SMART CAMERAS FOR REMOTE SCIENCE SURVEY

David R. Thompson [(1)], William Abbey[(1)], Abigail Allwood[(1)], Dmitriy Bekker[(1)], Benjamin Bornstein[(1)]
Nathalie A. Cabrol[(2)], Rebecca Castaño[(1)], Tara Estlin[(1)], Thomas Fuchs[(3)], Kiri L. Wagstaff[(1)]

(1) Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr. Pasadena, CA 91109, USA, *firstname.lastname*@jpl.nasa.gov
(2) SETI Institute, 189 Bernardo Ave, Suite 100, Mountain View, CA 94043, USA. *nathalie.cabrol@jpl.nasa.gov*
(3) California Institute of Technology, 1200 E. California Blvd, MC 136-93, Pasadena, CA 91125, USA. *fuchs@caltech.edu*

ABSTRACT

Communication with remote exploration spacecraft is often intermittent and bandwidth is highly constrained. Future missions could use onboard science data understanding to prioritize downlink of critical features [1], draft summary maps of visited terrain [2], or identify targets of opportunity for followup measurements [3]. We describe a generic approach to classify geologic surfaces for autonomous science operations, suitable for parallelized implementations in FPGA hardware. We map these surfaces with *texture channels* - distinctive numerical signatures that differentiate properties such as roughness, pavement coatings, regolith characteristics, sedimentary fabrics and differential outcrop weathering. This work describes our basic image analysis approach and reports an initial performance evaluation using surface images from the Mars Exploration Rovers. Future work will incorporate these methods into camera hardware for real-time processing.

1. INTRODUCTION

Astrobiologists have developed diverse in situ instruments for identifying paleohabitats, detecting spectral and chemical biosignatures, and characterizing isotopic abundances of life-related elements. However, during operations these instruments are universally constrained by limited communications windows, blackouts and narrow bandwidth data transfer [4,5]. Spacecraft rely on sporadic communications with mission operators to interpret observations and select new sampling targets. This operations bottleneck impacts nearly every data collection activity and significantly reduces a mission's lifetime science return [4]. Surface missions consume multiple-day command cycles for specific low-level actions like instrument placement [6,7]. Rovers autonomously traveling long distances collect limited science data en route [8]. This impacts astrobiology studies because accessible (paleo-) habitat indicators are most probably sparse, isolated, and invisible from a distance [9]. Future astrobiology missions must quickly constrain the geologic setting to identify sampling locations [10].

To this end, onboard science understanding can play an important role in fully utilizing each command cycle and preventing wasted observation opportunities. Onboard data understanding can benefit science return by summarizing terrain encountered during travel and directing autonomous instrument deployment to targets of opportunity. Recent innovations such as the AEGIS system have demonstrated the ability to discover science features of interest in images and automatically perform targeting for followup observations [3]. However, most mission concepts do not utilize onboard science analysis and many relevant features can pass undetected [11]. Onboard data understanding could assist Mars sample caching or return missions where the spacecraft must survey vast areas to identify potential sampling sites and habitat indicators [12]. It could also be highly valuable for determining the presence of chemical precursors and potential habitats for life in the outer Solar System. A Titan balloon, Europa submarine, or an Enceladus lander would likely have even less bandwidth and fewer opportunities to communicate with Earth [13]. In each of these scenarios, onboard data understanding could significantly improve reconnaissance for evidence of ancient climates, extinct life, and potential habitats.

This paper describes a new class of imaging instrument that aims to improve scientific return by increasing science autonomy and observation capabilities both when the spacecraft is stationary and when it is in motion. In this concept, image analysis uses *texture channels* to differentiate and map geologic surfaces. Here we use the term "texture" not in the geologic sense of formal physical properties, but rather in the computer vision sense to signify statistical patterns of image pixels. These numerical signatures can automatically distinguish geologically relevant elements such as roughness, pavement coatings, regolith characteristics, sedimentary fabrics and differential weathering in outcrop. Although one cannot predict these elements' appearance to a level that would give robots the subtlety of human interpretation, we propose that there is sufficient predictability — which improves during the course of a mission — to enable a useful level of scientific autonomy. Even rudimentary analysis can

still make certain predictable, relevant scientific

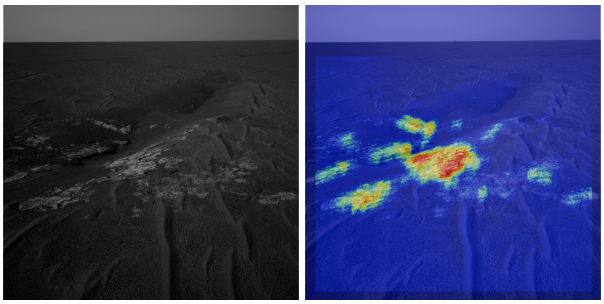

Figure 1: Meso-scale surficial mapping. A camera trained on previous data extrapolates to a new scene, identifying rock outcrop for target selection and/or summary downlink. Actual results on ground-processed data are shown. Image credit: NASA/Caltech/JPL, 1N134591492EFF1000P1803L0M1.

