

HYPERSPECTRAL DATA PROCESSING: AN OPPORTUNITY FOR END-TO-END PROCESSING

Marge Cole, Dr. Anne Wilson, Michael Little

SGT, Inc., LASP, NASA Earth Science Technology Office

ABSTRACT

The evolution and improvements in hyperspectral instrumentation are being matched by information technology improvements in science data processing and analysis. Research has improved techniques in both onboard and ground-based processing to support other high data volume instruments. Algorithms and hardware have evolved, permitting faster access to the observations. Cloud computing is taking the algorithms to the data. Technologies are being specifically designed to address high volume data sets and are an investment in the improvement of hyperspectral data processing.

Index Terms—hyperspectral, on-board processing, science data processing, algorithms

1. INTRODUCTION

The need for hyperspectral data has grown over the past few years, to the point where the 2017 Earth Science Decadal Survey has re-emphasized it as a contributor to the fourth priority observable, Surface Biology and Geology. It is also called out to support the Planetary Boundary Layer observable. Recent hyperspectral technology advances make it critical to improve estimates of primary production, nutrient, and carbon cycling. [1] Considerable discussion of hyperspectral instruments has also been held in considering the measurements in the Sustained Land Imaging mission.

The successful application of modern computer science and information technology is essential to

translating hyperspectral instrument output into scientific understanding of the natural phenomena. While substantial work remains to apply the fundamental concepts to the science problems, important inroads have been made. NASA'S Earth Science Technology Office (ESTO) addresses requirements by developing a broad range of advanced technologies for scientific observation and measurement of Earth. [2] A survey of ESTO technological developments is useful to mission designers.

2. ON-BOARD PROCESSING

Onboard computing hardware has evolved considerably, including capabilities such as:

- Improved onboard computing hardware, such as Field Programmable Gate Arrays (FPGA), which provide near ASIC performance while being reprogrammable,
- Increasingly higher resolution instrumentation, e.g., increasing spectral resolution from tens to thousands of bins,
- Dynamic interaction with other elements, e.g., satellites within a constellation, models, and integrated sensor webs.

Algorithmic approaches to leveraging these developments generally involve methods for data analysis and reduction using technologies such as computer vision, image analysis/processing, and machine learning. Examples include:

- Improved compression techniques and fault tolerance,
- Scene analysis, determination of regions of interest, used for instrument aiming and/or transmission,

- Ranking of importance of observations,
- Integration into a constellation or even autonomous swarms.

Several recent ESTO Projects have addressed on board processing, including A High Performance Onboard Multicore Intelligent Payload Module, a demonstration of end-to-end operations to create and deliver data products in near real time. [3] This involved configuring a flight package for airborne vehicles applicable for HypsIRI and most Earth Science Decadal Survey missions.

Recent investments in CubeSats [4] is creating greater functionality as well. IPEX, a 2013 – 2015 CubeSat mission funded by NASA’s ESTO validated onboard instrument processing, mission planning, and autonomous payload operations to be used on future hyperspectral missions, including for HypsIRI’s Intelligent Payload Module (IPM). IPEX demonstrated several technologies for data reduction: computer vision, image analysis, and operations autonomy. [5]

As FPGAs in space are susceptible to radiation effects, the use of commodity FPGA devices requires improved fault mitigation techniques. Information Sciences Institute’s Radiation Hardening by Software (RHBSW) [6] techniques address this issue by embedding code in processors that leverage the dual processors to implement fault mitigation techniques used in high performance computing: checkpoint and restart, heartbeat monitoring, control flow assertions and watchdog times. These techniques are extremely low overhead, typically <1.3%, enabling a 3.3x gain in processing performance as compared to the equivalent traditionally radiation hardened processor. [6]

Additionally, cooperative, cross organizational efforts may provide a big lift in these endeavors, such as that between NASA, Boeing, University of Michigan and US Air Force Research Lab to develop a general purpose Arm system-on-chip High Performance Spaceflight Computer (HPSC)

and attendant peripherals (data buses, memory, software development environment, storage). [7]

