

Onboard Detection of Active Canadian Sulfur Springs: A Europa Analogue

Rebecca Castano⁽¹⁾
Kiri Wagstaff⁽¹⁾
Damhnait Gleeson⁽²⁾
Robert Pappalardo⁽¹⁾

Steve Chien⁽¹⁾
Daniel Tran⁽¹⁾
Lucas Scharenbroich⁽¹⁾
Baback Moghaddam⁽¹⁾

Benyang Tang⁽¹⁾
Brian Bue⁽¹⁾
Thomas Doggett⁽³⁾
Dan Mandl⁽⁴⁾
Stuart Frye⁽⁵⁾

⁽¹⁾ Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA, 91109
<firstname.lastname>@jpl.nasa.gov
v

⁽²⁾ Laboratory for Atmospheric and
Space Physics and
NASA Astrobiology Institute
University of Colorado
Boulder, CO 80309-0392
<firstname.lastname>@colorado.edu

⁽³⁾ Dept of Geological Sciences
Arizona State University
Box 871404
Tempe, AZ 85287-1404
<firstname.lastname>@asu.edu

⁽⁴⁾ Goddard Space Flight Center
8800 Greenbelt Road
Greenbelt, MD, 20771
Dan.mandl@gsfc.nasa.gov

⁽⁵⁾ Noblis
3150 Fairview Park Dr
Falls Church, VA, 22042
Stuart.frye@gsfc.nasa.gov

Abstract

We discuss a current, ongoing demonstration of in-situ onboard detection in which the Earth Observing-1 spacecraft detects surface sulfur deposits that originate from underlying springs by distinguishing the sulfur from the ice-rich glacial background, a good analogue for the European surface.

In this paper, we describe the process of developing the onboard classifier for detecting the presence of sulfur in a hyperspectral scene, including the use of a training/testing set that is not exhaustively labeled, i.e. not all true positives are marked, and the selection of 12, out of 242, Hyperion instrument wavelength bands to use in the onboard detector.

This study aims to demonstrate the potential for future missions to capture short-lived science events, make decisions onboard, identify high priority data for downlink and perform onboard change detection. In the future, such capability could help maximize the science return of downlink bandwidth-limited missions, addressing a significant constraint in all deep-space missions.

1. Introduction

The Jovian moon Europa represents an intriguing potential for life within our solar system. Yet there are considerable challenges facing any mission to find such habitats. One such challenge is that of efficiently searching for evidence of life given limited ability to return data roundtrip light time delays, and the anticipated scarcity and limited scale of indicators. We describe an active demonstration of onboard data analysis for detection of surface sulfur deposits that originate from underlying springs. Classifiers on the Earth Observing-1 (EO-1) spacecraft distinguish sulfur from the ice-rich glacial background. This demonstration serves as a key component in validating that such signatures can be detected from orbit and that onboard decision-making capability could be used to increase the efficiency of such a search.

EO-1's existing Autonomous Sciencecraft [8] onboard decision-making capability enables the spacecraft to perform onboard analysis of data from

the Hyperion hyperspectral instrument and use positive or negative detections to take further actions (such as repeat imagery). Closing the control loop onboard greatly increases the efficiency of search and monitoring because data lacking events/signatures of interest need not be downlinked. This demonstration aims to validate the claim that such signatures can be detected from orbit and that onboard decision-making capability can be used to increase the efficiency of such a search by remote spacecraft.

Building on previous work that produced a snow, water, ice, and land (SWIL) pixel classifier [9,12], we have developed a Support Vector Machine (SVM) to automatically detect the presence of sulfur in a hyperspectral scene. We describe the process of developing this detector, including the use of a training/testing set that is not exhaustively labeled, i.e. not all true positives are marked. Restricted onboard capability limits any processing to 12 of the total 242 Hyperion instrument wavelength bands. A key design choice for a classifier is which 12 bands to use. We empirically compare performance when using manually selected bands to that obtained when using bands chosen by automated feature selection techniques.