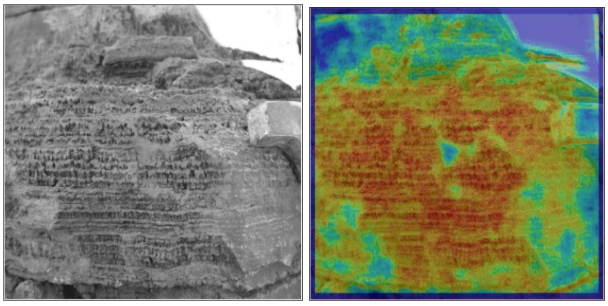

Figure 2: Outcrop mapping for fine-scale instrument placement activities. Here, layers of a stromatolite are automatically recognized by extrapolation from previous scenes [15].

distinctions during long traverses [14]. Scientists could use these distinctions to specify adaptive sampling and mapping behaviors, maximizing each command's science return. Autonomy would complement scientists' traditional options of "blind sampling" at scripted locations, or days of command cycles for each instrument placement action.

Onboard texture analysis can improve science return with new operational modes, of which two deserve special mention. In the *mapping operational mode* a texture-sensitive imager summarizes (and potentially provides a cursory evaluation of) the visited terrain. Summary maps, compressed numerical descriptions, exemplar textures and/or false color "cartoon" illustrations could characterize terrain during long-distance drives. These can alert scientists to the need for further investigation of potential (paleo-) habitat indicators [16] and maximize the useful information communicated in every transmission to Earth [17]. In the *adaptive sampling* operational mode, a texture-sensitive imager will identify distinctive surface elements or fabrics to guide automatic instrument placement during a single command cycle. This would enable the spacecraft to acquire, for example, targeted images, spectroscopy or microscopic imagery from representative surfaces without the need for scientists to see each target in advance and uplink incremental instructions. Each of these modes can be applied at multiple scales. Figure 1 illustrates a mesoscale application in which a rover autonomously identifies rock surfaces in a scene (similar to the AEGIS system) and identifies priority target locations for followup measurements. The heat map shows the probability of the "rock" surface, with higher probability pixels in warmer colors. Figure 2 shows a smaller centimeter-scale problem of identifying sample locations in rock outcrop. These outcrops could be approached from a distance during single cycle instrument placement; the ability to autonomously find the first sampling locations upon arrival could save many command cycles over the course of a long mission. Figure 2 shows the automatic identification of layered surfaces, from [15].

This paper describes one promising algorithm for automatic texture classification during exploration. A machine learning approach leverages labeled examples provided by a domain expert. These are used to train a *decision forest* classifier [18] to classify each pixel in any new scene. Section 2 describes our image pre-processing and training method in greater detail. Section 3 reports experimental results on the canonical task of rock surface mapping, applied to two Mars Exploration Rover panoramas consisting of dozens of images. Finally, we discuss future plans for parallelization and incorporation in Field Programmable Gate Array (FPGA) hardware. The ongoing *TextureCam* project aims to instantiate these algorithms in flight-relevant hardware for routine use on remote exploration platforms. These early validation efforts will be critical for guiding algorithm selection during later hardware development stages.

2. APPROACH

We formulate texture analysis as *supervised* classification of each image pixel from training examples. The workflow includes offline and real-time components. Figure 3 shows an example with images related to the specific challenge of automatic rock detection, which we will pursue at length in section 3. The input data sources might differ for other platforms or image analysis tasks, but the basic procedure would be the same. An offline stage builds a classifier model using previous images. System designers label a training set of image pixels according to geologic surface type, and train a statistical model that relates the local statistics of each pixel to its class. This model can be represented in a few tens of kilobytes for uplink to the remote explorer, or several models could be constructed in advance with selection at run time based

