Automatic Detection and Classification of Features of Geologic Interest

David Thompson, Scott Niekum, Trey Smith, and David Wettergreen
The Robotics Institute
Carnegie Mellon University
5000 Forbes Ave.
Pittsburgh, PA 15213
{drt,sniekum,trey,dsw}@ri.cmu.edu

Abstract—The volume of data that planetary rovers and their instrument payloads can produce will continue to outpace available deep space communication bandwidth. Future exploration rovers will require science autonomy systems that interpret collected data in order to selectively compress observations, summarize results, and respond to new discoveries. We present a method that uses a probabilistic fusion of data from multiple sensor sources for onboard segmentation, detection and classification of geological properties. Field experiments performed in the Atacama desert in Chile show the system's performance versus ground truth on the specific problem of automatic rock identification.

TABLE OF CONTENTS

1 INTRODUCTION
2 ROCK DETECTION AND CLASSIFICATION
3 EXPERIMENTAL METHOD
4 RESULTS
5 CONCLUSIONS
6 ACKNOWLEDGMENTS

1. INTRODUCTION

"Science autonomy" broadly describes a robotic vehicle's ability to analyze the scientific content of its observations and use this information to take intelligent actions. This onboard science data understanding is important for efficient allocation of a rover's time and bandwidth resources. Advances in rover navigation are increasing traverse range at a rate much faster than the increase in communications bandwidth [2]. Thus, much of the terrain the rover observes on a long traverse might never be examined by scientists. Onboard data understanding ensures that the rover reports the most interesting and valuable data. Our research objective is to develop a coherent framework for science autonomy that supports science during rover traverse and survey. In the scenario we are developing, the rover identifies significant observations from initial data collection and decides whether to perform follow-up measurements or instead continue with other tasks. After collecting data, the rover can then determine how much of its limited communications bandwidth to devote to transmitting each data product. Science data understanding also allows the rover to convey some important information (such as statistical summaries of terrain features) in a compact form [12]. Achieving these goals requires two key functionalities: the ability to identify targets of scientific significance and the capacity to plan and modify behavior based on science information. This paper addresses the first property.

Our initial research has focused on analyzing rocks in desert environments. Rocks have characteristics that make them an informative test case for autonomous science. They are rich science targets that contain useful information across a wide range of sensing modes. Autonomous rock analysis is easy to evaluate because the targets are discrete units that have measurable physical properties and occupy precise locations. Rocks are plentiful in terrestrial test environments and directly relevant to remote planetary geology. Unfortunately, rocks—like most features of geological interest—are difficult to find and classify with *a priori* designer knowledge. While one would like to identify features by specifying their physical properties, field conditions with sensor noise, variable lighting, and ambiguous terrain make these properties difficult to extract.

We believe a machine learning approach—where science targets are specified in terms of *examples* rather than properties—will yield reliable detection and classification performance. This paper begins with a detailed explanation of such a system. Then we discuss its performance with data obtained during a rover field experiment in the Atacama Desert of Chile. Detection performance is compared both to ground truth and to human performance on the same images. We find a strategy employing supervised learning for feature detection and both supervised and unsupervised learning for feature classification to be effective for flexible and reliable onboard data understanding.

