Essential Autonomous Science Inference on Rovers (EASIR)[1,2]

Ted L. Roush
NASA Ames Research Center
MS 245-3
Moffett Field, CA 94035-1000
650-604-3526
troush@mail.arc.nasa.gov
Mark Shipman, Raytheon Inc.
Robert Morris, SETI Inst.
Paul Gazis, San Jose State University Foundation
Liam Pedersen, QSS, Inc.

Abstract—Existing constraints on time, computational, and communication resources associated with Mars rover missions suggest on-board science evaluation of sensor data can contribute to decreasing human-directed operational planning, optimizing returned science data volumes, and recognition of unique or novel data. All of which act to increase the scientific return from a mission. Many different levels of science autonomy exist and each impacts the data collected and returned by, and activities of, rovers. Several computational algorithms, designed to recognize objects of interest to geologists and biologists, are discussed. The algorithms represent various functions that producing scientific opinions and several scenarios illustrate how the opinions can be used.

TABLE OF CONTENTS

1. INTRODUCTION

NASA's Office of Space Science is addressing several fundamental questions about our solar system and life. How did the universe begin and evolve? How did we get here? Where are we going? Are we alone?

One significant discovery about Mars stands out above all others: the possible presence of liquid water on Mars, either in its ancient past or preserved in the subsurface today. Water is key to supporting life, because almost everywhere we find water on Earth, we find life. If Mars once had liquid water, or still does today, it's compelling to ask whether life forms could have developed on its surface and if so, then does any evidence of it's presence remain? More provocative, if so, could any of these tiny living creatures still exist today?

To discover the possibilities for life on Mars--past, present or our own in the future-- NASA's Mars Program has developed an exploration strategy entitled *Follow the Water*. This effort begins with understanding the current environment on Mars. Observed features like dry riverbeds, ice in the polar caps, and rock types that only form when water is present must be explored. If ancient Mars once held a vast ocean in the northern hemisphere, as some scientists believe, then how did Mars transition from a more watery environment to the dry and dusty climate it has today?

To pursue these goals, all of NASA's future missions to Mars will be driven by rigorous scientific questions that will continuously evolve as new discoveries are made. On-going and future missions will continue to provide abundant morphological imaging, global compositional, and global topographic information. All these data allow scientists to develop hypotheses regarding the presence of certain materials and address mechanisms associated with their genetic and evolutionary history.

As illustrated on the left side of Figure 1. Historical mission operations have involved returned scientific data, their scientific evaluation, scientific recommendations for future mission activity, and development and relaying of commands to the vehicle. This traditional cycle of data evaluation and commands is not amenable to rapid long-range traverses, discovery of novelty, or rapid response to any unanticipated situations. Future Mars missions are envisioned to include rovers that carry imaging devices to

[2] IEEEAC paper #1328, Version 1, Updated October 9, 2003

characterize the surface morphology, and a variety of analytical instruments intended to evaluate the chemical and mineralogical nature of the environment(s) that they encounter. In addition to issues of response time, the nature of imaging and/or spectroscopic devices is such that tremendous data volumes can be acquired that simply can not be relayed to Earth within the time constraints imposed by communication opportunities.

The computational resources available on board current and planned rovers are quite limited. For example, the 2003 Mars Exploration Rover (MER) missions currently rely upon CPU's operating at ≤20MHz and have memory storage capabilities of approximately one hundred megabytes [1,2]. If these resource restrictions remain unchanged, then any effort to enable scientific evaluation of data acquired on-board these platforms must address these severe computational and memory limitations.

A large obstacle to achieving the scientific goals on Mars is the enormous distance that separates it from our own planet. The effectiveness of rovers is limited by the fact that communications are problematic over such large distances. Transmissions occur at extremely low data rates and may take over 20 minutes to travel one way. Another challenge is the harsh Martian environment quickly degrades rover hardware (the Sojourner rover lasted roughly three months), so time is essential once the rover is deployed.

The combination of limited hardware resources, communication time delays, and data transmission capabilities suggest crucial decisions regarding data analysis be made on-board these robotic explorers that require automating scientific analysis and discovery based upon data gathered by sensors. In order to address communication bottlenecks and reduce the dependence on ground based control there have been efforts in recent years to develop technologies that will enable rovers to act more autonomously. A rover employing autonomy technologies would be able to respond to high-level commands representing complex sequences of actions that may involve sensory feedback from the environment (e.g., test the hypothesis that this was/is an aqueous environment, or obtain spectra of the five largest rocks in the vicinity). This translates into fewer transmissions required to perform a given task and, ultimately, more mission time devoted to science and less to rover operation.