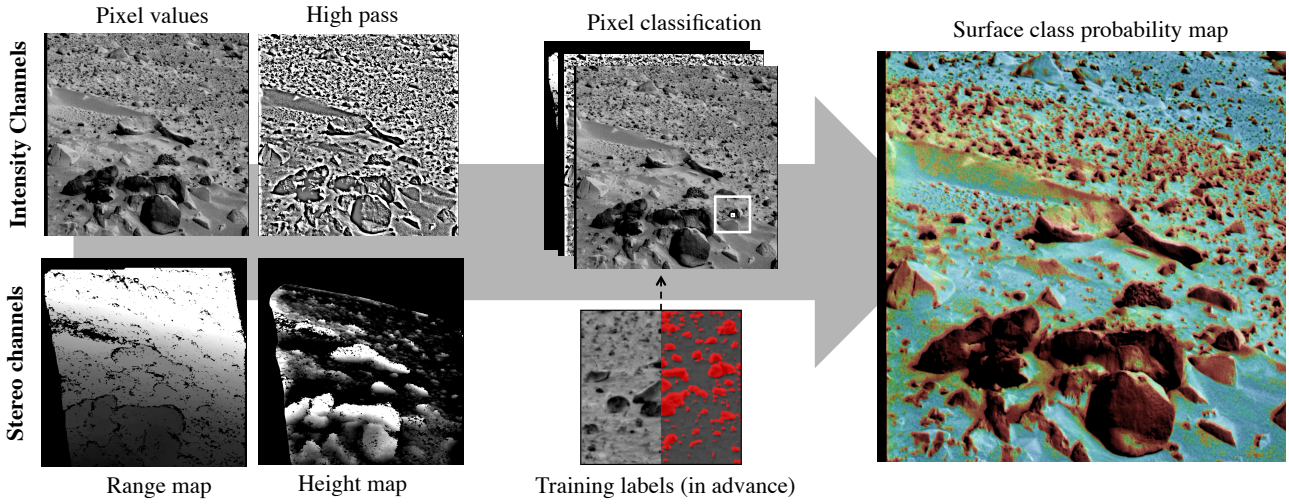

Figure 3: System architecture. We classify each pixel based on attributes computed from one or more image layers. The classifier uses learned patterns in a local window around each pixel to determine the posterior probability for each class.

on the needs of that day's command cycle. The second stage operates in real time, where the smart camera extrapolates these relationships to classify new scenes.

2.1. Preprocessing

The classifier models the probabilistic relationship P(Y|X) representing the probability of pixel class Y given the attributes X which characterize its local image neighbourhood. We learn this relationship from a training set of labeled pixels. Each pixel p is associated with an attribute vector $\boldsymbol{x} \in \mathbb{R}^n$ and class label $y \in \mathbb{N}$. In practice the attributes take the form of a stack of 2D image layers such as colors, multispectral image bands, the result of other preprocessing operations, or stereo products such as pixel ranges or Euclidean positions. In general it is desirable to use a rich set of attributes that incorporates domain knowledge where appropriate. Our tests use four channels appearing at left in Figure 2. The channels are:

Channel 1. Normalized Image Intensity records the raw radiance value for each pixel. We perform a simple contrast enhancement by stretching the middle span of image intensities, three standard deviations from the mean, to cover the entire pixel range. We compute this transformation based on the terrain in the image and excluding the bright sky. For convenience, we find terrain pixels using *k-means* clustering applied to the 3D vector of pixel row, column, and intensity. We choose the larger of the two regions as the terrain area, and compute a contrast-enhancing transformation that is then applied to all pixels in that image. Naturally this is simply one of many possible strategies for contrast enhancement and normalization.

Channel 2. Filtered image intensity Large shadows may still persist after histogram normalization. In principle surface texture analysis should be invariant to these incidental factors, so we include an additional channel that attempts to remove these large-scale intensity variations. We smooth the image with a rectangular moving average filter (size 32 pixels), and subtract this from the original. The result is akin to a high-pass frequency filter that emphasizes small-scale textures.

Channel 3. Stereo range data. Stereo range is a useful cue because texture appearance could vary with spatial resolution. A separate range channel encodes the distance from the camera center as pixel intensities. Ranges are suitably floored and ceilinged to cover the span of training data. This work is limited to ranges within 2 and 20 meters for which stereo data is available.

4. Stereo height data. Stereo data can be more directly useful expressed as height above the local ground plane. This helps to reveal protrusions such as rocks. Previous robotic investigations have used planar ground models, but this can be problematic in the presence of dunes and other natural undulations in the terrain. For this work, we use a more flexible model that (1) calculates Euclidean rover-relative coordinates for each stereo pixel, and (2) smooths the vertical direction to remove protrusions such as rocks from the ground model. Specifically, we compute the local median height value from a 200x200 pixel window. The difference between the pixel's actual measured vertical distance and the local median serves as a terrain-relative height measurement (Figure 4, Right). This occasionally introduces inaccuracies near the edges of the available range data.

We form the final attribute vector **x** by applying simple

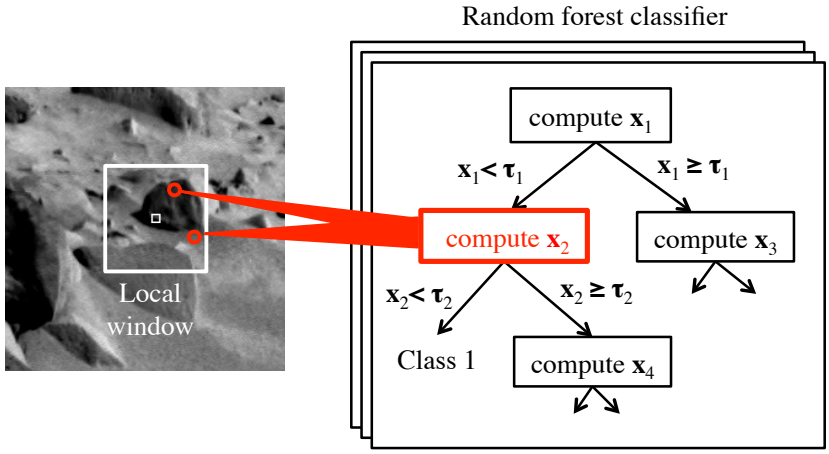

Figure 4: Classification approach. Each decision in the random forest classifier computes an attribute by testing one or predefined two pixel locations in a local window relative to the target pixel. It applies a simple arithmetic operation, such as a difference or ratio, and compares the result to a predetermined threshold. In this work each tree node uses just one image channel at a time. This need not be the case in general (see text).

