EFFICIENT SPECTRAL ENDMEMBER DETECTION ONBOARD THE EO-1 SPACECRAFT

Ben Bornstein[1], *David R. Thompson*[1], *Daniel Tran*[1], *Brian Bue*[2], Steve Chien[1], Rebecca Castaño[1]

[1] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA, 91109

[2] Rice University, Electrical and Computer Engineering, 6100 Main St., Houston, TX, 77006

ABSTRACT

Spaceflight and planetary exploration place severe constraints on the available bandwidth for downlinking large hyperspectral images. In addition, communications with spacecraft often occur intermittently, so mission-relevant hyperspectral data must wait for analysis on the ground before it can inform spacecraft activity planning. Onboard endmember detection can help alleviate these problems. It enables novelty detection and target identification for scheduling followup activities such as additional observation by narrow field of view instruments. Additionally, endmember analysis can facilitate data summary for downlink. This work describes a planned experiment of selective downlink by the EO-1 autonomous spacecraft. Here an efficient superpixel endmember detection algorithm keeps to the limited computational constraints of the flight processor. Tests suggest the procedure could enable significant improvements in downlink efficiency.

Index Terms— Hyperspectral Unmixing Applications, Endmember Detection, Spaceflight

1. INTRODUCTION

The Earth Observing One (EO-1) spacecraft was launched in November 2000 to a 700 kilometer altitude Earth orbit. EO-1 hosts the Hyperion Hyperspectral instrument, capable of resolving the Earth's surface to 30 meter resolution simultaneously in 220 distinct spectral bands ranging from 0.38 to 2.5 μm (near ultraviolet to short-wave infrared) [1]. The first hyperspectral sensor to operate from space, it captures image swaths measuring 7.5 kilometers (across orbital track) by approximately 100 kilometers (along orbital track). Hyperion imagery has been used in a wide range of applications including mineralogy, geology, forestry, agriculture, and environmental management through the classification of surface type and features. As of 2010, Hyperion has produced over 30,000 hyperspectral images.

In 2004, EO-1 began flying breakthrough autonomy software as its primary method of operations. This flight and ground software, called the *Autonomous Sciencecraft* [2] enables onboard processing of Hyperion instrument data to detect science events and features such as volcanism [3], cryospheric change [4], flooding [5], and sulfur springs [6, 7]. Detection of these phenomena onboard enables the spacecraft to respond autonomously via a range of options such as: (1) sending an alert; (2) downlinking science products or subimages; and/or (3) acquiring further imagery on a later overflight. These options are invaluable for spacecraft with a constrained communications budget that can only downlink a small fraction of their hyperspectral data. Autonomous detection enables the EO-1 spacecraft to prioritize key products for transmission, delivering data in a more timely fashion to ground teams. Higher-level derived products such as summary maps of thermal activity or ice coverage can reduce data volumes by orders of magnitude. Future missions such as HyspIRI will suffer even greater bandwidth constraints due to the larger volume of collected data; the EO-1 mission can validate these onboard analysis techniques for use in the future [8].

Current EO-1 data analysis classifies individual pixels into discrete types such as snow, water, ice, land, and clouds. Developing specific pixel-wise classifiers becomes evermore labor intensive as the variety and number of classes increases. Computational considerations also preclude many statistical analyses of the entire image dataset such as a general endmember search. We have developed a method to identify a large number of classes and enable scene-wide, onboard statistical analysis by decreasing the number of spectra used to describe each image. The two-stage procedure involves an initial segmentation of the dataset to several thousand superpixels, and a second endmember detection step that further reduces the spectra to a small fixed number of endmembers. The endmember list is intrinsically valuable as a downlink data product but also serves as a compressed representation of the scene diversity for applying more sophisticated and complex classification strategies. In the subsequent sections we describe this strategy in detail, discuss its adaptation to the EO-1 flight computer, and present initial results from a case study of Hyperion imagery of the mineralogically well-understood Cuprite, NV area. In the coming months, we expect the method presented here will be operational onboard EO-1.

2. SEGMENTATION AND ENDMEMBER DETECTION STRATEGY

This work utilizes the superpixel segmentation strategy of Thompson et al. [9, 10, 11], which fragments the image into small contiguous regions of constant spectral properties and then uses the mean spectrum of each segment in subsequent analysis. The *superpixel* regions range from approximately 10 to over 100 image pixels in size, and represent small physical features on the surface. This new representation has several benefits for onboard processing. First, it can potentially reveal subtle spectral signatures at or below the level of instrument noise, as long as those spectral features are present in multiple contiguous pixels. Additionally, it reduces the number of spectra for later processing by one or more orders of magnitude. This makes endmember detection in an entire image feasible for the EO-1 onboard processor.