Specifics of the Mars 2009 mission remain to be more completely defined [3]. However, the mission scenario still includes a long-range, long-duration rover implying a dynamic situation with limited communication abilities during traverses. The fundamental philosophical change needing to occur is illustrated in Figure 1 where, on the right, a science-enabled mission is shown. As indicated in Figure 1, a science-enabled rover does not eliminate input from the scientists; as new hypotheses or high-level directions are required.

It is not surprising that the planetary science community has been slow to embrace autonomy as a necessary component of rover missions. There are significant risks in trusting mission-critical decisions to a machine. Applying autonomy technologies indiscriminately is likely not realistic as some tasks are good candidates for automation, while others are better left to the judgement of human experts. Successful application of autonomy technologies will depend upon striking the proper balance between risks introduced by such technologies and the likely benefits of using them.

So far, most automation efforts have included tasks leading up to data acquisition such as navigation, health and safety monitoring, and autonomous sensor placement. There has been only modest research into the potential for autonomous interpretation of sensor data to enhance the scientific value of a mission [e.g. 4,5,6,7]. The on-going projects at NASA's Ames Research Center were initiated to address this gap in autonomy technology research.

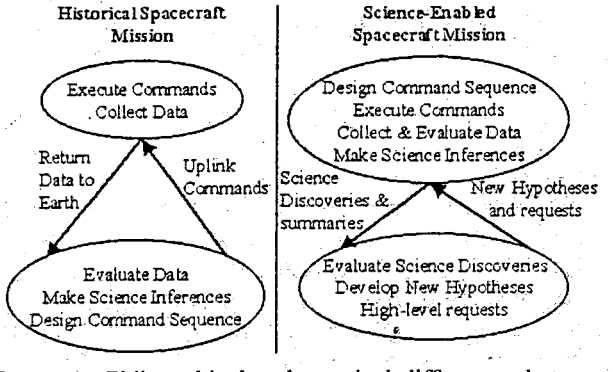

Figure 1. Philosophical and practical differences between historical (left) and Science-enabled (right) rover mission scenarios.

There are many levels of science autonomy spanning the range from assuring quality science data is collected to summarizing scientific content of sensor data [7]. Below we describe these various levels and illustrate their impact on future rover activities and data returned to scientists on Earth.

A basic level of autonomy, not described by [7], mimics the same questions addressed in well-controlled laboratory studies. Is the instrument working properly? Is the appropriate sample being measured? What actions can improve the accuracy of the measurement? All of these are intended to assure good quality science data is obtained and can impact the data collection strategy and hence, operational sequence planning and execution. Such considerations may also drive instrumental design to provide the rover with the ability to monitor actual voltages or resistances within an instrument. However, once the data is obtained, no further action is taken to change the data volume or priority for return to Earth.

The intermediate level of autonomy involves evaluating the science content of the sensor data; perhaps correlating or combining data from two, or more, different sensors. Based upon this evaluation, data might be reprioritized

(data providing evidence for aqueous activity has highest priority), or selectively compressed by differing amounts for transmission to Earth (return only average spectra of broad regions encountered). Regardless of these evaluations, the rover does not alter its planned activities, but the data returned may be different than originally intended, or in the extreme, never returned at all.

This intermediate level of autonomy illustrates how the issue of science autonomy can be viewed as a compression problem. The relative benefits and risks of techniques in science autonomy can be compared with those in the larger field of compression. For example, maximize the amount of information returned from a rover at a site, having a finite data collection time, during which it receives no guidance. The scientific yield of information is considered a constraint in addition to the typical resource constraints of command cycle frequency, data volume, and total mission life [8].

At higher levels of autonomy, evaluation of sensor data yields results that directly influence planned rover activities or provide terse summaries of sensor data. For example, if sensor data indicate the presence of a predefined important object (e.g. layers, water, or a fossil), then the rover may halt planned activities and await directions from Earth or alternatively, halt and obtain other sensor data for that object. If the rover continues to encounter the same materials during its traverse from point A to point B, then it may provide a very terse summary containing a few representative data rather than return the full complement of data collected; saving valuable data volume. In this scenario all rover activities can be affected. The rover may alter its planned activity, selectively eliminate data for return, and/or significantly compress data.

In general, the goal in the development of autonomous science data interpretation is to enhance the scientific return of Mars rover missions. Expanding on the analogy above, it should be possible to autonomously monitor navigation imagery as a rover traverses from point to point, and to recognize certain signature features that may be related to water or life (e.g., sedimentary layering). If any are encountered during a traverse, the rover may decide to stop and alert the ground control team. This is merely one of many possible situations in which so-called "science autonomy" may prove invaluable.

