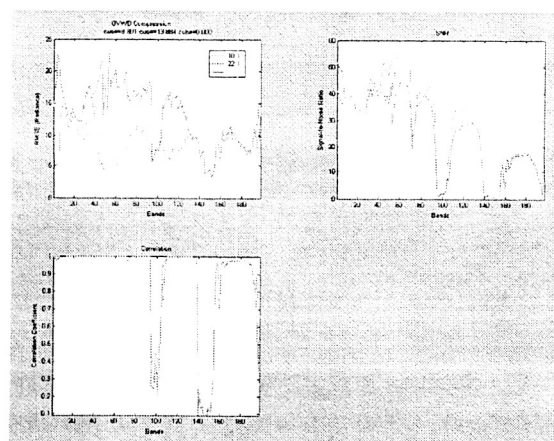


Figure 9: GVWD Error Analysis

In addition to analytical evaluations, it is also desirable to compare individual spectral features of compressed and uncompressed signatures. This type of comparison is being done in collaboration with EO-1 project scientists to ensure high value science information is not lost in the compression process. Results from two of the project scientists are presented in Figure 10.

The compression ratio does not take into account the size of the auxiliary data file because this file size is almost independent of the size of the data cube. Therefore for most operational strip maps the size of this auxiliary data file will be negligible in comparison to the compressed data cube.

### 3. ONBOARD PROCESSOR (OBP) REQUIREMENTS

Approximately 550 floating-point operations per pixel of input sensor data are required to implement the directed data compression processing flow specified in Figure 2. This information combined with specific sensor characteristics results in the OBP data processing requirements summarized in Figure 11.

For the EO-1 Hyperion sensor under the current operational environment the minimum requirement for an OBP to keep up with the data on a daily basis is 0.02 GFLOPS of equivalent computational power. Since our goal is to compress the data by a factor of 100:1 a Hyperion sensor could theoretically be tasked to increase its coverage by a factor of 100 without increasing the downlink requirements with the OBP processing capability of 2 GFLOPS equivalent, which we have chosen as the baseline requirements for the OBP system we plan to develop. If the Hyperion sensor is operated continuously then the OBP requirement becomes 6 GFLOPS equivalent. NASA Scientists at the May 2000 meeting of "Earth Science Enterprise Technology Planning Workshop on High Performance Spectrometry" specified a "future" HSI sensor with similar sensor characteristics to the Hyperion except with much improved noise characteristics and a swath increase from 7.5 km to 250 km. The corresponding OBP requirement to operate this future sensor continuously would be 160 GFLOPS. Analysis of the Directed Data Compression algorithms to be implemented by the OBP shows the single most computational intensive operation to be the matrix multiply operation. For a Hyperion type sensor this means operation with 256 pixel X 242 band sensor data matrices. A single generic matrix multiply operation (MMOp) on such data corresponds to approximately 16 million multiplies and adds, when reduced to 16-bit integer arithmetic. A 2 GFLOP equivalent OBP would be capable of performing approximately 120 such MMOPs/sec. For comparison, a desktop computer equipped with a 1.2 GHz Pentium™ processor can perform 10 MMOPs/sec. Our proposed OBP breadboard design for the Nov 2002 demo using 50 MHz FPGA based matrix multiply processing engines should be able to achieve at least 30 MMOPs/sec.

### 4. PROPOSED OBP ARCHITECTURE

In developing our proposed OBP architecture, several key points have been used as a guiding philosophy. First of all, the architecture should be modular in nature so as to be readily adaptable to future processing needs, both from a sensor type and overall throughput standpoint. A high degree of modularity also allows the maximization of parallel processing paradigms. Second, the concept should be highly FPGA based so as to maximize programmability and reconfigurability to accommodate changing mission computational requirements. Third, the use of COTS standards should be used where feasible to minimize development costs, however component selections should be done with caution so as to not preclude a clear path to a rad-tolerant implementations. Figure 12 illustrates such OBP architecture.

Comp Ratio	Evaluation Observations for the Colcambally Scene
99:1	Some Classification Error (<1%)
60:1	No Classification Error Some Minimum Noise Fraction (MNF) Analysis Error
28:1	No Observed Error
Analyzed by Dr. Jay Pearlman- EO-1 Science Team	

Comp Ratio	Evaluation Observations for the GVWD Scene
22:1	No Observed Error
Analyzed by Dr. David Goodenough - EO-1 Science Team	

Figure 10 EO-1 Scientists Evaluation Results

	Current Hyperion	OBP Enhanced Hyperion	Continuous / Real-time Hyperion	Future*
Area Coverage Per Day (KM <sup>2</sup> /day)	14.7 K	1.5 M	4.5 M	117 M
OBP Rate Requirement (GFLOPS)	0.02	2	6	160
Data Transmission Req. (Gigabits/day)	67	67	201	5360

<b>Sensor Characteristics:</b> • Pixel Spatial Resolution - 30m • # Spectral Bands - 242 • Sensor Data Generation Rate - 150 mbps	<b>Current Hyperion Operational Characteristics</b> • 6 Downlinks/day • 2 Image Swaths per Downlink • Image Swath Width - 7.65 KM • Image Swath Length - 185 KM
--	---