We favor the Felzenszwalb graph segmentation approach [12] to generate superpixels. In brief, an agglomerative clustering method represents the image as an 8-connected network with edges weighted according to the spectral distance between neighboring pixels. A generic measure of spectral distance can be used [9], while a task-specific distance metric learned from training data can potentially

improve segmentation quality [11]. The segmentation begins with each pixel as a separate cluster. It iteratively joins neighboring segments whenever a merger criterion is satisfied. Specifically, the largest segment weight of the minimum tree spanning the two clusters cannot be significantly larger than the maximum segment of either component's individual spanning tree. This results in some small pure regions that are never merged; a final *clean up* pass remedies this by merging all remaining small clusters below a minimum size threshold.

After the initial segmentation, we employ the Sequential Maximum Angle Convex Cone (SMACC) strategy for endmember detection [13]. SMACC is an efficient algorithm similar to Gram Schmidt Orthogonalization. It chooses endmember spectra in sequence that greedily maximizes the angular projection onto an ever-growing convex cone of endmembers. In principle, any endmember detection strategy would suffice but SMACC has several advantages for our application. First, it is *deterministic* so that its results are easily reproducible. Second, it is a *sequential* method that can return variable numbers of endmembers in rank-ordered priority. This permits the designer to vary the size of the resulting summary data product that is returned by the spacecraft based on the expected complexity of the scene and the downlink budget. One primary disadvantage of SMACC is its susceptibility to noise. However, we find that the superpixel segmentation helps mitigate this and in this application the algorithm offers solid performance relative to competitors like NFINDR [14].

3. FLIGHT ADAPTATION

While the implementation of both Felzsenzwalb graph segmentation and SMACC endmember detection algorithms are straightforward, care was required to fit within the limited memory and processing constraints of the EO-1 onboard compute environment. The flight computer is equipped with a Mongoose-V 32-bit microprocessor clocked at 12 MHz. Due to a lack of floating-point hardware, all calculations involving floating-point (decimal) numbers must be implemented purely in software. Also, the processor cannot be completely dedicated to onboard data analysis and autonomy tasks, but instead must be shared with the spacecraft control and flight software. Onboard memory capacity constrains both specific data analysis tasks and the amount of Hyperion data that may be processed.

A maximum of 16 MB of spare memory is available to onboard data analysis applications. Further, onboard software is limited to both a spectral and spatial subset of the entire scene. Downlinked and post-processed Hyperion data products contain a rich set of 220 distinct spectral bands at full 256x4096 pixel spatial resolution. By contrast, onboard data analysis is limited to a subset of 12 bands and only a 256x1024 pixel spatial subportion of the complete image. With such a limited amount of spectral data, band selection is critically important. The bands used for endmember detection will be selected by remote commands based on the observation goals; key considerations are their ability to discriminate features of interest and their susceptibility to noise. A wide range of automated band selection strategies exist in literature [15, 16], and a full treatment of this topic is outside this paper's scope.

Felzenzwalb segmentation leverages a disjoint-set graph representation with path compression whenever a node is accessed. This amortizes the cost of parent node searches and makes common segment joins operations highly efficient. With this optimization in place, the dominant computation and storage cost for segmentation is in the initial graph construction. The distance between each pixel and its eight neighbors must be computed and stored. We implement several common distance metrics for use onboard, but for each one we eliminate inessential floating-point operations. For instance, when computing Euclidean distance, the final square root computation is omitted; for spectral angle distance, only the dot product projection is computed, the inverse cosine is dropped. Such optimizations are justified in that the relative distance between pixels does not change. For a 256x1024x12 band image, calculating edge weights requires roughly 36 million floating-point operations. The total storage requirements for both disjoint-set nodes and corresponding edge weights is approximately 15 MB.

Compared to graph segmentation, SMACC endmember detection imposes a relatively small computation and storage load. During each SMACC iteration, the scene spectral residuals are projected onto each new endmember direction, subject to numerical constraints that preserve physical interpretability. The spectral residuals are updated to hold the new "remainder." This dot product projection dominates the computational budget of SMACC endmember detection. However, since superpixel segmentation has already reduced thousands of individual multispectral pixels to hundreds, by keeping the number of endmembers identified in the tens, the total number of dot product projection operations is limited to a few thousand. Similarly, storage is only required for each 12-band mean multispectral superpixel, plus some additional space to track endmember contribution and projection residuals. For onboard Hyperion data processing, SMACC, after segmentation, requires 500 KB or less of total storage.