Historically, spacecraft collect data and transmit it to Earth for analysis. There are two primary motivations for analyzing science data onboard a spacecraft. The first is that onboard analysis can enable prioritization of the data by identifying the highest priority data for transmission. For many spacecraft, in particular those in deep space beyond Earth orbit, downlink bandwidth is severely limited. It is common that instruments are capable of collecting considerably more data than can be transmitted to Earth, however they are usually tasked to collect only as much data as can be downlinked. By collecting data at the capacity of the instrument and analyzing it onboard, there is significantly increased opportunity to identify rare features or events of interest. The second motivation for onboard science data analysis is to enable the detection of and reaction to dynamic events. For example, with an eight-hour round trip light time, timely reaction to an event on Europa would be possible only if the event were detected onboard and the spacecraft equipped to react.

There are a number of challenges to analyzing science data onboard a spacecraft. First, processing power is very limited. The radiation hardened CPU's used in spacecraft lags desktop technology by several generations. Memory is also very limited. Second, onboard analysis needs to be performed in a restricted timeframe. The actual limitations depend on the application such as when downlink opportunities occur

or the potential reaction time of the spacecraft based on orbit geometries. The third challenge in developing onboard science data analysis algorithms is that the data available onboard the spacecraft is un-calibrated. In some cases, such as the domain described in this paper, it is only possible to access onboard a limited portion of the collected data.

2. Europa and Borup Fiord Pass, Canada

In the Borup Fiord Pass on Ellesmere Island in the Canadian High Arctic, sulfur-rich waters seep from the top of a 200-meter-thick glacier precipitating deposits of sulfur, gypsum and calcite that stain the ice. The physical and chemical conditions of the spring water and surrounding environment, together with mineralogical and isotopic signatures, suggest that micro-organisms are active in the system [1]. This site and nearby regions may provide significant information about supraglacial sulfur springs and potential associated biological activity, i.e. signs of life beneath the ice. The site is considered an analog to Europa where ruddy dark surface markings (Figure 1) are thought to contain sulfur-rich materials [2] that may reflect the chemistry of a subsurface ocean, and possibly of organic materials carried upward [3].

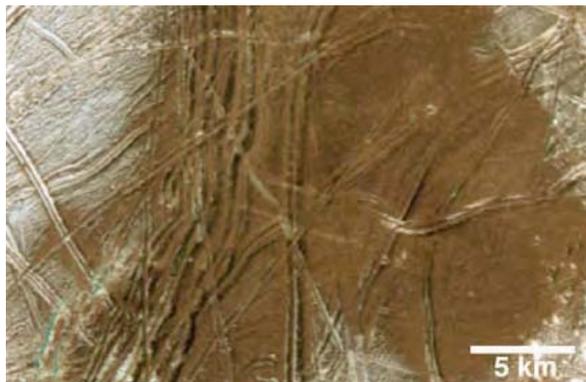


Figure 1. Ruddy regions on Europa as imaged by the Galileo spacecraft (from [3]).

Jupiter's moon Europa is an extremely high priority for spacecraft exploration because its probable subsurface ocean represents a possible habitat for microbial life. A variety of Europa mission studies have been conducted [4, 5, 6] and it remains a leading candidate for NASA's next Outer Planets flagship mission.

The spring system at Borup Fiord could potentially be analogous to Europa in a number of ways. The sulfur-rich chemistry of the supraglacial deposits may parallel that of Europa's ruddy spots. The glacial hydrology of the system operating at Borup may offer insights to the "plumbing" system at Europa. Borup's

microbiological environment could lend insights to possible microbiologic niches at Europa [3]. Thus, study of the Borup Fiord Pass site could be critical in the search for indicators of life beneath the ice of Europa.

Understanding how the complex system at Borup Fiord operates requires investigation of the geological, hydrological, geochemical, and microbiological properties of the site. One of the methods to conduct these investigations is through remote sensing. Remote sensing can provide details on the location and identification of the precipitates present on the ice in addition to aiding in mapping the regional geology and studying the temporal coverage of surface spring activity. Hyperspectral imagery from the Hyperion instrument onboard the EO-1 spacecraft provides high-resolution spectral data in narrow swaths along the area of interest. Hyperion and ASTER [7] coverage can be linked to spectral field observations and geochemical measurements that serve as ground truth.