arithmetic operations to two values drawn from a local window centered on the pixel, similar to Shotton et al. [19]. Each attribute applies one of five basic operations to any two pixels from this square region. The basic operations are: the *raw* value of one of the pixels; the *absolute difference* between the two pixels; the *difference* between pixels; the *sum* of the two pixels; and the *ratio* of the two pixels. These values can be computed efficiently to sufficient accuracy using integer arithmetic with a bit depth of 8. Each operation uses just one of the four image layers at a time. This provides a total of approximately 655360 potential attributes for each pixel (4 layers x 2 pixels x 1282 positions x 5 features). A natural extension of the method would be to operate on pixels across multiple image channels to let the classifier compute values such as multispectral band ratios. This has proved useful in other applications with richer spectral data products [18].

2.2. Training procedure

After assembling a suitable training set, we model the relationship P(Y|X) using a decision forest [19]. This is actually a collection of independent classifiers trained on random subsets of the training data. The classifiers are trees consisting of a sequence of nodes $l=\{l_1 \ldots l_n\}$. To find P(Y|X) for a new pixel, one begins at the root node of the tree, and moves left or right at each node l_i by comparing a specific attribute $\mathbf{x}_i$ with a threshold τ_i. Features can be computed from the raw image only as needed to progress down the tree, reducing the necessary computations from thousands to a few tens of operations per pixel.

Our training procedure grows each new tree from a single root node. We form a training set by random draws with replacement from a set of pixels that have been labeled in advance. For a class appearing *n* times in the training data, we weight each pixel by a factor 1/n as in Shotton et al [20]. This helps correct class imbalance in the training set, forcing an even prior P(Y) for all classes. Naturally, one can incorporate domain expertise in the form of biased priors if this information is known in advance. All training pixels in the subset for that tree start at the root. We find the optimal $\mathbf{x}_i$ and τ_i by trying random attributes $\mathbf{x}_i$, and for each candidate searching over all possible thresholds to find the one which maximizes an expected posterior entropy objective [20]. We apply the threshold test to separate the population at that node into a left-child subset L and a right-child subset R. We define weighting factors α_L and α_R to be the fraction of total pixels assigned to each group. A weighted sum gives the expected entropy upon testing the attribute at that value:

$$E[H(L,R)] = \alpha_L \sum_y p(y \mid \mathbf{x}_i < \tau_i) \log p(y \mid \mathbf{x}_i < \tau_i) + \alpha_R \sum_{y \in R} p(y \mid \mathbf{x}_i \geq \tau_i) \log p(y \mid \mathbf{x}_i \geq \tau_i) \quad (1)$$

This measures the expected information gain achieved by applying the threshold test. The training process searches 1000 random attributes for each node expansion operation, and use the best-scoring attribute/threshold pair as the final splitting criterion for that node. Each splitting operation generates two child nodes that can then be expanded recursively until the maximum number of expansions is exceeded or until some predetermined depth is reached. Each leaf node of the tree retains a count of the training pixels from each class that has reached that leaf, and the distribution at the destination leaf is a posterior probability estimate $P(Y|X=\mathbf{x}_i)$.

At runtime, one classifies a new pixel by propagating it down the tree until it reaches a leaf. Each tree gives a separate probability estimate. We average the result from all trees for the final classification result for the forest. Apart from its speed, we favor the random forest classifier because it has already been found effective in terrestrial scene understanding applications [20]. It is