0-7803-8870-4/05/$20.00/© 2005 IEEE
IEEEAC paper # 1251

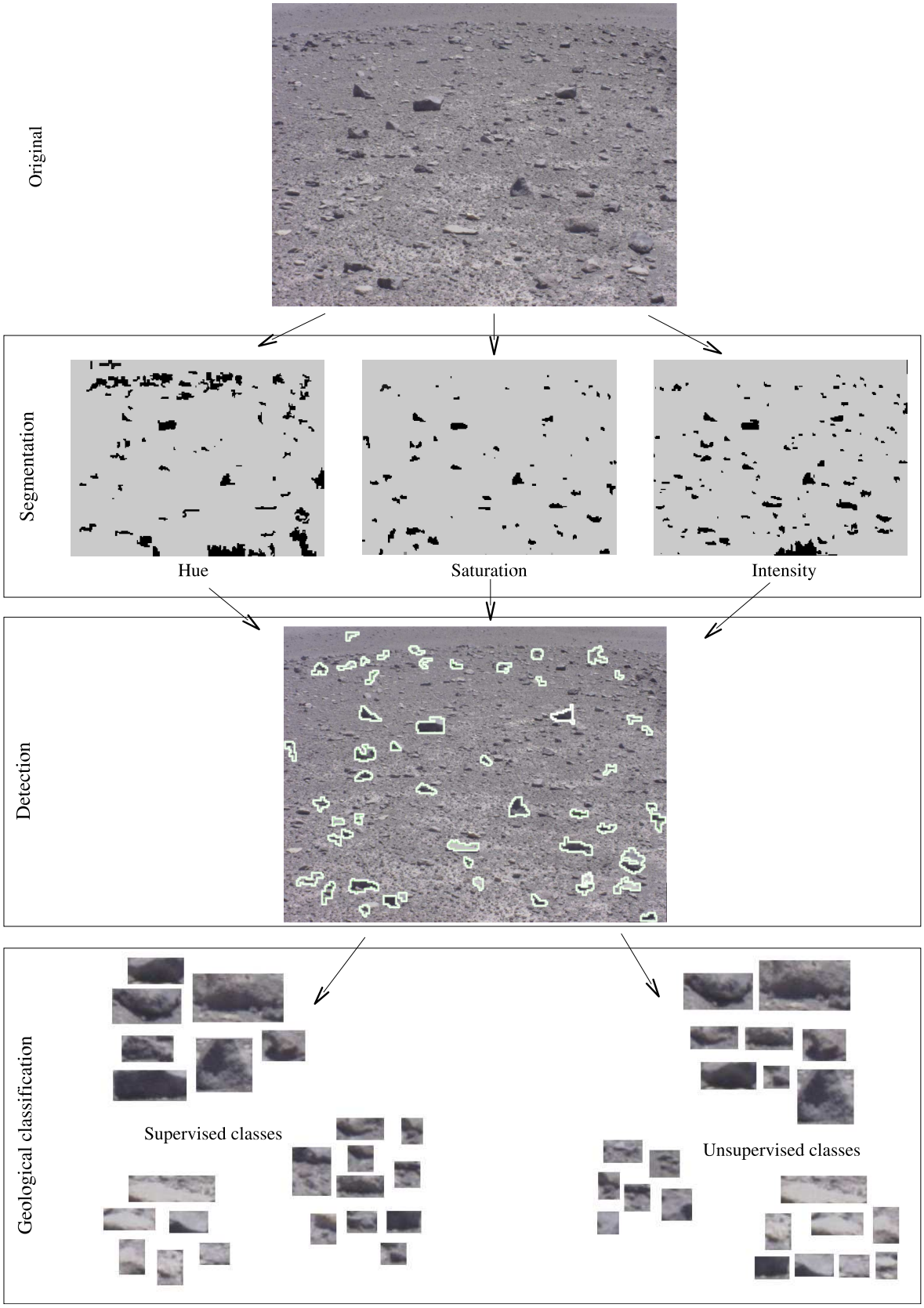

Figure 1. Procedure for detection and analysis of science targets. A belief network detects science targets out of candidates segmented from the original image. Then supervised and autonomous classification schemes categorize the geological features.

2. Rock Detection and Classification

Like many of the terrain features geologists find interesting, rocks have no single distinguishing signature for the standard sensors available to explorer robots. Researchers have tackled this difficult pattern recognition problem with a wide range of approaches. One strategy uses stereo images to find rocks based on their height above the ground plane [6]. This depth information yields precise detections—any sufficiently tall object is almost certainly a rock. Unfortunately stereo often misses small or distant objects. A second method reduces rock detection to the simpler problem of finding shadows. Given the relative position of the sun and the camera, shadows suggest the location of rocks [8]. A third strategy detects rocks directly in the image. This category includes the intensity-based method of Castaño *et al* that looks for closed shapes at varying resolutions with an edge detector [3]. The Nomad explorer in Antarctica is another example—it used color cues to segment meteorites from background ice [11]. Finally, one can compensate for the shortcomings of any one approach by using several simultaneously; Gor *et al* use stereo for detecting large rocks and image intensity to detect small rocks [7]. These diverse strategies rely on universal characteristics that distinguish the science targets from their backgrounds. However, unstructured environments and uncertain sensor data may demand distinctions too subtle to specify in advance. Our intuitive notion of a "rock" describes a range of observable phenomena that individual attributes fail to capture. Rocks can be darker or lighter than the base sediment. They can be rough or smooth, small or large, protruding or embedded in the surface. The interactions between these attributes—which together distinguish a science target from its background—are difficult to capture *a priori* .

Given this difficulty we approach detection not as the unconstrained task of designing general detectors for each geological feature but rather as a traditional classification problem to be solved through supervised learning techniques. Examples chosen by an expert teach the science autonomy system attribute relationships too subtle and specific for a human designer to exploit. In our method the pre-labeled examples train a belief network [9] that considers a probabilistic fusion of information from stereo, shape, shading, color, and texture sources. The analysis of a scene proceeds through three steps (Figure 1). The first step, *segmentation*, isolates potential science targets from the rover's stream of sensor data. The second step, *detection*, analyzes these candidates with a belief network that distinguishes target from non-target regions. The input to the belief network is an attribute vector that contains color, texture, shape and stereo disparity information. The output is a probability that each region is a science target. Finally the *geological classification* step classifies the geological features according to their observable attributes [2], [6].

Segmentation

The segmentation step processes sensor data to identify homogeneous regions in the rover's environment corresponding to possible science targets. The segmentation stage is not responsible for detecting targets but merely suggests candidate structures for further classification. This separation between segmentation and detection permits the science autonomy system to be quite permissive in the kinds of data it accepts as input. Any additional criteria, even from multiple measurement sources, can add candidate regions for consideration. The different sensors are processed separately to yield several segmentations of the same scene (Figure 1). A feature which is indistinguishable from the background in one channel might be clearly visible in another; as long as it is found in at least one segmentation the result should be recognized later in the detection stage. This approach is similar to the "symmetric" detection scheme used by Gor *et al* [7] in that it builds multiple object maps for each image. It is distinct, however, in that we do not require the different segmentations to be orthogonal—science targets may well appear in several segmentations at once. Overlaps not eliminated by the detection stage are in practice fairly easy to resolve. Moreover, removing the orthogonality constraint increases the designer's flexibility in choosing segmentation algorithms and ultimately helps the rover's chances of finding each terrain feature.

Our rock segmentation, detection and classification system begins with color images that segmentation splits into homogeneous pixel regions. First, each color channel from an image is preprocessed with a simple Gaussian blur operation. A region-merging algorithm shatters the original image into a grid of 5×5 squares and iteratively joins them back into regions of uniform properties (Figure 2). Their minimum 5×5 size is a reasonable compromise between a desire for high resolution segmentations and the requirement that each square be large enough to produce meaningful statistics about its interior pixel values. At each iteration we calculate the mean pixel values of all regions and merge neighbors whose means fall within a certain threshold of each other. This process repeats until no more merges occur. Finally we exclude regions whose sizes fall outside a given size window. The experiments that follow use a permissive window—every region between 20 and 500 pixels is a potential rock.

