

DATA MINING DURING ROVER TRAVERSE: FROM IMAGES TO GEOLOGIC SIGNATURES

David R. Thompson, Trey Smith, and David Wettergreen

The Robotics Institute, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh PA 15213, USA

ABSTRACT

Soon robotic explorers will be able to produce more scientific data than can be transmitted or interpreted efficiently. We present a method for characterizing geology during rover traverse using autonomous data analysis techniques. The strategy detects discrete geologic features in images; distributions of these features constitute a signature that correlates with the geology of each site. These numerical profiles reveal subtle trends and boundaries in geologic units that facilitate targeted sample selection and efficient data analysis. We demonstrate the system's use on field data collected during a field expedition to the Atacama Desert of Chile.

1. INTRODUCTION

Advances in planetary rover technology enable robotic explorers to produce data sets too large to analyze in detail. Future rovers will travel kilometers between command cycles while potentially crossing multiple geologic units and collecting hundreds or thousands of images and spectra. Meanwhile mission resources — time for taking measurements, bandwidth to return the data and the human resources to inspect it — will not keep pace with the growing data volume. The result is a new series of bottlenecks that prevent scientists from experiencing the full benefit of improved rover capability [1]. A smaller fraction of possible science targets will be investigated, a smaller fraction of gathered data will be returned to scientists on Earth, and a smaller fraction of returned data will be analyzed thoroughly.

These issues have motivated research into autonomous analysis of science data. “Science autonomy” describes the ability of a system to understand collected data in order to make more effective exploration decisions. Onboard science autonomy helps relieve resource bottlenecks by focusing rover activity on the most important data. By recognizing science targets and reasoning about goals a rover can autonomously prioritize key features for analysis and return. Meanwhile offboard science autonomy helps scientists on Earth to analyze incoming

data quickly. Our group and others are studying several aspects of science autonomy, including image analysis [1, 2, 3], science priority representation [4, 5, 6], and planning [7].

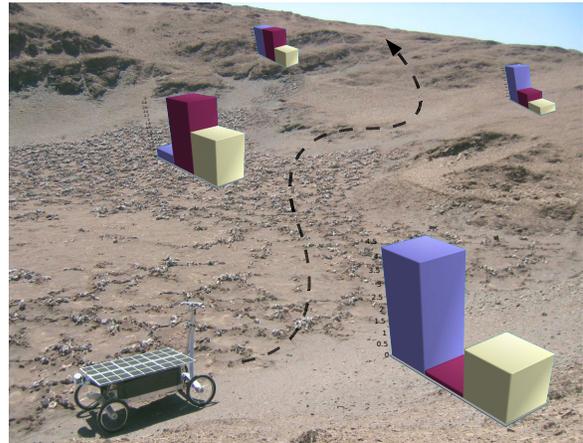


Figure 1. Concept illustration of autonomous geologic profiling in the Atacama Desert of Chile. “Zoë,” the rover used in the field experiments, appears in the foreground. The science autonomy system profiles different locations on the rover’s traverse by generating distributions of different classes of rocks.

Here we focus on the specific science autonomy problem of detecting geologic trends in rover images. In our model a rover collects image data at several coordinate locations, or “locales.” We profile a locale’s geology using the quantities of certain discrete features that are present. This reduces a large data set to a simple statistical signature characterizing each locale. These signatures reveal subtle geologic trends, the borders between geologic units, and anomalous geology that is different from neighboring locales (Fig. 1). Geologic signatures are also useful onboard; a rover equipped with information about unit boundaries could ensure that measurements from each distinct geologic region is returned to Earth. Finally, the signatures themselves provide compact statistical summaries of data sets that are far too large to downlink.

This paper demonstrates autonomous geologic profiling

by computing signatures from rover field data. Section 2 begins by presenting different feature extraction methods. We discuss a simple pixel-color approach and a more sophisticated strategy involving rock detection. Then section 3 tests these techniques on imagery from autonomous rover traverses in the Atacama Desert of Chile. The experiments use one data set made up of images concentrated at a few widely-separated locales, and another composed of periodic samples from a linear traverse. We compare the resulting geologic signatures to qualitative human interpretations. These tests suggest that autonomous profiling of regional geology is accurate enough to help resolve resource bottlenecks by summarizing data sets too large to downlink and assisting with their analysis after downlink.