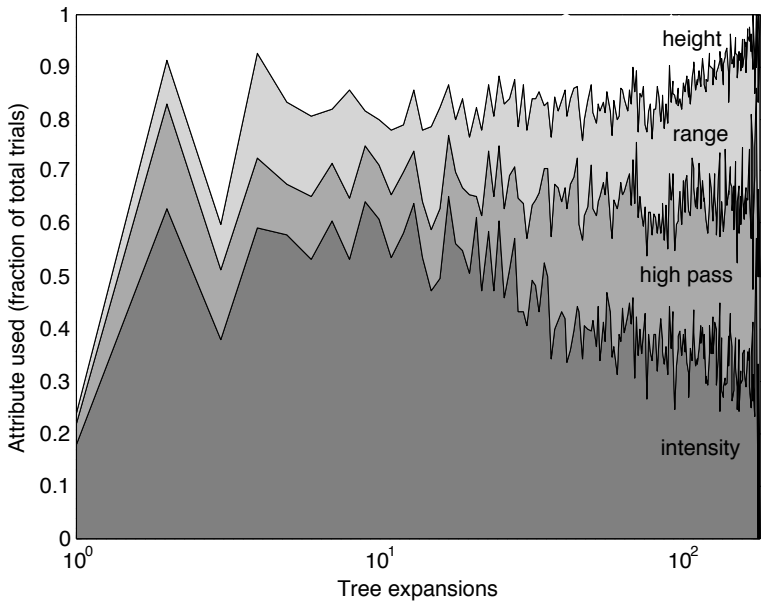

Figure 5: Different channels dominate the selected attributes selected at various tree levels. The first level almost always uses an attribute derived from the height channel. Note the logarithmic scale. Here 200 tree expansions were used. All channels contribute to the classification results, though the stereo channels diminish in importance for later expansions. These represent deeper layers in the tree where pixels are more ambiguous and subtler decisions are required.

simple to implement, and its posterior probability estimates are valuable for instrument placement or other activities involving risk to the spacecraft. It elegantly incorporates arbitrary input data channels, making it relevant for multispectral images or full imaging spectroscopy. It can accommodate multiple classes where needed. Training time scales with the number of attributes searched, rather than the number of potential attributes in the dataset itself. Thus, it is somewhat robust to redundant or non-informative image layers. Finally, the pixel-wise classification is amenable to hardware implementations in Field Programmable Gate Arrays (FPGAs) because it is parallelizable, based on local pixel data, and uses integer arithmetic. Designers can adjust the number and depth of decision trees for the best balance between classification performance and speed.

3. EVALUATION

This section reports a test involving the application of rock detection. This is a challenging image analysis problem; rocks exhibit varied morphologies and their appearance can change radically under different illumination conditions [21,22]. Nevertheless, the problem is important for autonomous geology due to rocks' value as targets for point spectroscopy and contact sampling. Rock detection is a useful test case because they are discrete objects with relatively objective boundaries, permitting a quantitative performance analysis. We focus this study on a set of panorama images from the Mars Exploration Rover (MER) catalog.

3.1. Data Sets

We consider two panoramas collected with the MER Pancam that have been studied extensively in previous work [23]. The Spirit rover acquired the "mission success" panorama during first week of rover operations. It acquired the "legacy" panorama over a multiple-day period beginning on Sol 59. We use several tens of sub-images from each panorama, including all terrain in the 2-10m range. These two panoramas provide training and test cases, showing how the surface classification can generalize to new environments. For consistency, we exclude images containing rover tracks since these features appear in only one of the two panoramas.

We obtained evaluation labels from a set of rock regions that had been manually labelled by analysts during the initial Golombek et al. study. This comprehensive analysis labelled all rocks within selected terrain areas down to particles 10 pixels or greater in size. The resulting list includes hundreds of individual rocks per image [23]. The resulting outline traces were re-projected into the frame of the MER radiance-corrected PDS MRL product [24], using SIFT interest point matching to fit a homography transform [25]. This produced excellent alignment with the original PDS products and pixelwise coregistration with stereo data. During the original analysis, the analysts only labelled rocks but not the background terrain that could be guaranteed to exclude rocks. Both are necessary for a Receiver Operating Characteristic (ROC) performance analysis. Consequently it was necessary to infer the background region. We defined the convex hull of pixels around the trace labels as the reviewed image area. Portions of this area within traces were classified as "rock," and others "terrain." Pixels beyond the convex hulls were treated as unclassified and excluded from the training procedure. The test dataset includes 46 different images. The rover panoramas were both acquired over several acquisition periods on different days, and each spans a range of different illumination conditions and directions.

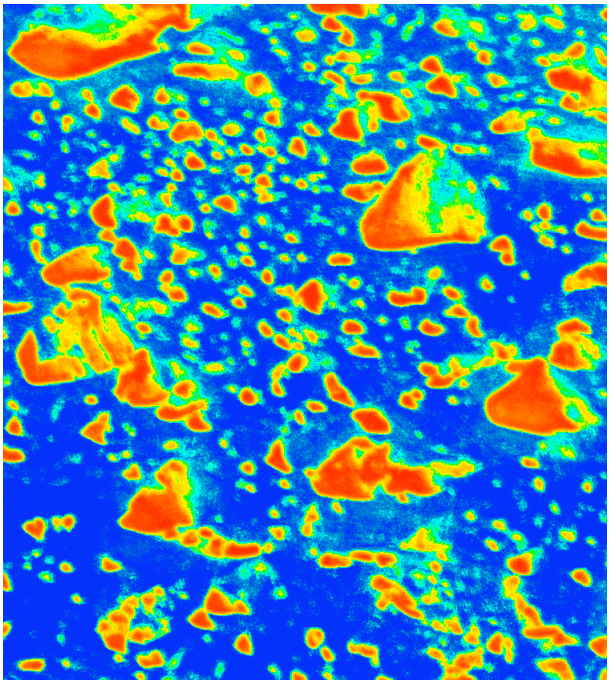

Figure 6: Rock probability map. Warmer colors signify a higher probability of the pixel lying on a rock surface. This certainty score permits rover followup decisions to be tuned to available resources, risk and target value.