We will see in the next section how rock classification exploits stereo information for later processing stages; stereo disparity offers another possible segmentation channel. Although segmentation of 3D data was not implemented for the following experiments, k-means clustering has been successfully employed to find rock shapes in stereo images [6]. These techniques generally involve fitting a ground plane to the pixels and segmenting the resulting height map.

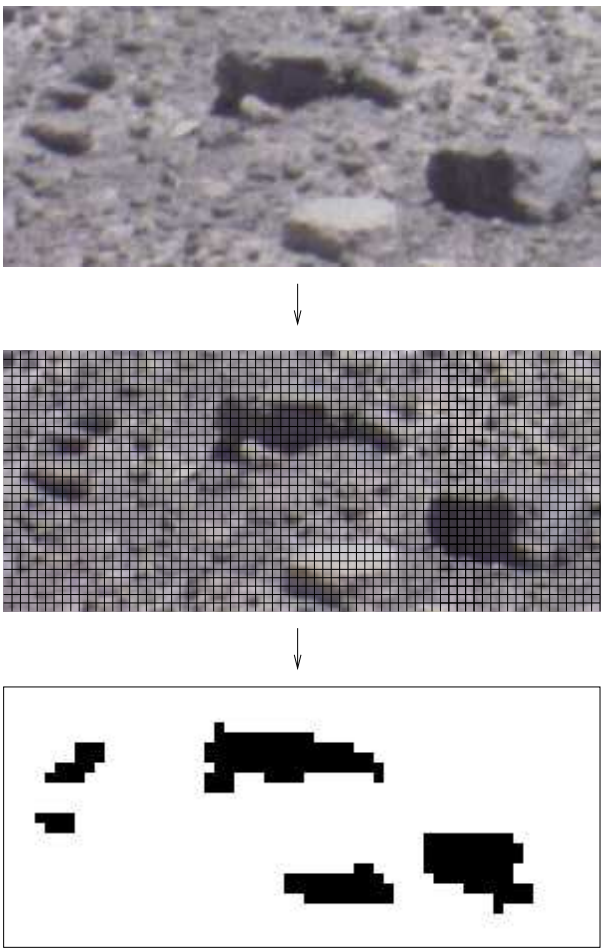

Figure 2. Segmentation of a sample image containing several rocks.

Detection

The detection stage identifies science targets from among the segmented regions by extracting a real-valued attribute vector from each candidate and labeling it with a Bayesian belief network [9]. For our application belief networks offer several advantages over other classifier techniques. First, they solve the problem of missing data that can occur when fusing multiple sensors with varying fields of view. In the case of absent data, such as a stereo mismatch or an undeployed sensor, the network provides posterior probabilities that are appropriate for the available information. Thus there is no penalty for training the network with time-consuming, seldom-used instruments like spectrometers. Another advantage to the Bayesian approach is that the explicit classification probabilities offer a principled way to tune the precision of the network. This accommodates the various performance standards that different science target detection applications require. For example, a first pass detector might use a lenient probability threshold to avoid missing any desirable science targets. After locating a high-value candidate for further inspection, the rover might deploy additional sensors and verify its decision using a stricter standard. By tuning a single value—the probability threshold for considering a target – an operator can elicit both high-precision and high-recall behaviors from a single network. A last advantage of belief networks is their computational efficiency relative to non-parametric classifier schemes like k-nearest-neighbor [1].

For a given attribute vector $X = \{x_1, x_2, ...x_n\}$, the network computes the probability $P(C|X)$ that the region is of class C. A naive Bayes approach, where all attributes are considered to be independent, yields the following decomposition:

$$P(C|X) = \alpha P(C)P(x_1|C)P(x_2|C)...P(x_n|C)$$

where $P(C)$ is the prior probability of the class and α is a normalizing constant. This formulation corresponds to a simple Bayes network with arcs running from the class node to each independent attribute. Although the naive Bayes strategy offers the advantage of simplicity, its conditional independence assumption is violated when there are correlations between some of the candidates' attributes. Because of this our implementation favors a more fully-connected Bayes network. The additional arcs in such a network represent relationships of conditional dependence to be reflected in the final probability calculation. For example, if x_1 were conditioned upon the value of x_2,

$$P(C|X) = P(C)P(x_1|C, x_2)P(x_2|C)...P(x_n|C)$$

Techniques such as belief propagation can calculate this joint probability even in cases where one or more attributes are unknown [9].

There are many ways one might represent the conditional probability distributions. For the rock detection task we quantize each attribute and use tables to count the number of times associated values appear. This is a non-parametric representation, requiring more training data than a more constrained parametric distribution function [1]. Nevertheless, the tabular distributions are more flexible and yield superior performance for the detection tasks we have tried. Smoothing techniques can also be used to fill out tabular probability distributions; Parzen density estimation facilitated autonomous meteorite classification for the Nomad antarctic explorer [10]. In the next section we describe geological classification techniques which exploit the same probability relationships but use Gaussian probability density functions that can be set with far fewer examples.

We extract attributes for each region from the interior pixels and the pixels immediately surrounding the image. An example appears in Figure 3 - note that the region's border appears somewhat aliased because it is composed from original 5×5 pixel blocks merged during segmentation. The attribute set extracted from each image region contains the following measurements:

Perimeter: The ratio of the region's squared perimeter to its

pixel area. Non-rock artifacts often have long, spidery shapes while rocks tend to be more convex and ellipsoidal.

Relative Color: The absolute value of the difference in mean pixel hue, saturation, and intensity between the interior of the region and the context region. Fisher distance (which weights the score according to the variance of each sample) performed better as a difference metric than a simple comparison of means.