3. TRANSMISSION

Multiple factors impact data transmission, including the bandwidth costs of housekeeping tasks and limited ground station reception capabilities. While the HypsIRI raw bit rate for the downlink is 800Mbps, the sustainable information rate will be significantly lower due to protocol overheads, error correction encoding, and the number of ground stations that can support the expensive dual X-band technology required to achieve those rates. With a data creation rate of 936Mbps, HypsIRI will require some amount of on-board compression in order to continually collect data over its entire orbit and successfully become a global mission.

While lossy compression is acceptable for some users, others can’t risk unrecoverable corruption or loss in compression, so lossless compression is important. New lossless compression algorithms are emerging. NASA JPL researchers have developed a domain specific Hyperspectral Fast Lossless (FL) compression algorithm for 16b single channel images that leverages knowledge of image depth to achieve similar compression rates as other lossless technique. [8] Adding fault tolerance with low overhead is critical to being able to realize the global mission real-time collection, compression, and downlink goals of HypsIRI.

4. GROUND BASED PROCESSING

Technologies such as cloud computing and machine learning are revolutionizing ground-based processing. The application of cloud technologies to Synthetic Aperture Radar (SAR) data greatly improved the ability to achieve production quality processing and analysis while keeping up with the information stream.

Machine learning processes have also been applied to SAR data improving quality control and work flow. By identifying a variety of errors and

alternative corrective processing streams, automated detection, assessment and pipeline processing have done much to reduce both the cost and clock-time of processing. The cost of this high volume computing, in the case of SAR, has been demonstrated to be reduced by a factor of 100 to 1000. Another major breakthrough is the turn-around time on the processing; the use of elastic capacity on cloud computing has translated months of processing into days or hours.

Hyperspectral instrument processing needs similar capabilities. The fundamental concepts and even much of the SAR code can be applied to hyperspectral data processing, so it is reasonable to assume that similar savings can be reaped in that arena.

Finally, once the data products have been created and stored, the ability to analyze them requires modern techniques around data modeling, machine learning, uncertainty quantification and more. NASA's Advanced Information Systems Technology (AIST) Program, under ESTO, has invested in several projects that are currently developing technologies involving:

- Learning methods for classification, learning frameworks
- Data modeling
- The integration of hyperspectral data into models
- Construction of themed databases, such as for vegetation

A current barrier is the volume of machine training data needed, especially when combining multiple types of features such as spatial and spectral. An ESTO project, Advanced Active Learning Methods for Robust Classification of Multi-Source Remotely Sensed Data [9] has advanced development in this arena by combining lidar and hyperspectral measurements to identify tree types. It employs an 'active learning' framework that utilizes components to 1) learn metrics, 2) regularize, and 3) project into a common feature space.

Integration of hyperspectral data into other data analyses tools or models is a common problem. E.g., the AIST funded project, Generative Models to Forecast Community Reorganization with Climate Change, integrates biological and remote sensing data to create decision-ready products. [10] This project integrates remote sensing variables with ecological data to provide accessible ecological forecasts translated directly to decision-ready products for the user community and managers.

Researchers using hyperspectral data need help using the massive volumes that these instruments produce. Hylatis is an AIST project from the University of Colorado's Laboratory for Atmospheric and Space Physics (LASP) [11] developing a suite of cloud based analysis tools for hyperspectral datasets. Leveraging the Spark big data engine to provide hyperspectral specific operations including subsetting on time, geolocation, and/or wavelength. A distinctive feature of Hylatis is its use of a domain agnostic, mathematical data model. Leveraging the rigor of functional programming, Hylatis will provide a domain neutral, mathematical platform that can support the integration and fusing of datasets from disparate domains.

Today many Earth science domains are creating unified databases for domain relevant data collected over the years. AIST's EcoSIS [12] is building an online, open source database for vegetation spectra, including the establishment of data collection standards and practices, and metadata standards. The ultimate goal is to facilitate question-based ecological and remote sensing research, as well as providing a foundation for community-driven collaboration.