Current directions of our research include identification of generic contour patterns in images that are likely to indicate scientifically interesting geologic features. For example, parallel contours in an image may be indicative of layering. Many geologic features of interest can be fairly well represented by line drawings as illustrated in section 2.

2. IMAGE ANALYSIS

Images can provide tremendous amount of information regarding morphological shapes of interest to both geologists and biologists. In 2D, gray-scale images of natural landscapes, there are a number of important geological features that have similar general characteristics. Subsequently, some important geologic features can be grouped into basic classes having an associated semantic label. Dendritic structures (constructive and erosional fluvial activity, biologic activity, mineral morphology, Fig. 2A); parallelograms (mineral morphology, structural patterns, Fig. 2B); elongated shapes (mineral and fossil morphology, fig. 2C); circular or elliptical patterns (impacts, volcanoes, biologic organisms, mineral morphology, Fig 2D); stellar patterns (constructive fluvial activity, mineral and fossil morphology, Fig. 2E); and concentric radial patterns (evaporite deposits, lava flows, Fig. 2F). Knowing the context in which the image was acquired, e.g. orbital versus microscopic, allows the best selection from among the various alternatives. To this end we have been implementing various automated image analysis algorithms to recognize a few key morphologies.

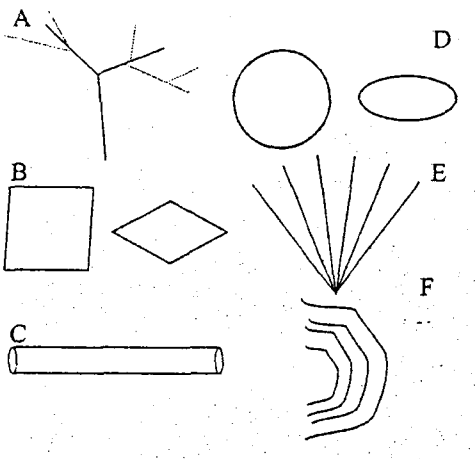

Figure 2. Various morphologic shapes of interest to geology and biology. (A) dendritic patterns; (B) parallograms; (C) elongated shapes; (D) circles and ellipses; (E) stellar patterns; and F) concentric radial patterns.

2.1 Edge Detector

Many of our image algorithms rely upon recognizing boundaries, and hence edges. We evaluated Sobel [9] and Canny [10] edge detectors. The following methods require the edge response at each pixel and the slope (orientation) of the edge. Because of its ability to localize edges with sub-pixel accuracy, and the additional capability of readily linking edges, the Canny edge detector was preferred over the Sobel method.

2.2 Layer Detection

The objective of the layer detection algorithm is to partition the rectangular image lattice (row, column) into "layered" and "non-layered" regions, producing a binary image representing these two regions. This is accomplished by searching a set of connected edges produced by the Canny

algorithm [10] for spatial groupings of approximately parallel and approximately straight edge segments [11]. For each point on the lattice, the first step is to examine all L-length edge segments in the surrounding NxN window, where L and N are user-defined parameters. If an edge segment is straight enough, its orientation is estimated from its endpoints. A histogram of segment orientations is maintained for each window along with a tally on the total number of approximately straight segments in the window. Next, a dominant orientation is computed for each window from its histogram along with a measure of dominance. Finally, a decision rule, based on the orientation dominance and density of edge segments in the surrounding window, is applied to each pixel.

To summarize, the three main steps leading to the final partition are: 1) generate statistics including the number of edge segments in each window and the distribution of their orientations; 2) determine the dominant orientation in each window and the associated degree of dominance; and 3) for each pixel, apply a decision rule to obtain the final labeling. Figure 3 illustrates the application of this algorithm to images containing and not containing layers.