Relative Color Variance: The absolute value of the difference between the pixel variance of the interior region and the pixel variance of the context region. This functions as a simple measure of texture, helping to detect situations like a smooth rock sitting in rough, high-variance gravel or a pitted rock on a smooth background.

Height Above the Ground Plane: While small rocks are generally below the noise threshold for our stereo system, height is a valuable attribute for detecting large rocks and excluding large non-rock regions.

Texture: We use a fractal dimension measure [4] of a binary intensity map to describe the detail of each region as resolution is increased. The result is an efficiently-computed value that corresponds somewhat to our intuitive notion of surface roughness. Like color variance, this helps to detect rocks in those cases where their roughness differs substantially from the background sediment.

Intensity Gradient: Rocks are three-dimensional protrusions that exhibit shading when illuminated by sunlight. We use least-squares regression to find the magnitude of the intensity gradient over the pixel surface of each region in the image, giving the overall strength of its shading.

Physical Location: While not implemented in the current field test, future versions will include an estimate of the rock's position in the world. A region's location should not affect its classification as a rock, but it should play into the autonomous geological classification—a familiar looking rock can still be geologically interesting if it is found in an unexpected place.

Absolute Color: Another attribute that informs geological classification. We exclude it from the attribute set for rock/non-rock classification in order to maintain generality.

Two additional attributes do not directly affect a region's chances of being a science target or its geological classification. Nevertheless, we include them due to their strong conditional dependence relationships with the other attributes. By considering these dependencies the Bayes network computes a more accurate posterior probability:

Absolute Range: Many of the differences between rocks and non-rocks become less apparent as range increases. In particular, texture is more difficult to see even when there are many pixels representing the distant region.

Pixel Area: There are varying degrees of conditional dependence between most observed attributes and the regions' pixel area. These dependencies are due to the way regions are represented as a finite number of pixel "samples." For example, normalized perimeter is generally small for regions of small pixel area because rough borders become less noticeable when the number of pixels used to describe them decreases. Similarly, texture is hard to recognize with few pixel samples.

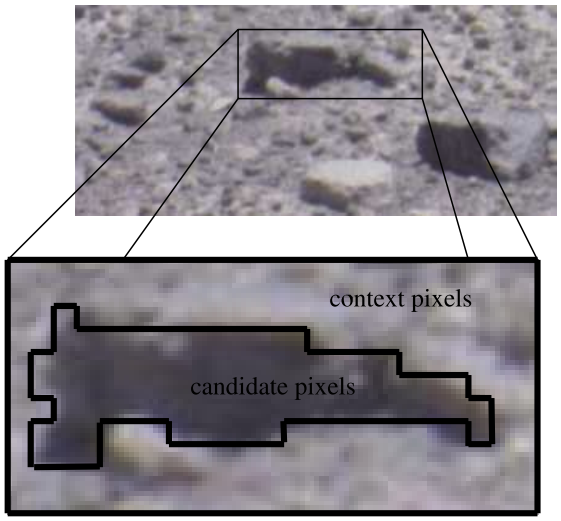

Figure 3. Attributes are extracted from the region and its immediate neighborhood.

The topologies of the networks used by the rock detector for detection and geological classification appear in Figure 4. In the detection step the belief network uses the attribute vector to calculate the probability of a region's belonging to each of five classes: rocks, uniform patches of soil, sky, shadows cast on the ground, and a final "everything else" class for regions that do not fit neatly into one of the other categories. The "everything else" class contains ambiguous candidates like a region comprising a small part of a larger rock or one that includes both a rock and a small patch of sediment. While we constrain our system to seeing just these five kinds of objects in the world, one could expand the set to include other kinds of geological features. The number and breadth of classes is limited only by their degree of separation in attribute space and the amount of training data available.

Geological Classification

After the detection stage discovers a science target the two geological classification modules classify it according to both predefined and synthesized categories. Here our classification method is similar to the one first suggested by Castaño *et al.* [2]. Predefined geological types are chosen at the outset by