Finally two areas that are often initially overlooked when adapting new algorithms and software technologies to flight code are quality assurance and integration and test. Integration and test (I&T) brings together separately finalized and validated hardware and software components to verify both inter-component interfaces and end-to-end system operation. While we have not yet started the formal I&T effort, we have taken a number of steps to ensure delivery of a high quality and easily verifiable product. Our flight code is written in C, based on, but not strictly ported from, both IDL and MATLAB prototypes. The code is backed by nearly 100 automated unit tests which provide complete code coverage. A smaller set of automated regression tests verify end-to-end system functionality. Unit and regression tests taken together provide an easy way to verify correctness during flight integration and, should later code changes be required, our test suite will quickly expose unanticipated ripple effects. Our flight code complies with the JPL Coding Standard for C, has undergone regular Coverity Prevent (TM) static analysis and formal code peer reviews are currently in-progress.

4. CASE STUDY: CUPRITE, NV

To verify our endmember detection method, we requested a Hyperion image acquisition of the well-studied Cuprite, NV area. Our request was fulfilled on February 6, 2011 (day-of-year 37) and was downlinked a few days later. Recall that a full Hyperion image swath is 256x4096 pixels by 240 spectral bands, but onboard processing is limited to a 256x1024 by 12-band pixel region. To match onboard conditions, we restricted endmember detection to a 256x1024 spatial subset of the full scene (scan lines 1925–2948). We also restricted analysis to 12 hand-selected spectral bands covering the 2.1 to 2.4 μm wavelength range. Cuprite is an acid-sulfate hydrothermal system exhibiting well-exposed kaolinite, alunite, silica, and some calcite [17]. All four minerals have a sufficiently rich set of absorption features in the 2.1–2.4 μm range to be detected and also distinguished from one another. Conveniently, this wavelength range avoids noise from atmospheric water absorption. While we strove to

select an equal spacing of bands covering the major absorption features, we omitted particularly noisy channels, as assessed by visual inspection. Simple, light striping artifacts were not sufficient for a band to be removed from consideration. The 12 bands we chose are shown in Figure 2. Hyperion band performance trade studies are an area of future work.

Onboard data is represented as a digital number (DN) (raw sensor) value. We started with this representation and applied the EO-1 Hyperion onboard algorithm to convert DN values first to radiance and then to reflectance. Briefly, onboard conversion assumes a standard bidirectional reflectance model on Lambertian (diffuse) surfaces, the same method used by GOES-8 [18]. We applied both superpixel segmentation using a spectral angle distance measure and SMACC endmember detection as already described. Intermediate superpixel segments are depicted in Figure 1 (right). SMACC filtered segments, i.e. those superpixel regions which correspond to likely pure endmembers, are shown in Figure 1 (bottom). For this scene, superpixel segmentation reduced 262144 individual multispectral pixels to 3478 intermediate superpixel segments, a 75-fold data reduction. SMACC endmember determination offered a further 100-fold data reduction.

Of the 30 superpixel endmember regions, over half, and by far the largest regions, correspond to exposed mineralogy of the Cuprite mine area in the south (right half of the scene in Figure 1). Several of these regions align nicely with those highlighted in both Kruse et al. [17] and Thompson et al. [9]. By contrast, superpixel detections in the northern mountainous region are much smaller in total area. This is consistent with exposed mineralogy being easily detectable from both aerial and orbital spectral imagery. In both northern and southern areas, we are continuing to assess which regions cover pure endmembers.

5. DISCUSSION

Spacecraft platforms presents unique challenges in the form of constrained communications, computation, and high instrument noise. We address these problems with data reduction strategies such as superpixel segmentation and band selection. Together, these measures make classical endmember detection algorithms feasible for use onboard the EO-1 autonomous sciencecraft. The Cuprite experiments demonstrate recovery of the basic mineralogical diversity in the Cuprite scene, suggesting that endmember detection is a good candidate algorithm for onboard data analysis. The Cuprite scene is uniquely well-studied, with a wide variety of minerals and many challenging compact deposits. The test scene was well-illuminated and mostly cloudless, so it provides a natural upper bound on expected performance of the system. Further study on the ground and in flight will explore performance for a wider range of targets and imaging conditions. We will also consider alternative application domains involving agriculture, land use, and ocean science.

A future generation of hyperspectral imagers such as HyspIRI will return unprecidented volumes of image data; communications bandwidth constraints and latency will be a key constraint on the total data yield. A similar issue confronts planetary exploration missions that must communicate with Earth over the Deep Space Network. Endmember detection can improve mission science return in both cases. It permits more sophisticated analyses such as novelty detection, change detection against historical catalogs, scene summary, and data reduction prior to onboard classification to find specific targets of interest.