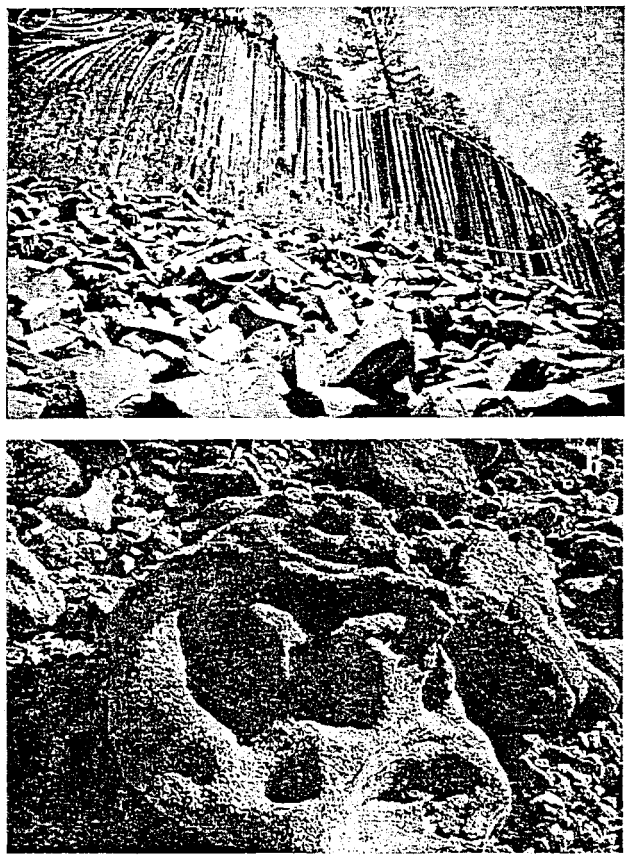

Figure 3. Results of layer detection algorithm. a) An image exhibiting obvious parallel structures b) An image lacking any such structures (b).

Good correspondence between layered regions identified by this algorithm and layered regions identified by a geologist depends upon the validity of the following assumptions: 1) If the edge set was derived from an image, then the layers must be resolved in the image and must be of sufficiently high contrast that an edge or line detection procedure can extract them; 2) edges representing layer boundaries are approximately parallel (i.e. perspective doesn't make them appear to diverge significantly); and 3) there are no parallel and linear structures in the scene caused by objects or phenomena other than geological layers.

2.3 Horizon Detection

The algorithm implemented to locate the apparent horizon in a grayscale image consists of the following: 1) determining a feasible region in an image to perform a search for the apparent horizon 2) a heuristic search method; and 3) evaluating candidate solutions reached by the search. The methods used were devised in an effort to minimize computational complexity while ensuring robustness.

If information describing the camera's orientation and field of view is available, then it is used to determine the feasible region where the apparent horizon is expected. All image points above the geometric horizon (i.e. the plane tangent to the planet's surface at the point of observation) define the feasible region. If only points below the geometric horizon have been imaged, then the assumption is that the apparent horizon has not been imaged. Otherwise, the feasible region is searched for.

The search method used for the apparent horizon belongs to a class of image analysis algorithms known as Active Contours. An initial estimate of the horizon's location in the image is made (usually the uppermost row) and then deformed over time. A physical analogy is used here to describe the deformation algorithm. Points (pixels) along the active contour are treated as a sequence of particles connected by inelastic strings that move in response to a force field. The motion of each particle is restricted to discrete downward steps in the vertical direction (i.e. along a column). The force field comprises two components: a uniform downward force analogous to gravity, and an upward buoyant force that is a function of local image edge intensity and direction. In addition, a particle's neighbors may exert forces on it if it moves to the end of either the left or right connecting string.

Initially the downward force is set to zero. Then it is gradually increased until the contour begins to move. The contour is allowed to propagate until it reaches an equilibrium, at which point it is considered as a potential candidate and a confidence value is computed and the potential candidate is either approved or rejected. Currently confidence is defined as a heuristic function of contour intensity and smoothness. The algorithm terminates when the active contour reaches the lower boundary of the feasible image region. Approved candidate solutions are stored in a list for retrieval.

So-called Snakes [12,13] are another type of active contour. To evaluate how well our algorithm performed relative to this alternative, we implemented a version of our algorithm that uses the efficient snakes of Williams and Shah [12] in place of our active contour. The snake-based version was less accurate and took minutes to complete the search as opposed to a few seconds using our method. Figure 4 illustrates the application of our algorithm to images where a horizon is present and absent.

Strong assumptions are: 1) if the horizon is visible in the image, then each image column contains a single horizon pixel; 2) camera roll is small (<20°). If camera roll is large, then the image should be rotated before processing so that the effective roll is small; 3) the horizon's slope predefined upper limit (currently 45°) for significant portions of its length. Weak assumptions are: 1) there are no salient horizontal contours above the horizon in the image; and 2) the horizon is the most salient contour connecting the left and right image boundaries, where contour salience is approximately a function of edge strength and smoothness. If occasional errors are acceptable, then the weak assumptions may be violated. If strong assumptions are violated, poor performance should be expected.