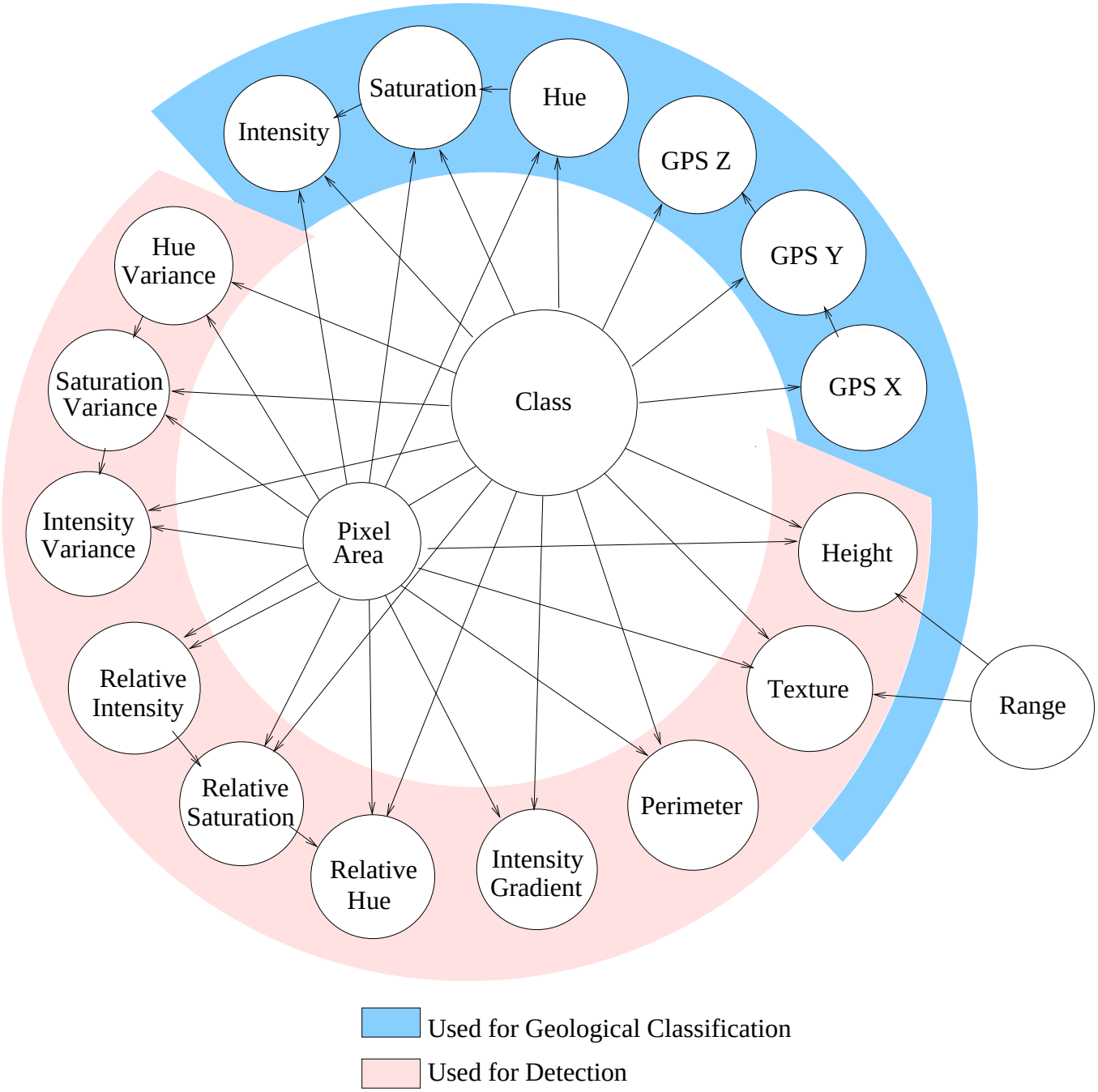

Figure 4. Bayes network topologies for rock detection and analysis. Texture, color, shape, shading and stereo data are all probabilistically related to the candidate region's class. The candidate's range and pixel area do not depend on its class but still affect the interpretation of its other attributes.

an expert. Autonomous classes, on the other hand, come from an unsupervised clustering algorithm that runs whenever the rover collects new data. Multiple classifications give operators a rich palette of options for defining the rover's behavior. For instance, they could instruct the rover to focus only on novel outliers that have a low probability of being generated by any autonomous cluster. This way the rover could find the most interesting science targets in areas that operators had not yet seen. Alternatively, operators could instruct the rover to examine a single representative example of each geological type that it sees. A third alternative would be to search only for a specific predefined geological type and to ignore everything else. Operators can choose the mix of these behaviors that is most appropriate for the rover's mission.

Geological classification differs in several ways from the initial detection step discussed above. One difference is the input attribute vector—because geological classification is only interested in properties of the science target itself, relative measures of the region's properties versus its background are ignored. Another difference is the representation of the conditional probability densities; geological classification models classes utilize Gaussian probability density functions. This permits the use of the EM algorithm [5] to find efficient maximum-likelihood estimates during autonomous clustering. Moreover, Gaussian PDFs are more effective with sparse training data. They allow both the rover and remote experts to create a geological class with just a few examples.

We use a cross-validation procedure to choose the best number of autonomous geological classes that explain a given dataset. Before field operations, a geologist estimates the appropriate number of clusters i in terms of its own separate probability density, $P(M_i(X))$, where $M_i(X)$ is a model that fits a list of observed targets X to i distinct classes. If only a handful of geological classes are interesting, this probability density might favor low single-digit values of clusters. If there are many different classes of interest or the operator wishes the rover to make fine distinctions between autonomous classes, the expected number of clusters could be much higher. During a traverse, the rover divides its set of observations into a training set T and a validation set V. Then it performs expectation-maximization clustering on the training half using various numbers of clusters, yielding models $M_1(T), M_2(T), ...M_n(T)$. The prior probability of each number of clusters, together with the likelihood of the validation half given that model, describes the comparative probabilities that each number of clusters generated the observed data.

$$P(M_i(X)) = P(M_i)P(V|M_i(T))$$

After calculating the best number of clusters, the algorithm clusters the data one last time and re-labels any previously-observed rocks according to their new class membership probabilities. There exist other non-probabilistic methods for classification and clustering, but the explicit uncertainty in the classification can be used to help the scientists calibrate their reliance on the results or to inform onboard planners. Moreover, low probability densities indicate the interesting outliers that have little in common with other observations.

3. Experimental Method

In this section we examine a field test performed in the Atacama desert in Chile. The Atacama is a common place for testing exploration robots because of its resemblance to Mars. In particular, its central arid region is practically devoid of macroscopic life. These factors make it an appropriate place to evaluate the rock detection system. The following experiment was designed to test the utility of the described rock detection strategy in the uncertain lighting and terrain conditions found in the field. In particular we hoped to evaluate the rock detector against a human's own identifications given the same image data.