3. Autonomous Sciencecraft

The Autonomous Sciencecraft (ASE) is a JPL-led, NASA New Millennium Program mission containing new technology in the form of software which has been flying on the Earth Observer-1 (EO-1) satellite since the fall of 2003 [8]. This new technology facilitates autonomous science-driven capabilities. Among the ASE flight software is a set of onboard science algorithms designed for autonomous data processing, to identify observed science events [8, 9]. Using the output from these algorithms, ASE has the ability to autonomously modify the EO-1 observation plan, retargeting itself for a more in-depth observation of a scientific event in progress with current response times on the order of hours. Several onboard science algorithms are associated with ASE for detecting dynamic events. Events detected include volcanic activity [10], floods [11] and cryosphere events [12]. In this work, we have expanded the detected features to include the presence of sulfur on an ice background.

3.1 Spacecraft and Instrument

EO-1, managed by NASA's Goddard Space Flight Center and also part of the same New Millennium Program, is designed to validate new technologies for remote sensing. It was launched from Vandenberg Air Force Base on 21 November 2000 and placed in a sun-synchronous orbit with an altitude of 705 km and a 10:01 AM descending node. The EO-1 payload is comprised of three instruments: Hyperion, Advanced Land Imager (ALI) and the Linear Etalon Imaging

Spectral Array (LEISA) Atmospheric Corrector. ASE analyzes data from the Hyperion instrument onboard the spacecraft.

The Hyperion instrument [13] consists of two imaging spectrometers, covering the visible/near infrared (VNIR) and short-wave infrared (SWIR), respectively, which share a common telescope, producing hyperspectral images with a 30 m/pixel spatial resolution and 10 nm/band spectral resolution. Hyperion images are 7.5 km in width, with an along track length that depends on the duration of the data collect, but typically 60 km (8 seconds) or 90 km (12 seconds). Due to onboard memory and data transfer limitations, we analyze a 7.5 km by 15 km subset of the captured image when detecting sulfur signatures. The VNIR spectrometer has 50 calibrated bands, ranging from 0.43 to 0.93 μm , and the SWIR spectrometer has 148 calibrated bands, ranging from 0.91 to 2.4 μm . Onboard constraints permit access to only 12 of the bands of the Hyperion instrument, although these 12 are selectable from the full complement.

There are two identical processors onboard the EO-1 spacecraft, one for the primary spacecraft operations and the other for the payload. ASE uses the payload processor. It is a Mongoose V CPU with a processor speed of 8 MIPS and 256 MB of RAM. With this hardware constraint, the Hyperion data cannot be fully processed from Level 0 (raw) data to Level 1 (calibrated) data [14]. Instead the data are partially processed to an onboard product designated Level 0.5, using data from a dark calibration image collected within a few minutes of the actual image.

Features of Level 1 data processing [14] not performed in the onboard processing include smear and echo correction to the SWIR bands, as well as interpolation between pre- and post- dark calibration images before dark image subtraction. While both Level 0.5 and Level 1 data are identical in VNIR, they diverge in SWIR, where the lack of smear and echo correction in Level 0.5 gives higher values than in the fully processed data. Because Level 0.5 data are not fully calibrated, the radiance and reflectance values for SWIR bands calculated onboard the spacecraft can be considered as pseudo-radiance and pseudo-reflectance.

4. Detector Implementation

The sulfur detectors in this effort were developed using supervised learning methods. Supervised classification employs *a priori* knowledge of a site and the identity of surface cover materials of interest in a training image or image set. Training areas are used to

develop or train a classification algorithm to recognize land cover classes based on their spectral signatures. In this instance, a labeled training data set was provided by a domain expert who had first-hand knowledge of the site. A small set of pixels in a scene of the target location was manually labeled as ice, land, or sulfur by the domain expert. Linear Support Vector Machine (SVM) classifiers [15, 16] were then trained using this data. Since there are limited examples of the presence of sulfur and few scenes available, the training set was necessarily limited. Further, as the sparse sulfur pixels had to be located manually, it is possible that there could be (rare) unlabeled pixels that do truly contain sulfur. The objective of such a classifier is to correctly identify the class of new pixels, in this case the focus is on sulfur present or not. The design goal was to correctly classify all labeled pixels with highest consideration given to not missing any labeled sulfur pixels. The secondary consideration was to minimize false alarms, where unlabeled pixels were all considered negative (no sulfur).