2. FEATURE EXTRACTION

Geologic profiling is a technique that describes the geology of a locale using the discrete features found there. We posit an unknown stochastic function which maps a locale’s geologic class onto the quantity of each feature. These features in turn generate noisy detections in the rover’s onboard feature detection algorithm. The ability to recover a geologic signature from detections relies on the precision and invertability of these two mappings.

$$\textit{Geology} \rightarrow \textit{Features} \rightarrow \textit{Detections} \quad (1)$$

This formulation underscores the importance of a good set of features. Ideal features would follow reliably from geologic class and produce reliable detections. In practice the designer must compromise between these two goals. On one extreme the system might extract “simple features” like pixel colors that are trivial to detect but geologically ambiguous. A large enough sample of simple features might still yield a reliable geologic signature. An attractive alternative is to find geologically relevant features like rocks and patches of soil in the images. These “complex features” are similar to what a human geologist might consider when describing the locale. They combine many attributes in complex, empirically guided ways. We have examined both simple and complex feature techniques for generating geologic signatures from rover field data. Section 2.1 presents the simple feature technique, while Section 2.2 describes the complex feature method by considering rock detection and classification in detail.

2.1. Minimal Features: Color

A common method for database indexing computes a histogram of the colors of all the pixels in an image [8, 9]. The distribution of pixel colors becomes the statistical fingerprint that permits comparisons between images. While the simple features themselves are geologi-

cally uninteresting, combining many of them might still result in a discriminating signature.

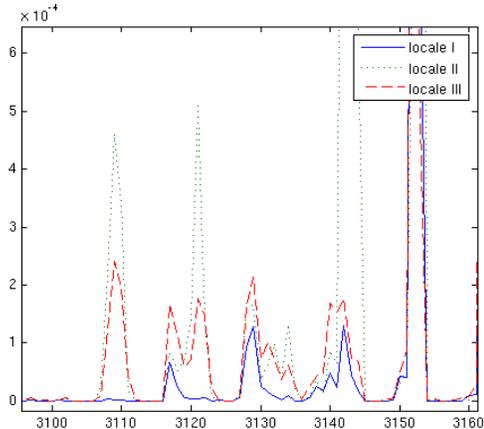


Figure 2. A small portion of the color histogram used for comparing locales I-III in the first field experiment. The vertical axis represents the percentage of image pixels in each color bin.

Our “simple feature” signature utilizes a color histogram. We calculate each pixel’s Hue/Saturation/Value (HSV) color and place the result in a 3-dimensional histogram with 32 bins along each axis. Only a small portion of the possible colors appear in any image so our feature vector is a selected portion of this histogram. In the following experiments we discount colors associated with the sky. The other 4875 colors that appear become features for image comparison (Fig. 2).

A color-based feature extraction should ideally use a formal color calibration to guard against illumination changes. While the following experiments relied on the automatic gain and balance features of the cameras, we compensated by restricting experiments to short tests under constant lighting. The bright sky washes out terrain colors near the horizon so we constrained color profiling to only use images where no skyline was visible.

2.2. Complex Features: Rock Detection and Classification

A complex feature method uses *geologic* features like rocks and soils [7, 10, 1]. It is not immediately clear that such a description would yield better geologic fidelity because detecting these features in field data is difficult. In the case of rocks, current detection strategies often find less than 80% of the actual features [2]. Additionally structural detection error might favor some features over others and distort regional statistics. For example, large rocks are easier to detect than small ones and shadows can be mistaken for rocks with some methods. In the worst case, complex features could result in a statistical descriptor that ignores geologic boundaries or hallucinates them where none exist.

If detection is accurate enough however then complex

features should provide a superior match to ground-truth geology. Complex features incorporate designer knowledge in the detection process, forcing the system to focus on geologically relevant data (like rocks) and ignore irrelevant changes (like the rover’s own shadow). Moreover, signatures based on complex features are themselves scientifically informative. Consider for example a count of rock sizes, which is useful to geologists but time-consuming to gather manually. A signature based on rock size distributions or other complex features can suggest possible geologic interpretations.

For simplicity and ease of ground-truthing we focus on distributions of rocks. The overall procedure has several stages (Fig. 3). Initially the system detects rocks in the images in a manner similar to [2] but with the addition of stereo data to provide information about scene geometry¹. An image pair is segmented into a set of candidate features that might turn out to be rocks. A region-merging segmentation algorithm processes multiple channels for each image, considering each color and stereo channel separately. The color segmentations find pixel regions which have a uniform hue, saturation or intensity that differs from the background. Stereo segmentation identifies an estimated ground plane with a least-squares fit, and then finds regions of homogeneous height that lie well above the plane as in [1]. This multichannel segmentation yields a set of pixel regions which might or might not be science targets.

The system calculates a vector of numerical attributes for each candidate region. These attributes include characteristics like the estimated size, height above the ground, color, and shape. Then a Bayesian belief network [15] trained on human-labeled data analyzes each vector to provide an output probability that the candidate region is a true rock. In this manner the system generates a list of detected rocks in the image.