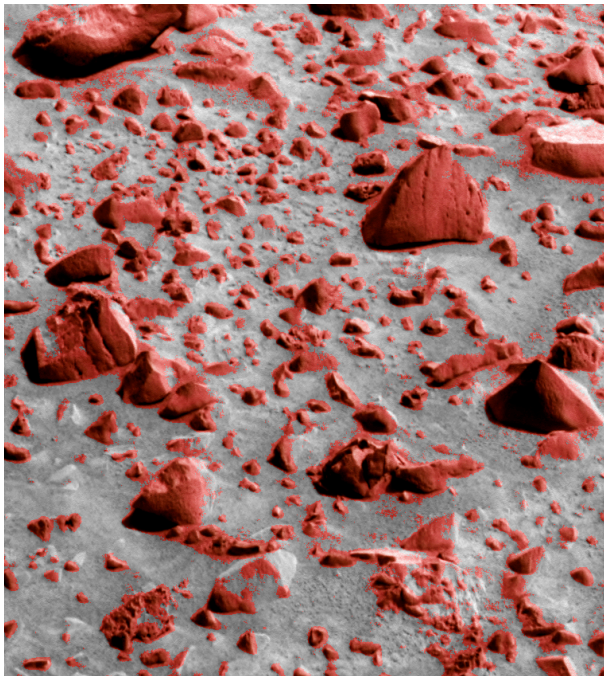

Figure 7: Thresholded classification result (typical). An example showing automatic classification of MER Pancam image 2P131596386MRL1155P2216L2M1, NASA/Caltech/Cornell.

These tests describe the result of training a random forest classifier on the "mission success" panorama and applying the resulting classifier to the "legacy" panorama. We selected 23 images for the training set and constructed a decision forest using a random subset of 10^6 pixels. The entire training procedure requires less than an hour to complete on a standard consumer-grade processor. As noted, there is flexibility in the number and depth of trees used for the classification, which permits variable time requirements. These tests intended to evaluate best-case performance when time was not a factor, so very lenient forest sizes were used. We trained a random forest of 200 trees, permitting 200 expansions. However, in our experience these results are similar to those of much smaller forests with 8-16 trees. Figure 5 shows an area plot to illustrate the features that were favoured at each expansion. Curiously, the initial root node nearly always used the height channel. Inspection of the resulting trees suggests that the trees generally threshold these pixels at a high value, e.g. it used the first decision to partition "tall pixels" from the rest. This is reasonable as such pixels are the most unambiguous classifications. After this coarse height partitioning, later levels rely increasingly on image-space attributes. The height channel quickly declines in relevance until the very end of the training procedure where it is almost never used. Height and intensity features are most important overall, but all the channels seem to contribute to the classifications to some degree.

3.2. Results

We apply the trained decision forest to the "legacy" panorama. After applying range constraints the data set contains 23 images, with many thousands of rocks and millions of classified pixels. Figures 6-7 show typical results, with rock probability and classification maps produced by the full set of image features. The dark regions of shadowed rocks are most obviously rocks, but the system is also capable of using stereo data to discriminate the lighter faces of large rocks. Figure 8 shows a Receiver Operating Characteristic (ROC) curve detailing the performance of the rock detection system for a system trained on intensity channels alone (e.g. the intensity and high pass channels), and another using the full set of channels with stereo data. Precision/recall rates compare favourably to those reported previously in the literature for purpose-built rock detection algorithms [23], though one should be careful in drawing a direct comparison due to the difference in evaluation methodology described above. Precision and recall also varies by image: the error bars show the standard deviation across 20 images in the test set, excluding the extreme deciles. An actual confidence interval on the mean performance for the panorama dataset would be smaller. We also note that small rocks were not labelled during manual analysis. These rocks are easily identified by the classifier, and often appear as false

positives. This could cause the scores reported here to understate actual performance.

The most confident rock classifications are nearly always *bona fide* rock surfaces, suggesting that performance as a target selector for followup spectroscopy or imaging would be very good. Figure 9 quantifies the benefit over random selection, showing the target selection failure rate. This is defined as the probability of selecting a terrain pixel at the strictest reasonable threshold. To evaluate the random approach we simply count the actual number of rock and terrain pixels wherever there is valid data within our 2-10 m range window. A random target selection approach would fail about three times as often as it would succeed. Performance for some sparse scenes is considerably worse; some images require 10 or more misses for each hit. We score the adaptive methods based on the pixels with highest probability of being rocks. We perform a percentile grid search over probability scores, retaining the first threshold that results in a nonzero number of rock pixels detected. The rock mapping system trained on image intensity data improves performance considerably over random selection. 90% of selected pixels would be correct on the first try, for a net improvement of approximately 30x. Including stereo data as an input channel gleans additional benefit, with a true false positive rate of 0.3% in the median. This represents an improvement of two orders of magnitude over random target selection. Each of these differences is statistically significant (p=0.05). For clarity the illustration floors all failure rates at 10^{-2}, but both trained methods have images for which no false positives are returned at all.