4. FUTURE WORK

While many advances are on-going considerable work is still required to apply modern computer science and information technology to hyperspectral data science. The National Academy

of Sciences Decadal Survey – Thriving on Our Changing Planet, identifies as some of its Most Important goals Ecosystem Change, Reducing Climate Uncertainty and Informing Societal Response, and Surface Dynamics, Geological Hazards, and Disasters, for greater understanding. [1] These are all areas in which hyperspectral remote sensing will be a key provider of data. Additionally, while many advances are being made in gathering, transmitting, and analyzing these data the near future will bring greater volumes of this data and a substantial need for improvement in all technologies related to hyperspectral data administration. The AIST Program continues to assemble technology needs within the hyperspectral remote sensing domain and these requirements still stretch across the topics covered in this paper. On-board processing, transmission, and ground-based processing, including data fusion and analytics advancements will all benefit from continued innovative technology development.

5. REFERENCES

- [1] Thriving on Our Changing Planet: A Decadal Strategy for Earth Observation from Space, <https://www.nap.edu/catalog/24938/thriving-on-our-changing-planet-a-decadal-strategy-for-earth>.
- [2] NASA Earth Science Technology Office, <https://esto.nasa.gov/>.
- [3] D. Mandl, A High Performance Onboard Multicore Intelligent Payload Module, https://sed.gsfc.nasa.gov/tech_products/581/high_perf_onboard/.
- [4] C. Norton, S. Chien, P. Pingree, D. Rider, J. Bellardo, J. Cutler, M. Pasciuto, NASA investment in CubeSats, Maturation, <https://digitalcommons.usu.edu/cgi/viewcontent.cgi?referer=https://www.google.com/&httpsredir=1&article=2987&context=smallsat>.
- [5] J. Doubleday, S. Chien, C. Norton, Autonomy for remote sensing – Experiences from the IPEX CubeSat, IEEE Xplore, Nov 12, 2015, [10.1109/IGARSS.2015.7327033](https://doi.org/10.1109/IGARSS.2015.7327033).
- [6] A. Schmidt, M. French, T. Flatley, IEEE Xplore, June 08, 2017, Radiation hardening by software techniques on FPGAs: Flight experiment evaluation and results, <https://ieeexplore.ieee.org/abstract/document/7943651/>.
- [7] NASA, Boeing, US Air Force Research Lab, University of Michigan developing a general purpose High Performance Spaceflight Computer (HPSC), <https://news.engin.umich.edu/2018/03/new-computing-system-to-enable-deep-space-missions/>.
- [8] Fast and adaptive lossless on-board hyperspectral data compression system for space applications, Aranki, N., et al, <https://ieeexplore.ieee.org/document/4839534/>.
- [9] H. Yank, M. Crawford, IEEE Xplore, January 27, 2014, Learning a joint manifold with global-local preservation for multi-temporal hyperspectral image classification, <https://ieeexplore.ieee.org/document/6721343/>.
- [10] J. Swenson, Generative Models to Forecast Community Reorganization with Climate Change, https://cce.nasa.gov/cgi-bin/bio/pi_list.pl?project_group_id=3762.
- [11] A. Wilson, D. Lindholm, O. Coddington, P. Pilewski, T. Baltzer, C. Pankratz, A Functional Approach to Hyperspectral Image Analysis in the Cloud, https://figshare.com/articles/A_Functional_Approach_to_Hyperspectral_Image_Analysis_in_the_Cloud_Presented_at_the_Winter_2018_ESIP_Meeting/5782071.
- [12] Spectral Data Discovery, Access and Analysis through EcoSIS Toolkits, Townsend, P., https://hyspiri.jpl.nasa.gov/downloads/2013_Symposium/day2_pm/20130530_NASA-TE_EcoSIS-HyspIRI-sjh.pdf.