Finally detected rocks are categorized into geologic classes. The choice of which categories to use is highly significant; it determines the distinctions that the system can make. To offer scientists more flexibility we permit categories according to both supervised and unsupervised definitions as in [4]. This scheme permits two general types of feature classes. The first, “interval classes,” define a binary decision boundary along specific attribute values chosen by the scientist. For example, quartz rocks might be identifiable by their high albedo; a scientist could instruct the system to label as quartz any rock which had an albedo greater than two standard deviations above the mean. While finding appropriate values might require some trial and error on the part of the scientist, the interval-based classifications are simple and predictable.

If a feature does not fall within a specific attribute interval it becomes a candidate for a second-tier of classification using a probability density model. The model uses several types of classes. Example-based classes are a supervised categorization that fits a multivariate

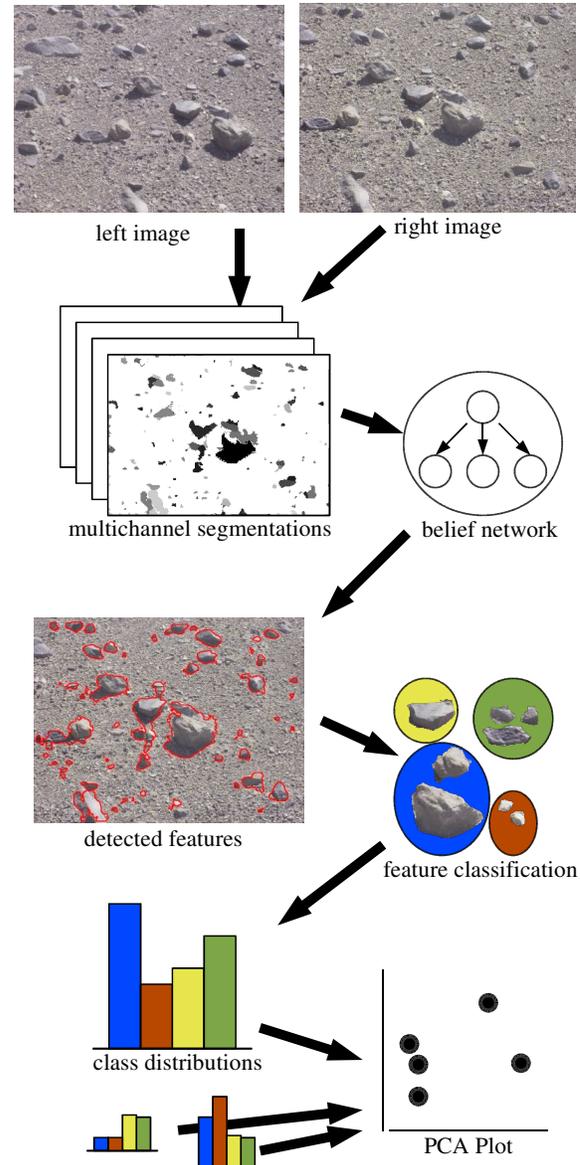


Figure 3. Complete procedure for rock detection and classification. A belief network classifies segmented image regions, and a classification routine categorizes the resulting features to generate class distribution signatures for each locale.

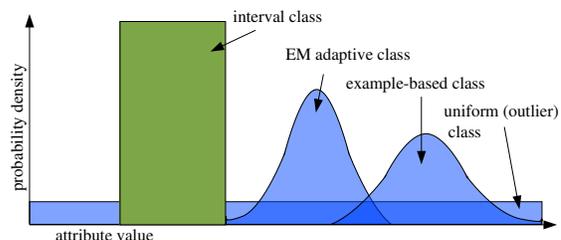


Figure 4. The interval class defines an absolute decision boundary, while the other classes constitute a probability density model of the remaining features. Only the EM-based Gaussian clusters are adaptive.

¹Other plausible rock detection strategies appear in [11, 12, 13, 14]

Gaussian distribution to representative examples provided by the scientist. The scientist could identify several samples of quartz in previous days' data; standard maximum-likelihood techniques [16] find an appropriate probability density distribution. These manually-defined classes are complemented by unsupervised expectation/maximization (EM) clustering [17] that fits a mixture of Gaussian distributions to remaining features. Finally, an "outlier" class exhibits uniform probability over the entire feature space. It captures novel features that have little in common with any other rock.

The resulting model gives the quantity of each different class at a locale, where classes are defined in terms of a heterogeneous mixture of manual and adaptive components (Fig. 4). This flexibility means that the scientist might favor supervised categories, unsupervised categories, or some mixture of the two. While the classifications do not always correspond to the distinctions made by a human geologist they are a useful metric for comparing different locales. Subtle changes like a shift in the density of a certain class can identify important border regions and suggest areas for further exploration.