Acknowledgments: We thank the EO-1 team and for accommodating our observation request. The work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

6. REFERENCES

[1] S.G. Ungar, J.S. Pearlman, J.A. Mendenhall, and D. Reuter, "Overview of the Earth Observing One (EO-1) mission," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 6, pp. 1149–1159, 2003.

[2] S. Chien, R. Sherwood, D. Tran, B. Cichy, G. Rabideau, R. Castano, A.G. Davies, D. Mandl, S. Frye, B. Trout, S. Shulman, and D. Boyer, "Using autonomy flight software to improve science return on Earth Observing One," *Journal of Aerospace Computing, Information, and Communication*, April 2005.

[3] A.G. Davies, S. Chien, V. Baker, T. Doggett, J. Dohm, R. Greeley, F. Ip, R. Castano, B. Cichy, R. Lee, G. Rabideau, D. Tran, and R. Sherwood, "Monitoring active volcanism with the Autonomous Sciencecraft Experiment (ASE)," *Remote Sensing of Environment*, vol. 101, no. 4, pp. 427–446, 2006.

[4] T. Doggett, R. Greeley, A.G. Davies, S. Chien, B. Cichy, R. Castano, K. Williams, V. Baker, J. Dohm, and F. Ip, "Autonomous on-board detection of cryospheric change," *Remote Sensing of Environment*, vol. 101, no. 4, pp. 447–462, 2006.

[5] F. Ip, J. Dohm, V. Baker, T. Doggett, A.G. Davies, R. Castano, S. Chien, B. Cichy, R. Greeley, and R. Sherwood, "Development and testing of the Autonomous Spacecraft Experiment (ASE) floodwater classifiers: real-time smart reconnaissance of transient flooding," *Remote Sensing of Environment*, vol. 101, no. 4, pp. 463–481, 2006.

[6] L. Mandrake, K. Wagstaff, D. Gleeson, U. Rebbapragada, D. Tran, R. Castano, S. Chien, and R. Pappalardo, "Onboard SVM analysis of Hyperion data to detect sulfur deposits in arctic regions," *Proceedings of the IEEE WHISPERS Workshop*, August 2009.

[7] D.F. Gleeson, R. Pappalardo, S. Grasby, M. Anderson, B. Beauchamp, R. Castano, S. Chien, T. Doggett, L. Mandrake, and K. Wagstaff, "Characterization of a sulfur-rich arctic spring site and field analog to Europa using hyperspectral data," *Remote Sensing of Environment*, 2010.

[8] S. Chien, D. Silverman, A. G. Davies, and D. Mandl, "Onboard science processing concepts for the HyspIRI mission," *IEEE Intelligent Systems*, vol. 24, no. 6, pp. 12–19, 2009.

[9] D.R. Thompson, L. Mandrake, M.S. Gilmore, and R. Castaño, "Superpixel endmember detection," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 48, no. 11, pp. 4023–4033, 2010.

[10] M. Bunte, D. R. Thompson, R. Castano, S. Chien, and R. Greeley, "Enabling europa science through onboard feature detection in hyperspectral images," *LPSC*, March 2011.

[11] B. Bue, D.R. Thompson, M. Gilmore, and R. Castano, "Metric Learning for Hyperspectral Segmentation," in *Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*. IEEE, 2011 (submitted).

[12] P.F. Felzenszwalb and D.P. Huttenlocher, "Efficient graph-based image segmentation," *International Journal of Computer Vision*, vol. 59, no. 2, pp. 167–181, 2004.