4.1. Band (Feature) Selection

The Hyperion instrument has 242 spectral bands, however due to limited processing power only 12 bands can be accessed onboard. The selection of which 12 bands to use represented a key design decision. Identifying the optimal 12 bands is a non-trivial, NP-hard ‘242-choose-12’ subset selection problem with over six million trillion possible combinations. With this many possibilities, an exhaustive search would require over 100 million years of computation. This being infeasible, we considered three practical approaches to selecting bands. The first is the traditional manual selection by a domain expert. We also looked at two automated feature selection methods: recursive feature elimination (RFE) and Greedy Sparse Linear Discriminant Analysis (GSLDA).

4.1.1. Domain Expert. The foundational approach to selecting which bands or features to use was to employ the domain knowledge of an expert. The expert was skilled in spectroscopy as well as possessing an extensive knowledge of the field site and the target sulfur signatures. Based on experience with the site, the full spectra of candidate pixels were manually inspected along with representative spectra of non-target material. By studying the spectral features, key bands that could best be used to discriminate the signatures of interest from what is considered the background were identified.

4.1.2. Recursive Feature Elimination (RFE). The goal of Recursive Feature Elimination (RFE) [17] is to identify a small subset of highly discriminative features. The RFE algorithm begins with the full set of features and recursively removes the feature with the minimum variation in a cost function. The key to the algorithm is the insight that the weights multiplying the inputs of a given classifier can be used as feature ranking coefficients. Inputs that are weighted by the largest value have the greatest influence on the classification decision. Therefore, if the classifier performs well, those inputs with the largest weights correspond to the most informative features. The algorithm proceeds by training a classifier using the current available set of features. The feature with the smallest weight is then removed and a classifier is then trained on the reduced set of features. This process is repeated until the desired subset is achieved, in our case 12 features. Note that, although a classifier is trained at each step of algorithm, the method is relatively efficient in addressing the combinatorial problem of identifying the best subset of features by greedily eliminating features.

4.1.3. Greedy Sparse LDA (GSLDA). Greedy Sparse LDA (GSLDA) [18, 19] is a state-of-the-art feature selection technique that uses a bi-directional (forward/backward) greedy search algorithm to find feature subsets which jointly maximize the Fisher linear discriminant class-separability criterion for a binary classification problem. It formulates and solves this NP-hard combinatorial optimization problem as a sparse generalized eigenvalue decomposition. It does so by first estimating two covariance matrices from the raw data: the between-class scatter and the within-class scatter matrices. It then maximizes the corresponding Generalized Rayleigh Quotient (or principal eigenvalue) with a sparse generalized eigenvector whose cardinality matches the number of features desired.

5. Experimental Results

5.1. Band Selection Discussion

The bands identified by the three feature selection approaches are shown in Table 1. Inspection of the table shows a significant overlap in the actual bands selected. In particular, seven bands (bolded) were selected by all three methods. Interesting to note is that the two automated methods each selected two higher wavelength bands while the expert did not choose any in this wavelength region.

Table 1. Band Selection Results.

Expert	RFE	GSLD A	
Band	Band	Band	Wavelength (λ , nm)
8	8	8	426.8
9	9	9	437.0
10	10	10	447.2
11	11	11	457.3
12			467.5
14			487.9
16			508.2
18			528.6
	19		538.7
20	20	20	548.9
	21	21	559.1
22	22	22	569.3
	23	23	579.5
24	24	24	589.6
		26	610.0
28			630.3
	189		2042.5
		193	2082.7
		194	2092.8
	219		2345.1

We compared the performance of classifiers trained using each of these three subsets. As mentioned earlier, the set of labeled data is extremely limited. There were a total of 151 pixels labeled, of which 18 were sulfur and the remainder non-sulfur. Each of the three classifiers correctly classified 150 of the 151 pixels (99.34% correct). All three misclassified the same non-sulfur pixel. We also evaluated the full image containing primarily unlabeled pixels. Results are shown in Table 2.