We favor two techniques for visualizing and comparing class distribution signatures. If there are only a few classes a simple histogram offers an intuitive summary of the locale. With many different classes, however, the histograms may be hard to compare. Applying principal component analysis (PCA) and projecting these distributions onto their first two principal components generates a 2D plot that serves as a compact visual comparison of the different locales. This is especially appropriate for a simple features like pixel colors that could easily have over a thousand distinct classes.

3. EXPERIMENTS

We performed rover field experiments in the Atacama Desert, a region spanning several hundred kilometers in the North of Chile. Conditions of extreme desiccation mean that the Atacama is nearly devoid of macroscopic life. This together with its Mars-like terrain make it a good testbed for planetary rover technology.

The experiments in this section aim to address some of the aforementioned design choices with empirical data. Foremost we compare the geologic fidelity of complex feature extraction (i.e. rock detection) against simple features. In addition we test geologic models utilizing both supervised and unsupervised classification schemes. Finally we compare data collection strategies: a "concentrated" method that collects large image sets at a few well-spaced locales and a "high-frequency sampling" method that collects single images at many short intervals along a traverse.

The hardware platform used for the experiments is Zoë, an exploration robot developed at Carnegie Mellon [18]. Zoë is a solar powered rover with a dual passive axle de-

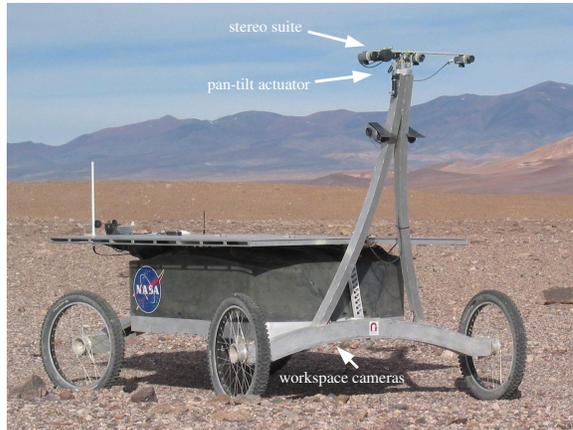


Figure 5. Zoë's science camera configuration.

sign that permits it to travel up to 1m/second and traverse slopes and rough terrain [19]. Onboard navigation uses stereo imagery and local path-planning to navigate between scientist-specified waypoints. Zoë also carries a variety of science instruments. It incorporates a 2-meter mast with a pan-tilt actuated stereo suite (Fig. 5). These cameras individually provide a 21-degree field of view, but we often generate mosaic panoramas to give complete coverage of a locale. Also used for these experiments are "workspace" cameras mounted under the body of the rover, which complement the large-scale panoramic views of a scene with close-up images of the ground.

3.1. Concentrated Data Collection

The first experiment consisted of a series of concentrated samples from a few well-separated locales. We tested this strategy during an autonomous traverse to the top of a rock-strewn hill. The rover began a distance from the base and traveled forward in 50 meter intervals, collecting panoramas and workspace imagery at each locale. Five locales were visited. The first three were situated on the approach to the base of the hill and the fourth was part-way up the hill to the top. Halfway between the fourth and fifth locales Zoë encountered steep terrain thick with obstacles that confounded autonomous navigation, so the fifth locale at the peak of the hill was reached manually.

To an untrained human observer the first three locales all appeared alike, with occasional patches of white sediment and few significant rocks. The hillside terrain at the fourth locale was different, however - here the sediment contained some large dark rocks along with many small white rocks. The white material was no longer present at the fifth locale, but the peak of the hill was covered with large gray rocks. Fig. 6 shows some of the collected data. Row A corresponds to the ground-truth human interpretation with dots representing rocks of different albedos. Row B shows sections of the 95-image panoramas collected at each locale. Row C shows underbody "workspace" imagery. While underbody images were not used in the autonomous analysis they provide some insight into terrain conditions at each locale.