The rover platform is "Zoë," an exploration robot built at the Carnegie Mellon Robotics Institute (Figure 5). Zoë is a solar powered robot developed for long-duration autonomous traverse. The sensor utilized for the experiment was a 30 centimeter baseline stereo pair of CCD cameras mounted on pan-tilt actuator at 2 meters height. These cameras captured 1280×960 full color images. Their $21.1° \times 15.9°$ field of view provides an angular resolution similar to the human eye and the Pancams used in the Mars Exploration Rover missions.

Figure 5. Zoë in the Atacama Desert of Chile.

A hill strewn with rocks of various sizes was chosen for the test. Data collection occurred at two locations: a training set captured at the base of the hill and a test set captured at its peak. At each site the rover used its stereo imaging suite to capture partial panoramas comprising 180 degrees of azimuth and 40 degrees of elevation. The segmentation outputs from the first panorama were labeled manually and used to train the region classifier. Then, the fully-trained classification algorithm detected rocks in the second panorama.

The detection procedure for a typical test image appears in

Figure 1. This scene contains rocks of various sizes, albedos and textures. Some of the smallest segmented regions are ignored because they fall underneath the minimum size threshold for our analysis. The detection stage processes the remaining regions and classifies some of them as rocks.

We evaluated these autonomous detections against the locations of real rocks in the image set. Because manually identifying every rock would be infeasible and error-prone, we labeled a random sample of 50 rocks from each image. Software for manual rock selection displayed each test image and prompted the user with a random location in the scene. The user would draw a bounding box around the true rock closest to the random prompt point. This defined not just the real science targets but also "no-rock regions," circular areas centered on the prompt point within which there were no science targets (Figure 6). Thus the labeling operation resulted in rock, no-rock, and uncoded areas. In accordance with our detector's size window, we restricted the our definition of a rock to those that had a short axis longer than 20 pixels. Not only are the smallest rocks less informative science targets, but they are often ambiguous even to a human observer.

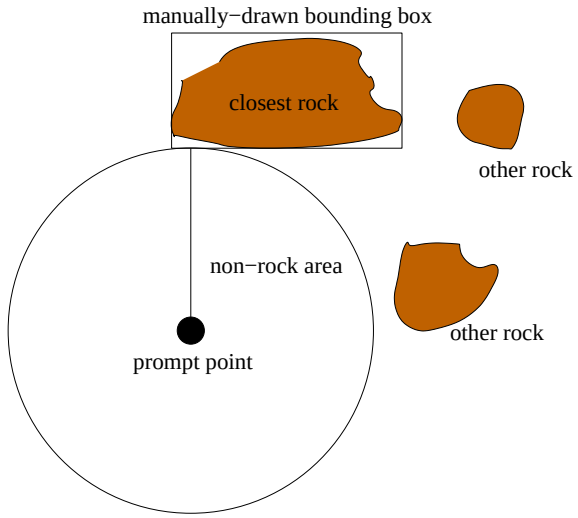

Figure 6. Method for coding test images.

We employed two different evaluation methods to test different aspects of the autonomous detection (Figure 7). The first method evaluates the detector's performance at locating rocks in the scene, an ability that facilitates tasks like targeted sensor deployment, rock distribution analysis and selective data return. To calculate this performance score we checked that the center of each detected bounding box intersected a manually-drawn true rock region. If the center fell within the bounding box of a true rock that was not yet accounted for the detection was a true positive. We labeled regions with centers falling in non-rock regions false positives, and discarded regions whose centers fell outside coded areas. A stricter evaluation method tested the detector's performance at localizing the exact outline of the targets. This ability is important for attribute extraction and geological classification. Here we checked the bounding boxes of autonomously-detected regions against the manually-drawn bounding boxes by considering the ratio of their intersection to their union. An intersection area that comprised 50% of the union area was sufficient to consider the detected region well-localized (Figure 7).

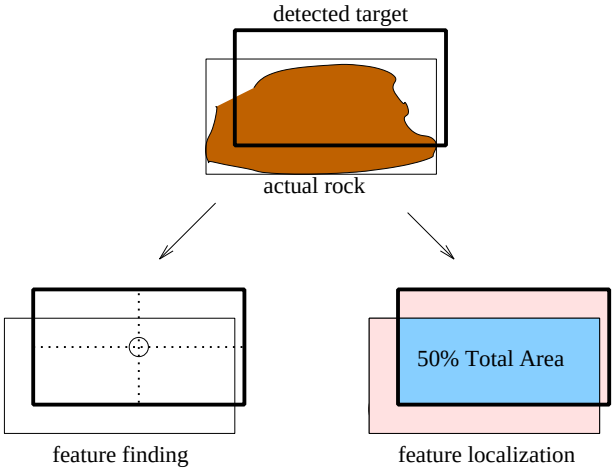

Figure 7. Lenient and strict evaluation strategies.

The 30 images in the test set together contained a total of 1078 coded rocks that varied in size, color, range, and illumination. We evaluated the detector on the entire test set (which contained image regions over 20×20 pixels in size) and the subset containing only large features (regions over 75×75 pixels in size, of which there were 68 in the data set). Figure 8 shows typical near- and far-field images with templates that illustrate these sizes.

4. Results

Figures 9—12 illustrate the detector's performance on the test data. Varying the prior probability for the "rock" class results in different performance characteristics. *Precision*, the fraction of detected targets in coded regions that correspond to actual rocks, is in general inversely related to *recall*, the fraction of actual targets that are found. Figure 9 illustrates detection precision for varying choices of the class prior. As the prior probability increases a greater number of uncertain regions are classed as science targets and precision drops. Figure 10 shows the recall rate for each prior probability. Note that the detector never found all of the science targets, even for the case where the prior was set to 1. The reason for this is that the evaluation measures not simply region classification but rather the overall system's ability to find and recognize rocks in the original images. A rock that does not appear in the segmentation stage has no chance of being classified later. Figure 11 illustrates the superior recall performance for science targets over 75 pixels in size. Precision was also improved for large regions; in every median case 100% of the detected regions corresponded to real rocks.