Table 2. Classification of unlabeled pixels.

Classification	Manual	RFE	GSLDA
Sulfur	460	660	880
Non-sulfur	869533	869333	869113
Potential false alarms	0.053%	0.076%	0.10%

Finally, we mention the time to perform the band selection. Manual identification required an expert with years of spectral data analysis and extensive field experience to take several days to inspect the data and select the key bands. The RFE method took around 10 seconds to compute (Python) while the GSLDA took 1.23 seconds (Matlab). RFE computes an ordered ranking of features from least to most significant and

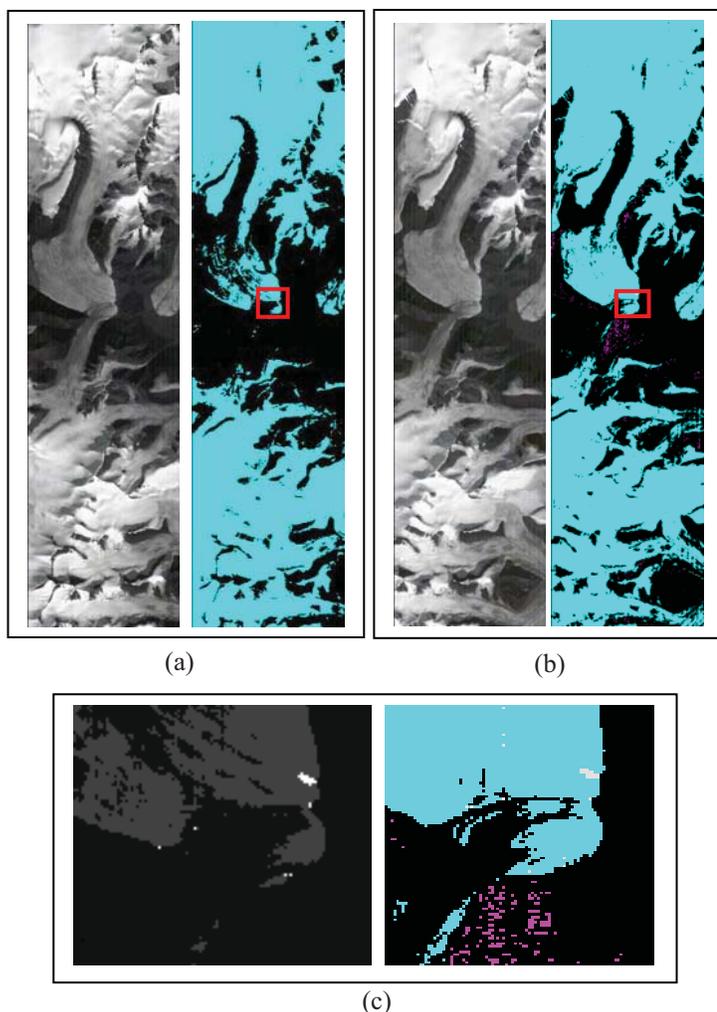


Figure 2. Results from onboard classifiers. For the classification images, black is land, cyan is water, white is sulfur, and magenta is unclassified. (a) Classifier using expert-selected bands. Image collected 7/24/2007. (b) Classifier using RFE-selected bands. Image collected 8/5/2007. (c) Close-in view of areas shown in red, which represent the field experiment site with known sulfur. As can be observed, the sulfur deposits were identified in both cases.

the GSLDA algorithm identified subsets of all sizes (not just 12 bands). Thus, both of the automated methods implicitly compute subsets of sizes other than twelve, should more or fewer bands be available. Both of the automated methods do, however require labeled data for training.

5.2. Data Analysis

In this section, we show a more extensive comparison between the expert-band SVM and the RFE-band SVM. Figure shows the number of pixels classified as sulfur in a set of 12 images for two different classifiers. Ten of the images are from 2007 while two are from the summer of 2006. Manual indicates the classifier was trained and run using data from the bands manually selected by an expert, while RFE indicates the classifier was trained and run using the bands selected automatically via the RFE algorithm. For these images, the manual-based classifier typically identifies more pixels as sulfur.