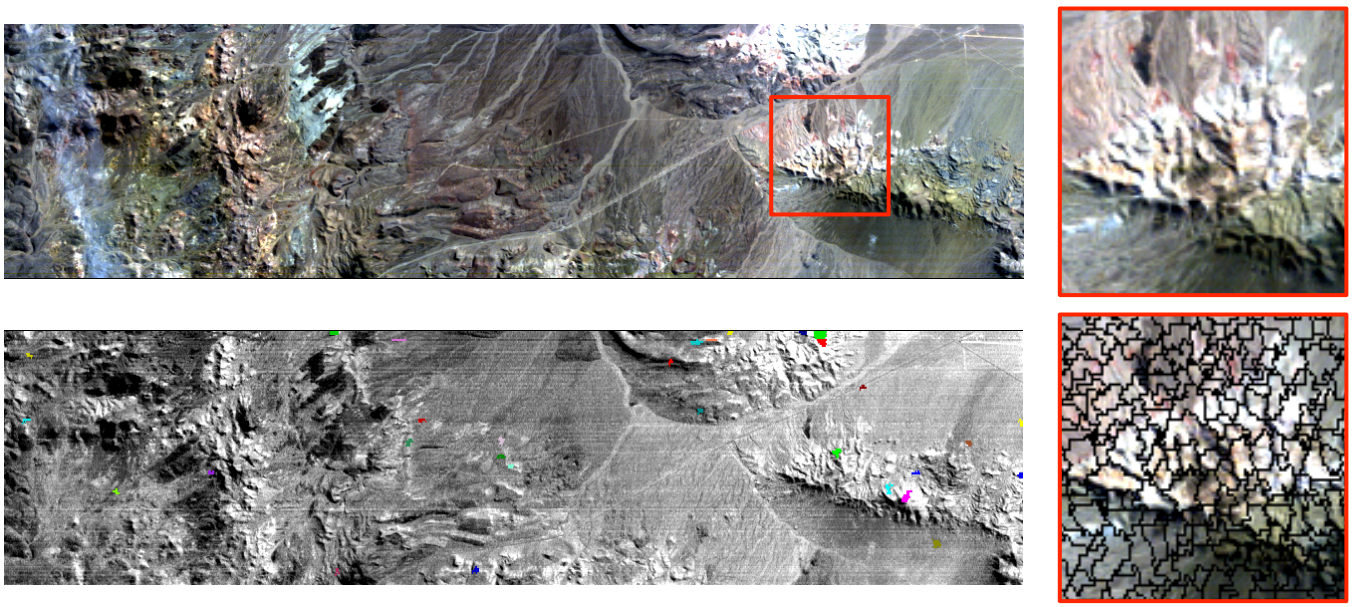

Fig. 1. (Top) Hyperion image of Cuprite, NV 256x1024 pixels (rotated) at 30 m/pixel (R: 640.50 nm, G: 548.92 nm, B: 457.34 nm). (Right) Detail region with and without intermediate superpixel segment boundaries superimposed. The entire scene was segmented into 3478 superpixel regions, a 75-fold reduction in data size compared to the original image. (Bottom) The top 30 SMACC filtered endmember segments, overlaid on Hyperion band 196 (2113.04 nm); some regions in the Cuprite mine area (right half of scene) correspond to endmember regions highlighted in both Kruse et al. [17] and Thompson et al. [9]. Further investigation is underway to identify which regions correspond to pure endmembers.

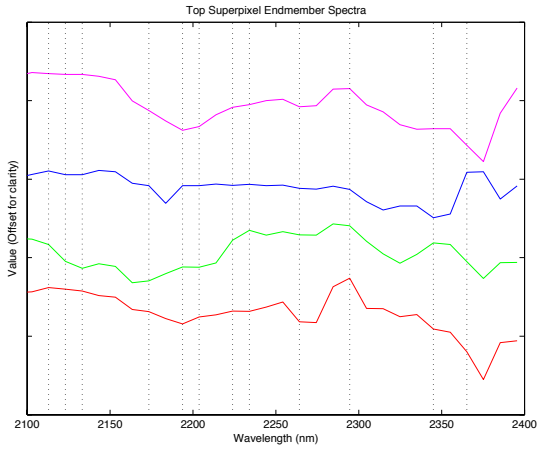

Fig. 2. Top four mean superpixel endmember spectra. For clarity all 30 Hyperion bands in the wavelength range 2.1–2.4 μm are shown. Spectra were ratioed using a flat spectral superpixel region. Dashed vertical lines denote the 12 bands used for superpixel segmentation and SMACC endmember determination.

[13] J. Gruninger, A.J. Ratkowski, and M.L. Hoke, "The sequential maximum angle convex cone (SMACC) endmember model," *Proceedings SPIE, algorithms for multispectral and hyperspectral and ultraspectral imagery*, vol. 5425, pp. 1–14, 2004.

[14] M.E. Winter, "N-FINDR: an algorithm for fast autonomous spectral end-member determination in hyperspectral data," in *Proceedings of SPIE*, 1999, vol. 3753, p. 266.

[15] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," *Machine Learning*, vol. 46, pp. 389–422, 2002.

[16] B. Moghaddam, Y. Weiss, and S. Avidan, "Generalized Spectral Bounds for Sparse LDA," in *International Conference on Machine Learning (ICML)*, 2006.

[17] F. A. Kruse, J. W. Boardman, and J. F. Huntington, "Comparison of airborne hyperspectral data and EO-1 Hyperion for mineral mapping," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, pp. 1388–1400, June 2003.

[18] K.R. Knapp, "Radiance to Reflectance for GOES-8 Channel 1," Tech. Rep., Colorado State University, August.