20x20 pixels
(minimum size considered)

75x75 pixels
(large feature)

Figure 8. Typical near- and far-field images from the test data set.

Finally, Figure 12 illustrates the fraction of detected targets that were well-localized given different choices for the prior.

While we did not attempt a quantitative evaluation of geological classification, both supervised and autonomous classification showed reasonable results for the tested images. The example of Figure 1 shows a supervised classification based on albedo. Three predefined categories were generated by grouping rocks of like albedos from a different image. Then the classifier applied these user-defined categories to classify the rocks in the test case. Representative samples of each class are shown in the diagram. Note that high-, medium-, and low-albedo rocks are accurately classified. Curiously, the autonomous geological classification clusters the test set into similar categories. We believe this is due to the limitations of the attribute set used for geological classification. There was little color or texture data in the test scenes; this resulted in classifications that favored size and albedo characteristics.

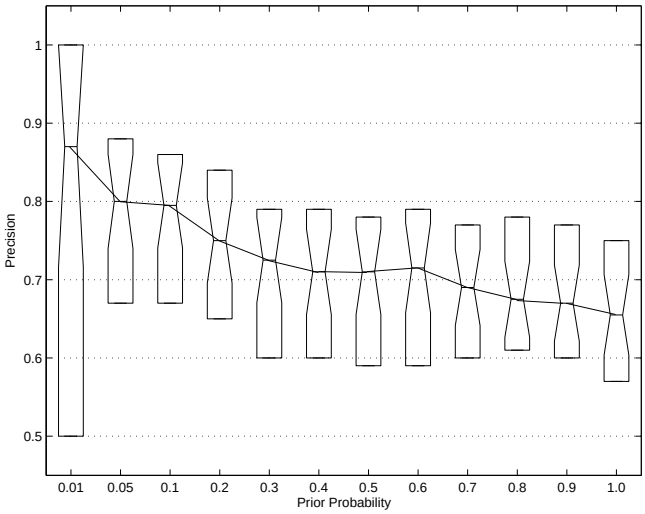

Figure 9. Detection precision—the fraction of detected targets corresponding to actual rocks. Median rates appear along with boxes illustrating the extent of second and third quartiles. Notches in the boxes show intervals of 95% confidence for the median.

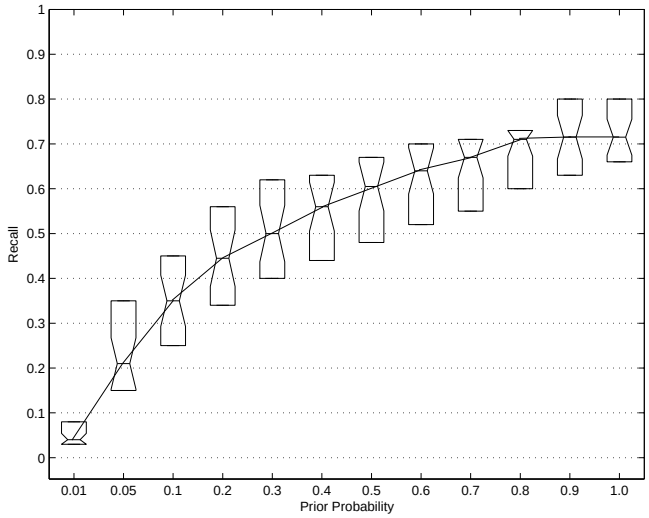

Figure 10. Detection recall—the fraction of manually-labeled rocks that were found by the detection algorithm.

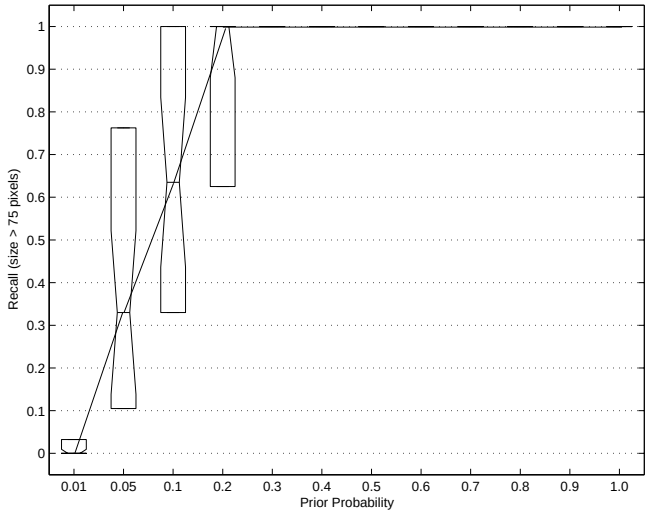

Figure 11. Detection recall for the subset of science targets that were over 75 pixels in size.

5. Conclusions

This experiment constitutes a preliminary test of the presented algorithms for automatic detection and classification of geological features of interest. A comprehensive evaluation will involve a larger data set representing diverse geological regions. Nevertheless, this first test is sufficient to suggest some initial conclusions. It reinforces the notion that while detection and localization performance are related it is possible to get the former without the latter. The evaluated system detects rocks well, making it a suitable candidate for tasks like locating science targets and pointing instruments. However, poor localization performance suggests that its geological classifications will contain some inaccuracy. These classifications might still be useful for recognizing outliers or performing general representative sampling, but precise small-scale maps of rock distributions would require improvements in the segmentation portion of the algorithm.