\* "Earth Science Enterprise Technology Planning Workshop on High Performance Spectrometry"  
May 12, 2000 (Hyperion Sensor with 250 KM Swath Width and SNR = 1000:1 VNIR and 500:1 NIR)

Figure 11 Onboard Processor Requirements

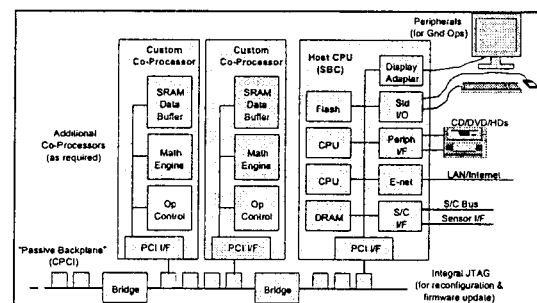


Figure 12 System Architecture

Figure 13 depicts the functionality of the board that has been built. This single co-processor card is FPGA based, and designed using Xilinx XCV1000 and XCV300 components that are currently available in a rad-tolerant version, so as not to preclude eventual design migration to an OBP. The PCI backplane interface, external data port control, and local data bus transfers are all implemented in FPGA logic to demonstrate the independent bus mastering, multiprocessor architecture operating in a Windows 2000 environment. The matrix multiply math function, along with necessary data sequencing, row and column data buffering, and overall instruction execution logic are also implemented in FPGA logic to demonstrate the performance advantages possible from dedicated processing hardware. An ample amount of high-speed SDRAM provides local storage of matrix data necessary to support any given calculation.

The matrix multiply co-processor card receives its instructions from the host, which consists of an "op-code" which defines the exact operation to be performed, the size of the data arrays, and pointers to the data locations in host memory. The co-processor card then initiates DMA operations to retrieve the operand data, execute the matrix multiply instruction, return the results to host memory, and generate an interrupt to the host processor signifying instruction completion. Although initially implemented in a conventional PC, this approach serves to demonstrate the advantages of parallel co-processors within the overall OBP architecture, and does not preclude a simple migration path to a real-time host operating system and fully flight qualified hardware in the future

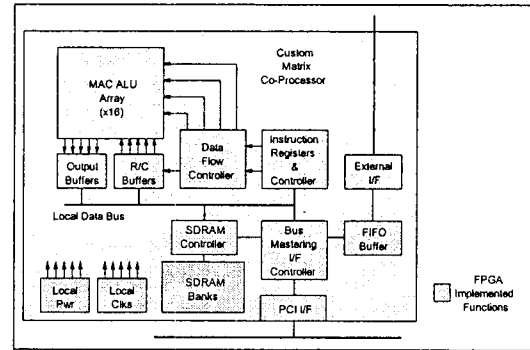


Figure 13: Matrix Multiply Co-Processor

## 5. CONCLUSIONS

We have now defined and demonstrated an approach to compress HSI data, and have shown that we can compress these data 100:1 with only a small amount of spectral degradation. When we compare the spectra of corresponding pixels between compressed and uncompressed data cubes we observe angular differences and other slight changes in spectral features. The significance of these differences depends upon the application. Currently NASA scientists are supporting us in evaluating the significance of the differences between the compressed and uncompressed data cubes, so that important science data are not lost. Two NASA scientists have concluded that at a 25:1 compression ratio they cannot observe any degradation.

We believe that digital signal processor (DSP) and field programmable gate array (FPGA) technology applied in a modular parallel processing architecture as shown in Figure 12 can satisfy both near-term and long-term flight-worthy OBP requirements. Other enabling technologies; e.g., application-specific integrated circuits (ASIC), optical processors, and high-density interconnect (HDI) packaging, are available to address future OBP requirements as volume and power constraints become more prohibitive.

## ACKNOWLEDGMENTS

This work is being performed under the NASA/ESTO/AIST NRA contract NAS5-00216.

## REFERENCES

- Harsanyi, J. C., (1993), *Detection and Classification of Subpixel Spectral Signatures in Hyperspectral Image Sequences*, Ph.D. Dissertation, University of Maryland, Baltimore County, 116 pp.
- Harsanyi, J. C. and C. Chang, "Hyperspectral Image Classification and Dimensionality Reduction: An Orthogonal Subspace Projection Approach," *IEEE Transactions on Geoscience and Remote Sensing*, July, 1994.
- Harsanyi, J. C. and C. Chang, "Hyperspectral Image Dimensionality Reduction and Pixel Classification: An Orthogonal Subspace Projection Approach," *Proceedings of the Twenty-seventh Annual Conference on Information Sciences and Systems*, pp. 401-406, Mar 24-26, 1993.
- Harsanyi, J. C., W. H. Farrand, C. Chang, "Detection of Subpixel Signatures in Hyperspectral Image Sequences," *1994 ASPRS/ACSM Annual Convention and Exposition Technical Papers, Reno, Nevada*, vol. 1, pp. 236-247, Apr. 25 -28, 1994