4. DISCUSSION

These initial tests show promise on several fronts, including: learning geologic terrain types from labelled examples; generalization across different scenes, terrain types, and illumination conditions; high-precision classifications for target selection; and area classification for mapping. This approach is entirely generic, but its performance on the rock detection task seems fully competitive with specialized state-of-the-art approaches. The ability to train a single algorithm for many different image analysis tasks is a valuable characteristic of the approach. Future work will continue to mature these technologies for other tasks such as outcrop analysis and instrument placement. Another important direction for future development is the incorporation of these algorithms into FPGA hardware. In coming months the on-going *TextureCam* project will encode the decision tree approach into a parallelized FPGA implementation that could run at near frame rate while placing very little computational load on the spacecraft avionics system. It is our hope that this approach, or one similar to it, could someday reduce design costs and pave the way for routine science image analysis by robot spacecraft.

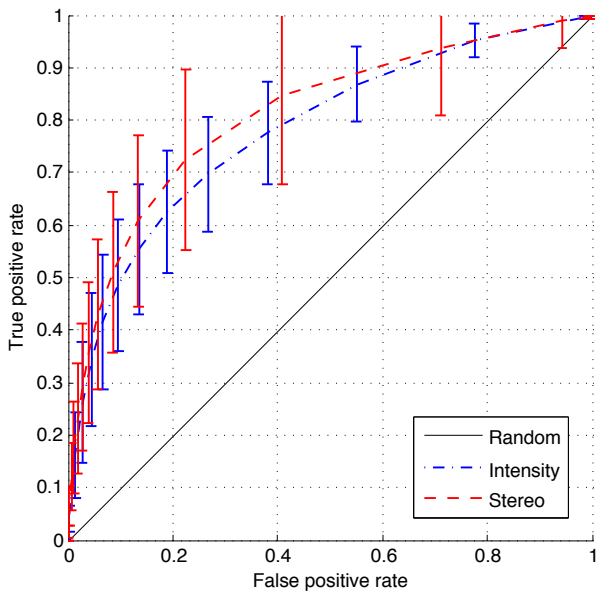

Figure 8: Performance ROC for the Legacy panorama. The classifier was trained on the Mission Success panorama. Both datasets consist of 23 images.

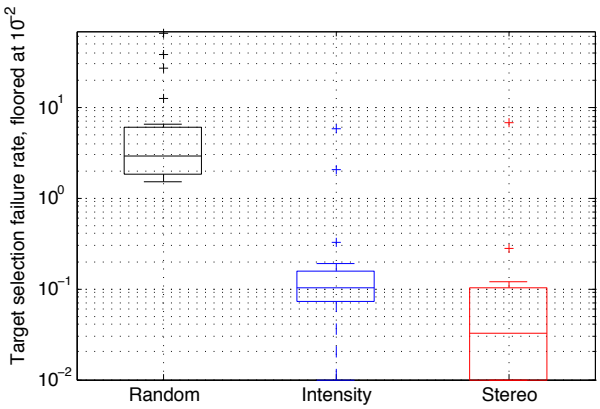

Figure 9: Target selection performance for random pointing, single image features, and full stereo data. Onboard processing improves target selection performance by a factor of 30-100x.

5. ACKNOWLEDGEMENTS

We thank the members of the OASIS team, including Robert Anderson, Mike Burl, Dan Gaines, and Michele Judd, for their advice and inspiration. This research was carried out at the Jet Propulsion Laboratory, California Institute of Technology. Copyright 2012 California Institute of Technology. All Rights Reserved; U.S. Government Support Acknowledged. The TextureCam project is supported by the NASA Astrobiology Science and Technology Instrument Development program (NNH10ZDA001N-ASTID).