Both detection and localization performance vary with respect to the pixel size of the region under consideration. As a result, small rocks are noticeably more difficult to detect and classify than large ones. Accuracy also drops for regions in the extreme far-field. This suggests one could gain a substantial accuracy boost by considering only the science targets above a certain pixel size. It also implies that active perception may play an important role in detecting science targets. After locating a possible target, the rover could zoom

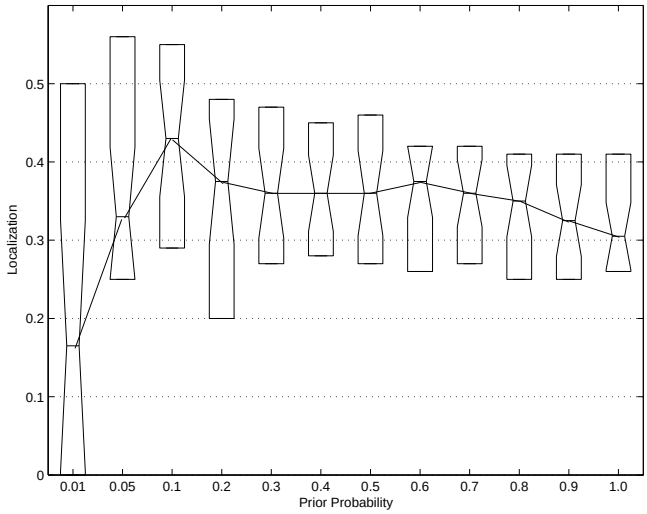

Figure 12. Localization performance—the fraction of accurately-detected science targets that meet the localization criteria of Figure 7.

in a camera, target a spectrometer or try some other means of gathering more information and increasing detection accuracy.

We believe the "detection as classification" strategy holds unique advantages for finding science targets. It can draw on multiple segmentation methods and fuse information from multiple sensors when available. It can adapt to different terrain by exploiting attribute relationships specific to a new training set. Finally, it expresses its final decision through an explicit probability value. While there are predictable scenarios for which this example performs suboptimally, results suggest that the method as a whole is sound and that performance will improve as refinements are made to its component parts. Nevertheless it is unlikely in the near term that this —or any other detection technique—would approach 100% recall for all science targets of interest. Thus, quantitative evaluaiton will remain an important aspect of future science autonomy research. Through regular evaluation we can track progress while maintaining reasonable expectations for the system's field performance.

There are several avenues for further improvement in the example system presented here. Accurate segmentation of stereo disparity data is one obvious way to improve localization accuracy. Another important advance will be the inclusion of other science targets into the detection and classification scheme. While this paper focuses solely on rocks, the system itself should generalize to other features of scientific interest like soil patches, gravel beds, and salt deposits. This should be verified experimentally. Similarly, the attribute set for geological classification should be expanded to include more sophisticated measures of angularity and texture. Finally we hope to integrate the science autonomy system discussed in this paper with a planner that will use this data to plan experiments and command the rover. Together these modules would constitute a comprehensive science autonomy system that not only detects likely targets but also plans appropriate experiments "on the fly." Such a system will be important to exploit the extended traverse range of future exploration robots.

6. Acknowledgments

We thank those who assisted in data collection during the Atacama field experiment: Michael Wagner, Dominic Jonak, Chris Williams, and Allan Lüders. Many thanks are also due for the aid and expertise of Kim Warren-Rhodes and Nathalie Cabrol. This research was supported by NASA under grants NNG0-4GB66G and NAG5-12890 with program executive Michael Meyer.

References

[1] C. M. Bishop, *Neural Networks for Pattern Recognition* (Oxford: Oxford U. P.), 1995.

[2] 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 Conf.*, Big Sky, Montana, 2003.

[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] B. B. Chaudhuri, N. Sarkar, "Texture segmentation using fractal dimension," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2:1, 72—77, January 1995.

[5] 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, 1-38.

[6] 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.

[7] 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.

[8] 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, 7745–7764, 2001.

[9] J. Pearl, *Probabilistic reasoning in intelligent systems: networks of plausible inference* (San Francisco: Morgan Kaufmann), 1988.

[10] L. Pedersen, *Robotic Rock Classification and Autonomous Exploration*, PhD thesis, Robotics Institute, Carnegie Mellon University, CMU-RI-TR-01-14.

[11] 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, 1742—1749.

[12] K. L. Wagstaff, R. Castaño, S. Dolinar, M. Klimesh, R. Mukai, "Science-based region-of-interest image compression," *Lunar and Planetary Science*, 35, 2004.

David Thompson *is a Ph.D student at the Carnegie Mellon University Robotics Institute. His research interests include machine learning and its applications in exploration robotics. He received a B.A. in computer science from Carleton College in 2001 and an M.Sc. in informatics from the University of Edinburgh in 2002.*

Scott Niekum *is an undergraduate Computer Science and Cognitive Science student at Carnegie Mellon University. His research interests include computer vision, machine learning, parallel probabilistic problem solving, and efficient knowledge representations. He expects to graduate in May 2005 and hopes to continue on to a Computer Science Ph.D program.*

Trey Smith *is a Ph.D student at the Carnegie Mellon University Robotics Institute, studying under a GSRP Fellowship with NASA Ames Research Center. His research interests include probabilistic planning, rover navigation and science operations, and multi-robot systems. He received B.S. degrees in computer science and math from Carnegie Mellon in 1999.*

David Wettergreen *is an Associate Research Professor in the Robotics Institute at Carnegie Mellon University. His research in space robotics advances capabilities for autonomous navigation and exploration. Enabling robots to reason about science objectives and to modify behavior based on science observations is a current focus. Dr. Wettergreen is the Principle Investigator for the ASTEP Science on the Fly project.*