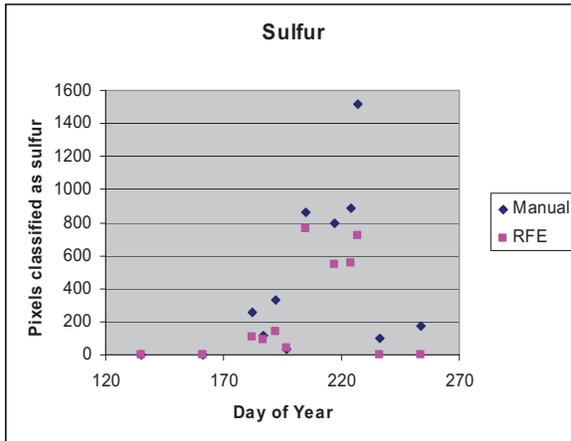


Figure 3. Comparison of number of pixels classified as sulfur on a set of 12 Hyperion images.

Table 3. Confusion matrix for classification of set of twelve test images. Rows indicate the class assigned by the manual-band based classifier while columns indicate the class assigned by the RFE-band based classifier.

		RFE		
		<i>Land</i>	<i>Ice</i>	<i>Sulfur</i>
Manual	<i>Land</i>	709557	111072	1139
	<i>Ice</i>	2349	2315921	6
	<i>Sulfur</i>	318	2843	1823

Table 4. Confusion matrix for classification of set of twelve test images given in percentage of total pixels. Rows indicate the class assigned by the manual-band based classifier while columns indicate the class assigned by the RFE-band based classifier.

		RFE		
		<i>Land</i>	<i>Ice</i>	<i>Sulfur</i>
Manual	<i>Land</i>	22.56	3.53	0.04
	<i>Ice</i>	0.07	73.62	0
	<i>Sulfur</i>	0.01	0.09	0.06

Tables 3 and 4 provide more insight into the consistency of pixel classification of the two classifiers. It can be seen from the tables that there is considerable agreement, although the RFE-based classifier deemed a number of pixels ice that the manual-band based classifier selected as land. Finally, for this section, we show the progression of the coverage over the course of the 2007 summer in Figure 4. A small subimage, corresponding to the subregion shown in Figure 2 has been focused on as it is the location of the sulfur deposits.

5.3. Onboard Experiments

Classifiers based on the domain expert selected bands and the RFE selection method have been uploaded and run onboard the EO-1 spacecraft. Initial results from these two classifiers run while there was still daylight at the site are shown in Figure 2. These results are significant in demonstrating not only the feasibility of identifying the signatures of interest in the hyperspectral data, but the practicality and effectiveness of performing the analysis onboard a spacecraft in the restricted environment available.

6. Conclusions

A broad objective of this study is to demonstrate the potential for future missions to capture short-lived science events, make decisions onboard, and identify high priority data for downlink. Ultimately, such capability can help maximize the science return of bandwidth-limited downlink channels, addressing a significant constraint in all deep-space missions.

7. Acknowledgments

This work was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. Funding was provided by the Interplanetary Network Directorate Technology Development program and NASA's New Millennium Program.

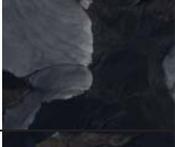
Date	Sub-Image	Manual	RFE
2007.05.15			
2007.06.10			
2007.07.01			
2007.07.11			
2007.07.24			
2007.08.05			
2007.08.12			
2007.08.15			
2007.08.24			
2007.09.11			

Figure 4. Progression of Surface coverage over a from May through September 2007.
 For the classification images, green is land, cyan is water, and yellow is sulfur.