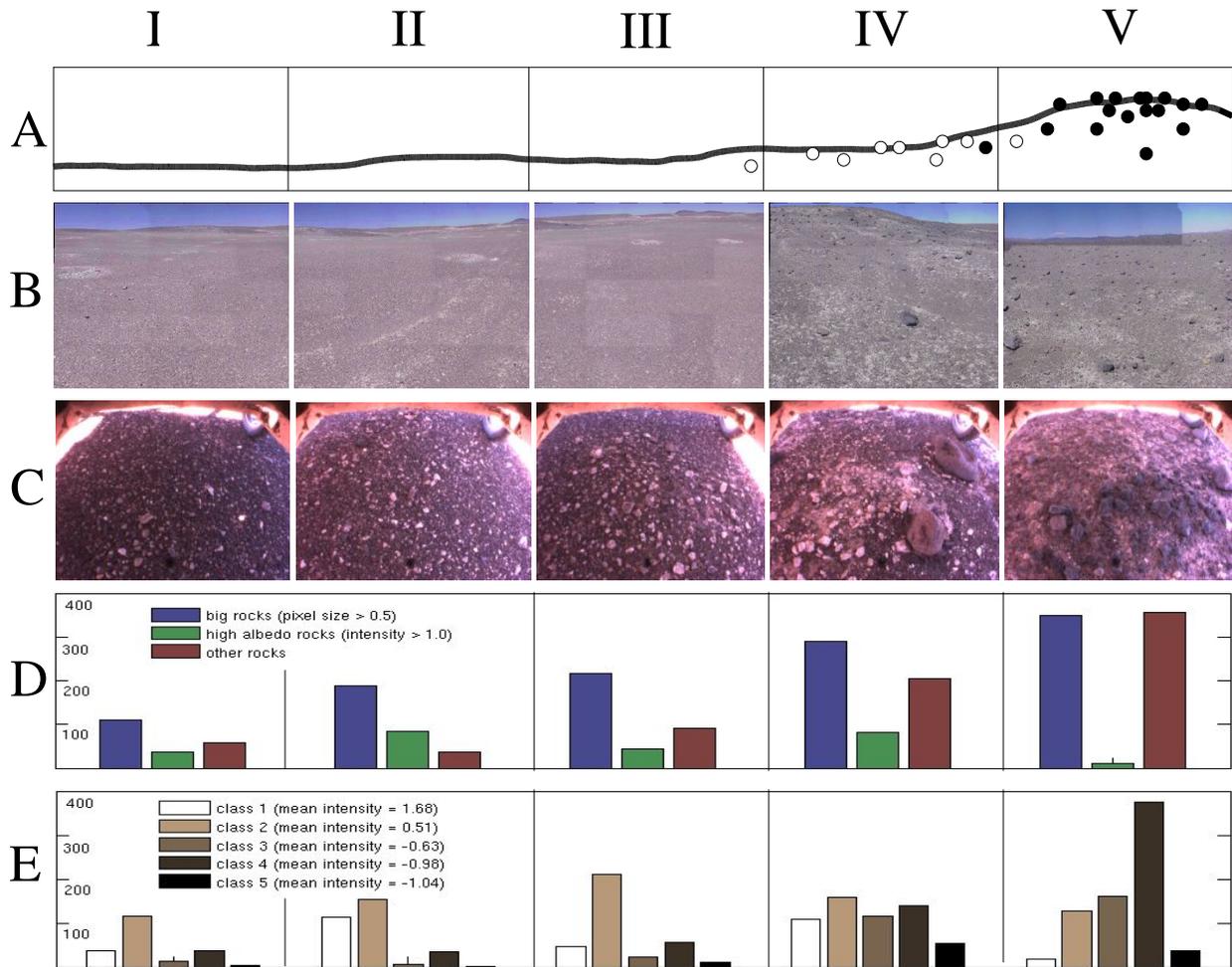


Figure 6. Locomotion profiles in the first field experiment. Row A: Qualitative human interpretation. The peak of the hill has the highest concentration of rocks. Row B: A portion of the panorama from each locale. Row C: Underbody images from each locale. Row D: Rock detection with manually defined “large rock,” “bright rock,” and “outlier” classes. Row E: Unsupervised model using rock detection; classification employs EM clustering over color and size features. The mean intensity of each class is shown.

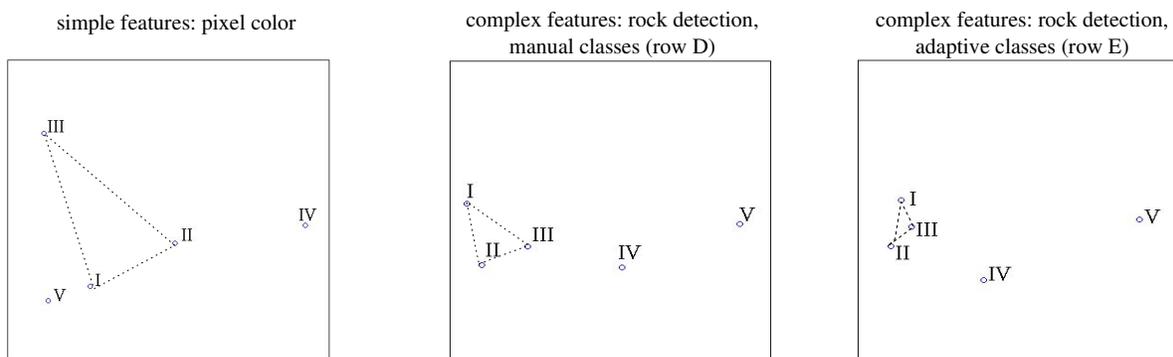


Figure 7. Feature distributions from each of the profiling strategies projected into 2 dimensions with PCA. Geologic signatures from the color histogram method (left) are misleading; they suggest a large difference between the first three locales. Rock detection with manual classes (center) gives better signatures. Rock detection with adaptive classification (right) best parallels the human interpretation.