6. REFERENCES

1. Chien, S. A. et al. (2005) Using autonomy flight software to improve science return on Earth Observing One. Journal of Aerospace Computing, Information, and Communication 2:4.
2. Thompson, D. R., Wettergreen, D. and Calderon P., F., (2011). Autonomous Science for Large-Scale Robotic Survey. *Journal of Field Robotics,* 28(4):542.
3. Estlin, T., Bornstein, B., Gaines, D., Anderson, R. C. Thompson, D. R., Burl, M., Castaño, R., Judd. M. (2012) AEGIS Automated targeting for MER Opportunity Rover. *ACM Transactions on Intelligent Systems Technology*. To appear.
4. Gulick, V. et al. Autonomous image analyses during the 1999 Marsokhod rover field test (2001). *Journal of Geophysical Research*, vol. 106:E4,pp.7745-7763.
5. Castaño, R., Anderson, R.C., Estlin, T., DeCoste, D., Fisher, F., Gaines, D., Mazzoni, D., and Judd, M. (2003), "Rover traverse science for increased mission science return," *Proceedings of the IEEE Aerospace Conference.*
6. Arvidson, R.E., Bonitz, R.G., Robinson, M.L., Carsten, J.L., Volpe, R.A., Trebi-Ollennu, A., Mellon, M.T., Chu, P.C., Davis, K.R., Wilson, J.J., Shaw, A.S., Greenberger, R.N, Siebach, K.L., Stein, T.C., Cull, S.C., Goetz, W., and Morris, R.V. (2009), "Results from the Mars Phoenix Lander Robotic Arm experiment," *Journal of Geophysical Research,* vol. 114, E00E02, doi:10.1029/JE003408.
7. Baumgartner, E.T., Bonitz, R.G., Melko, J.P., Shiraishi, L.R., and Leger, P.C. (2005), "The Mars Exploration Rover instrument positioning system," *Proceedings of the IEEE Aerospace Conference*, p. 1-19, doi:10.1109/AERO.2005.1559295
8. Carsten, J., Rankin, A., Ferguson, D., and Stentz, A. (2007), "Global path planning on-board the Mars Exploration Rovers," *Proceedings of the IEEE Aerospace Conference.*
9. Warren-Rhodes, K. et al. (2007). Robotic ecological mapping: Habitats and the search for life in the Atacama Desert, *Journal of Geophysical Research,* 112:G04S06.
10. Pratt, L.M. and the Mars Mid-Range Rover Science Analysis Group (2010), "The Mars Astrobiology eXplorer-Cacher (MAX-C): A potential rover mission for 2018." *Astrobiology,* 10(2):127-163.
11. Castaño, R., Judd, M., Estlin, T., Anderson, R.C., Gaines, D., Castaño, A., Bornstein, B., Stough, T., and Wagstaff, K. (2005b), "Current results from a rover science data analysis system," *Proceedings of the IEEE Aerospace Conference.*
12. Borg, L. and the Next Decade Science Analysis Group (2008), Science Priorities for Mars Sample Return, *Astrobiology,* 8(3):489-535.
13. Hayden, D., Chien, S., Thompson, D. R. Castaño, R., (2012). Using Clustering and Metric Learning to Improve the Science Return of Remote Sensed Imagery. *ACM Transactions on Information Systems Technology*.
14. Smith, T., Thompson, D.R., Wettergreen, D., Cabrol, N.A., Warren-Rhodes, K.A., and Weinstein, S.J. (2007), "Life in the Atacama: Science autonomy for improving data quality," *Journal of Geophysical Research*, vol. 112:G04S03
15. Thompson, D. R. et al. (2012). TextureCam: Autonomous Image analysis for Rapid Astrobiology Surveys. *Lunar and Planetary Science Conference, 2012.*
16. N. Cabrol, et al. (2007). Life in the Atacama: Searching for life with rovers (science overview). /Journal of Geophysical Research, vol. 112, G04S02, 25pp.
17. Thompson, D. R., Smith, T., Wettergreen, D (2008). Information-Optimal Selective Data Return for Autonomous Rover Traverse Science and Survey. *IEEE International Conference on Robotics and Automation*.
18. McLaren, D., Thompson, D. R., Davies, A. G., Gudmundsson, M. T., Chien, S. (2012). Automatic identification of volcanic ash plumes using WorldView-2 imagery. *Proceedings of the SPIE*. In press.
19. Breiman, L. (2001). Random Forests. *Machine Learning* 45(1):5.
20. Shotton, J., Johnson, M., Cipolla, R., (2008). Semantic Texton Forests for Image Categorization and Segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* 1–8.
21. Dunlop, H., Thompson, D. and Wettergreen, D. (2007). Multiscale Features for Detection and Segmentation of Rocks in Mars Imagery. *IEEE Conference on Computer Vision and Pattern Recognition*.
22. Thompson, D. R. and Castaño, R. (2007). A Performance Comparison of Rock Detection Algorithms. *Proceedings of the IEEE Aerospace Conference.*
23. Golombek M. P. et al. (2005) Assessment of Mars Exploration Rover landing site predictions. *Nature* 436, 44-48.
24. Maki, J. et al. (2007). Mars Exploration Rover Software Interface Specification. Camera Experiment Data Record (EDR) and Reduced Data Record (RDR) Operations and Science Data Products. JPL D-22846. Jet Propulsion Laboratory, California Institute of Technology.
25. Vedaldi, A., Fulkerson, B (2010). VLFeat: An open and portable library of computer vision algorithms. *Proceedings of the international conference on Multimedia*. ACM pp1469—1472.