8. References

- [1] Grasby, S.E., Allen, C.C., Longazo, T.G., Lisle, J.T., Griffin, D.W., and Beauchamp, B., 2003, Supraglacial sulfur springs and associated biological activity in the Canadian High Arctic--Signs of life beneath the ice: *Astrobiology*, v. 3, p. 593-596.
- [2] McCord, T.B. et al., *Science*, 280, 1242-1245, 1998.
- [3] Gleeson, D.F., Pappalardo, R.T., Grasby, S.E., and Spear, J.R., 2006, "Borup Fiord: A unique glacial environment of astrobiological significance and potential analogue to Europa exploration." 37th Lunar and Planetary Science Conference, p. 1854.
- [4] Johnson, T. V., K. B. Clark, R. Greeley and R. T. Pappalardo, "Europa Exploration: Challenges and Solutions, Lunar and Planetary Science Conference, March 2006, Abstract #1459.
- [5] Europa Explorer Mission Study: Final Report, JPL D-41283, November 1, 2007.
- [6] Europa Explorer Study report, (D-34054), May 2006.
- [7] Yamaguchi, Y., A. B. Kahle, H. Tsu, T. Kawakami and M. Pniel (1998), Overview of Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), *IEEE Trans. Geosci. Remote Sens.*, 36(4), 1062-1071.
- [8] Chien, S., Sherwood, R., Tran, D., Cichy, B., Rabideau, G., Castano, R., Davies, A., Mandl, D., Frye, S., Trout, B., Shulman, S., and Boyer, D., Using Autonomy Flight Software to Improve Science Return on Earth Observing One, *Journal of Aerospace Computing, Information, and Communication*, April 2005.
- [9] Castano, R., D. Mazzoni, N. Tang, R. Greeley, T. Doggett, B. Cichy, S. Chien, and A. Davies, "Onboard classifiers for science event detection on a remote sensing spacecraft," in *Proc. Of Twelfth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Philadelphia, PA, Aug 2006.
- [10] Davies, A. G., Chien, S., Baker, V., Doggett, T., Dohm, J., Greeley, R., Ip, F., Castaño, R., Cichy, B., Rabideau, G., Tran D., and Sherwood, R., Monitoring Active Volcanism with the Autonomous Spacecraft Experiment (ASE), *Rem. Sens. Environ.*, Vol. 101, Issue 4, pp 427-446, Apr. 2006.
- [11] Ip, F., Dohm, J.M., Baker, V.R., Doggett, T., Davies, A.G., Castaño, B., Chien, S., Cichy, B., Greeley, R., and Sherwood, R., Development and Testing of the Autonomous Spacecraft Experiment (ASE) floodwater classifiers: Real-time Smart Reconnaissance of Transient Flooding, *Rem. Sens. Environ.*, in press, 2006.
- [12] Doggett, T., Greeley, R., Chien, S., Castano, R., Cichy, B., Davies, A. G., Rabideau, G., Sherwood, R., Tran, D., Baker, V., Dohm J., and Ip, F., Autonomous Detection of Cryospheric Change with Hyperion on-board Earth Observing-1, *Rem. Sens. Environ.*, Vol. 101, Issue 4, pp. 447-462, Apr. 2006.
- [13] Pearlman, J.S., Barry, P.S., Segal, C.C., Shepanski, J., Beiso D., and Carman, S.L., "Hyperion, a Space-Based Imaging Spectrometer," *IEEE Trans. Geosci. Rem. Sens.*, 41, 2003, 1160-1172.
- [14] Barry, P. *EO-1 / Hyperion Science Data User's Guide, Level 1_B*, Redondo Beach, CA: TRW Space, Defense & Information Systems, 2001.
- [15] Boser, B., I. Guyon, and V. Vapnik. "A training algorithm for optimal margin classifiers." In Fifth Annual Workshop on Computational Learning Theory, pages 144--152, Pittsburgh, ACM. 1992.
- [16] Vapnik, V. N. *Statistical Learning Theory*. Wiley Interscience. 1998.
- [17] Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine Learning*, 46, 389--422.
- [18] Moghaddam, B., Y. Weiss, S. Avidan, "Generalized Spectral Bounds for Sparse LDA," *International Conference on Machine Learning (ICML)*, 2006.
- [19] Baback Moghaddam, B., Y. Weiss and S. Avidan. Fast Pixel/Part Selection with Sparse Eigenvectors. *IEEE International Conference on Computer Vision (ICCV)*, October 2007.