We detected rocks in the five panoramas and classified the resulting features using a variety of different schemes. The first classification scheme favored manually-chosen interval classes: an albedo feature interval to account for the white material, a second interval based on rock size, and an outlier cluster to account for everything else. These distribution histograms appear in row D. The clusters show a general trend toward an increase in the number of rocks as one approaches the peak together with a corresponding drop in the proportion of white material. Note that in both unsupervised and supervised cases, a larger amount of white rocks appear in locales I-III than the ground-truth records suggest. This is due to patches of white sediment (visible the first three panoramas of row B) that can be mistaken for rocks.

The second classification used a completely unsupervised model. Five multivariate Gaussian clusters were initialized to random data points. Then EM clustering iteratively adjusted the clusters' parameters to converge on values that maximized their likelihood with respect to the entire data set. The resulting distribution histograms appear in Fig. 6 row E. Three color-channels and size features were used in clustering, but clusters divided mainly along the intensity axis. Because of this the legend provides mean intensity information to distinguish the clusters. This classification succeeds in detecting the general trend in number and types of rocks — fewer rocks in locales I-III, white material that disappears as one climbs the hill, and a profusion of dark rocks at locale V.

Fig. 7 shows the locale signatures plotted along their first two principal components together with a third option utilizing color histograms. The simple color features fare worst; the relationship between color pixels from different locales has little to do with the geologic ground truth. This inaccuracy is highlighted by the first three locales, which are similar geologically but widely separated in the space of geologic signatures. This suggests that our color histogram features are poor correlates of geologic type. The distribution of colors is influenced by many factors — such as features on the horizon or lighting changes — that have little to do with the locale's geology. Any subtle differences in pixel counts caused by different concentrations of rocks are overwhelmed by these other factors.

Detecting rocks in images provides better geologic signatures. The PCA plot of manual categories suggests a linear gradient of change between locales I and V. The ground truth geologic change is sudden, however, so this signature is still slightly misleading. The error is probably due to the bias from an overly-rigid model. Much of the variation between locales occurs within the "other rocks" category and is thus invisible to the classification. Rock detection with unsupervised clustering yields the best result, matching the human interpretation that locales I - III are similar while locales IV and V are each different from all others.

3.2. High-Frequency Data Collection

The second experiment tests an alternative data collection strategy that captures single images at high frequency during an extended traverse. Unlike the meticulous locale-based data collection the rover is in motion during the entire procedure. This means that data can be collected more quickly but dramatically reduces the sample size for characterizing each locale's geology.

Here the rover traveled 100 meters across an open plain and into a field of mid-sized (10-50cm) rocks. The pan-tilt unit was fixed straight ahead at a -15 degree inclination. The resulting images showed the far-field in front of the rover with enough horizon to provide some visual context. Images were captured every 2 meters. The cameras were not synchronized during this experiment so stereo geometry data was unavailable; rock detection relied on pixel intensities alone.

Figure 8 shows the result of detecting rocks in each image and clustering them into three unsupervised categories. To reduce detection noise in the single-image samples a smoothing operation averaged class counts among all images in a 2-neighbor radius. Again, clusters varied most along the intensity attribute. The bar graph provides rock counts for each class in each of the traverse images. Note that the count is very small in the open plain and rises rapidly as the rover enters the rock field near image 20. The data suggest that a well-localized rover using this technique should be able to autonomously place these sharp boundaries to within 10 meters.

The decrease in the rock count around image 34 is not due to a change in geology, but simply in rover heading; for a moment the cameras pointed toward the well-lit side of the rocks which made them difficult to distinguish from the background sediment. Synchronizing cameras for stereo should give the detector more data and help to alleviate this problem. Nevertheless, the error underscores the challenge of using complex features for geologic profiling; the detection step can introduce additional error. This problem is more obvious in the high-frequency sampling case where there are fewer images available to determine the signature of each locale.

4. CONCLUSIONS

This paper describes a method to characterize a locale's geology by detecting and classifying features in rover images. While our "simple feature" profiling results in ambiguous signatures, the complex features generated by rock detection are accurate enough both for finding region boundaries and for providing basic summaries of local geology. Our comparison of data collection policies suggests that both concentrated and high-frequency sampling are useful for different purposes. Concentrated data collection accumulates large sample sizes for more accurate feature counts. High-frequency sampling offers

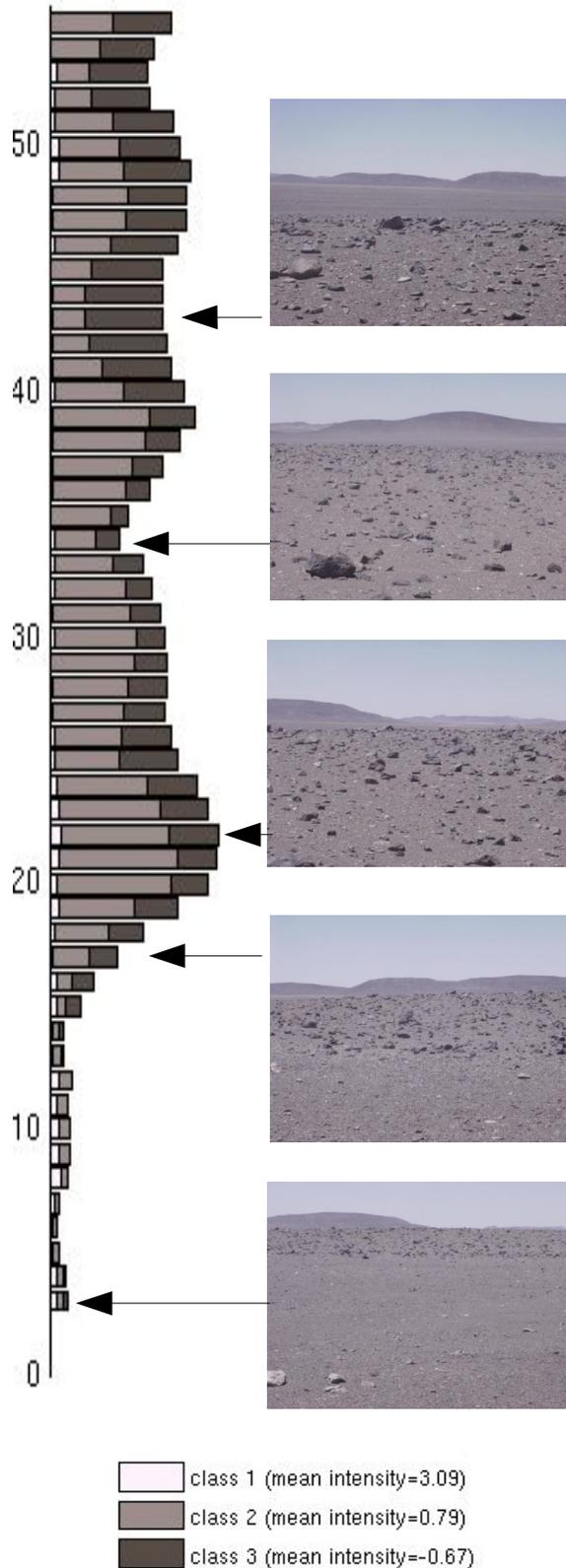


Figure 8. Periodic sampling for data collection “en route.” Images are spaced at 2 meter intervals during an autonomous traverse into a dense rock field. The transition is visible in feature class signatures by image 15.

expedient summaries with higher spatial resolution. Geologic signatures will become more reliable as rock detection and feature extraction continue to improve.

The approach we advocate here makes several simplifying assumptions that future work might address. Spatial smoothing aside we have presented discrete locales as independent of each other. In fact the geologic class of each location is correlated with other nearby locales. One could glean additional accuracy from a higher-level description like a Markov Random Field that represents probabilistic relationships between different locales. Another shortcoming of our geologic signature profiling is that locales are compared in an unsupervised fashion. Regions are grouped purely on the basis of feature counts without any consideration of what the resulting geologic regions *mean* or the processes that generate the features in the first place. This unsupervised description is advantageous in some respects — the first field experiment shows the difficulty of translating human understanding into numbers that generalize. However, unsupervised profiles forgo the descriptive power of generative models that would permit reasoning about the causal relationships between the geology of the terrain and the rocks that appear.

Geologic signatures would benefit from data sources apart from camera imagery. In the long term introducing additional data sources like spectroscopy and microscopic imaging will provide additional information for more complete models. The system would also profit from orbital data. Scientists engaged in mission planning could schedule traverses to investigate specific boundaries that are visible from orbit. During unsupervised modeling, satellite data provides additional information to characterize locales. Finally, orbital imagery could corroborate rover-based geologic signatures during data analysis.

As rovers travel longer distances they will offer new opportunities for planetary science, but also new challenges in the form of resource bottlenecks. Autonomous science answers these challenges with better selectivity from the rover and better data analysis on the ground. Testing in field scenarios will continue to be an important part of understanding and validating these new technologies.

ACKNOWLEDGMENTS

Thanks are due to Scott Niekum for his work developing algorithms for image segmentation. Dominic Jonak, Mike Wagner, Chris Williams, and Joseph Flowers provided indispensable help in data collection and analysis. We would also like to thank Nathalie Cabrol for her ideas and insight. This research was supported by NASA under grants NNG0-4GB66G and NAG5-12890.

REFERENCES

- [1] R. Castaño, R. C. Anderson, T. Estlin, D. DeCoste, F. Fisher, D. Gaines, D. Mazzoni, and M. Judd, "Rover Traverse Science for Increased Mission Science Return," *Proceedings of the IEEE Aerospace Conference*, 2003.
- [2] D. R. Thompson, S. Niekum, T. Smith and D. Wettergreen, "Automatic Detection and Classification of Geological Features of Interest," *Proceedings of the IEEE Aerospace Conference*, March, 1995.
- [3] A. Castaño, R. C. Anderson, R. Castaño, T. Estlin, and M. Judd, "Intensity-based rock detection for acquiring onboard rover science," *Lunar and Planetary Science*, 35, 2004.
- [4] T. Smith, S. Niekum, D. R. Thompson and D. Wettergreen, "Concepts for Science Autonomy during Robotic Traverse and Survey," *Proceedings of the IEEE Aerospace Conference*, March, 1995.
- [5] R. Castaño, M. Judd, R. C. Anderson, T. Estlin, "Machine Learning Challenges in Mars Rover Traverse Science," *International Conference on Machine Learning*, Washington, D.C., August 2003.
- [6] M. desJardins, K. L. Wagstaff. "DD-Pref: A Language for Expressing Preferences Over Sets," *Proceedings of the Twentieth National Conference on Artificial Intelligence*, July 2005.
- [7] R. Castano, M. Judd, T. Estlin, R. C. Anderson, D. Gaines, A. Castano, B. Bornstein, T. Stough, and K. Wagstaff. "Current Results from a Rover Science Data Analysis System," *Proceedings of the IEEE Aerospace Conference*, March 2005.
- [8] B. V. Funt, G. D. Finlayson, "Color Constant Color Indexing," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 17:5, May 1995, pp. 522—529.
- [9] M. J. Swain and D. H. Ballard, "Color Indexing," *International Journal of Computer Vision*, 7:1, 1991, pp. 11—32.
- [10] L. Pedersen, *Robotic Rock Classification and Autonomous Exploration*, PhD thesis, Robotics Institute, Carnegie Mellon University, CMU-RI-TR-01-14.
- [11] V. Gor, R. Castaño, R. Manduchi, R. C. Anderson, and E. Mjolsness, "Autonomous rock detection for Mars terrain," *Proceedings of AIAA Space 2001*, Albuquerque, August 2000.
- [12] M. D. Wagner, D. Apostolopoulos, K. Shillcutt, B. Shamah, R. G. Simmons, W. Whittaker, "The Science Autonomy System of the Nomad Robot," *ICRA 2001*, 2, pp. 1742—1749.
- [13] J. Fox, R. Castaño, R. C. Anderson, "Onboard autonomous rock shape analysis for Mars rovers," *Proceedings of the IEEE Aerospace Conf.*, Big Sky, Montana, 2002.
- [14] V. C. Gulick, R. L. Morris, M. A. Ruzon, and T. L. Roush, "Autonomous image analysis during the 1999 Marsrokhod rover field test," *J. Geophysical Research*, 106, No. E4, 2001, pp. 7745—7764.
- [15] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference* (San Francisco: Morgan Kaufmann), 1988.
- [16] C. M. Bishop, *Neural Networks for Pattern Recognition* (Oxford: Oxford U. P.), 1995.
- [17] A. P. Dempster, N. M. Laird, D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society*, B, 39, pp. 1—38.
- [18] D. Wettergreen, N. Cabrol, S. Heys, D. Jonak, D. Pane, M. Smith, J. Teza, P. Tompkins, D. Villa, C. Williams, M. Wagner, A. Waggoner, S. Weinstein, W. Whittaker. "Second Experiments in the Robotic Investigation of Life in the Atacama Desert of Chile," *ISAIRAS* 8, 2005.
- [19] M. Wagner, S. Heys, D. Wettergreen, G. Kantor, J. Teza, D. Apostolopoulos, W. Whittaker, "Design and Control of a Passively Steered, Dual Axle Vehicle," *ISAIRAS* 8, 2005.