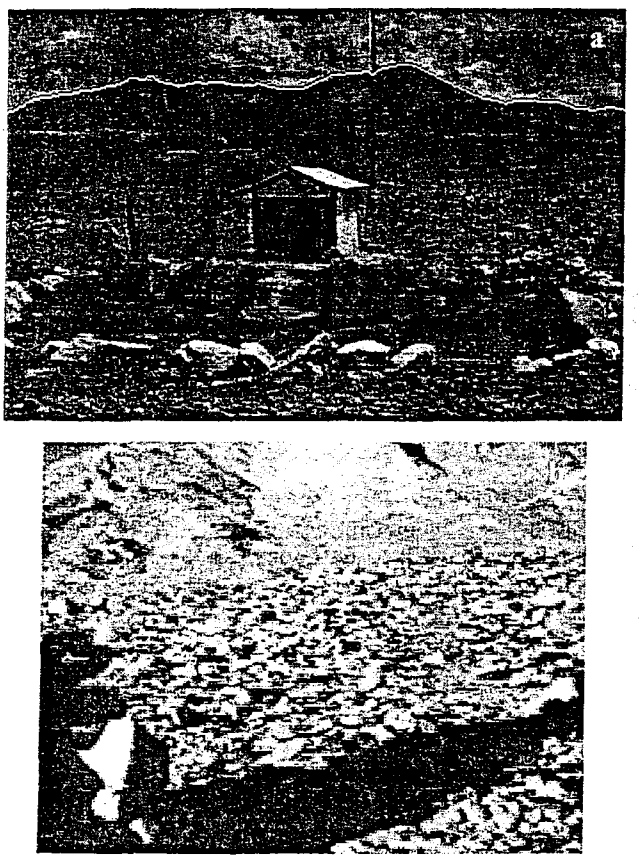

Figure 4. Results of horizon detection algorithm. The apparent horizon has been delineated in image (a), while use of camera pointing information enabled the algorithm to determine that the horizon was not visible in image (b).

2.4 Terminator Detection

In astronomy a terminator is the dividing line between the illuminated and the unilluminated part of the moon or a planet's disk. Here, we extend this definition to include curves dividing illuminated and unilluminated portions of any convex object. In scenes of Mars-like terrain obtained by rovers, such objects are usually rocks, so the algorithm can be used to infer the likely positions and, to some degree, the sizes of rocks in such scenes. The terminator detection algorithm attempts to quickly identify terminators on convex objects in a scene illuminated by a distant point source [11]. Although not pursued here, the algorithm could be used to look for craters by readily changing the sign and looking for concave shapes.

In addition to a grayscale image, input to the algorithm consists of camera orientation and position (planetary latitude and longitude) and sun position. Using these, the expected orientation (in the image) of the terminator on a spherical object is calculated. The mean orientation of edge points along the terminator of a spherical rock is given by the orthogonal projection onto the image plane of a vector directed toward the sun (Figure 5).

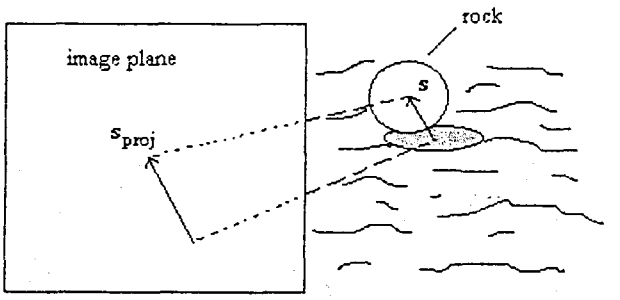

Figure 5. If s is a vector pointing toward the sun, then its orthogonal projection onto the image plane gives the expected orientation of terminators on spherical objects in the image.

An object's orientation will, in general, affect the orientation of its terminator in the image. A spherical object model is used because all orientations are equivalent. Once the expected terminator orientation has been calculated, Canny's edge operator [10] is used to find candidate edge points in the image. The strength of each candidate is then attenuated according to the deviation of its orientation from the expected terminator orientation. The hysteresis step of the Canny algorithm [10] is performed to extract connected edges from the set of edge point candidates. Edge segments which are at least some user defined threshold length and which have roughly the expected orientation are considered valid terminator candidates.

Finally, a confidence measure is assigned to each candidate. A heuristic function of intensity means and variances of image regions on either side of the candidate boundary is used to calculate a confidence measure and candidates whose associated confidences are below a

predefined threshold are eliminated. Although this method of assigning confidence is naive, results from field test images indicate good detection.

For a particular application, such as finding rocks, results can be improved by using prior knowledge about the image in order to limit the search to a feasible region. Edge points outside the feasible region are eliminated from consideration. For example, the horizon detector results discussed above have been used to limit the search space in past efforts to detect rocks [11].

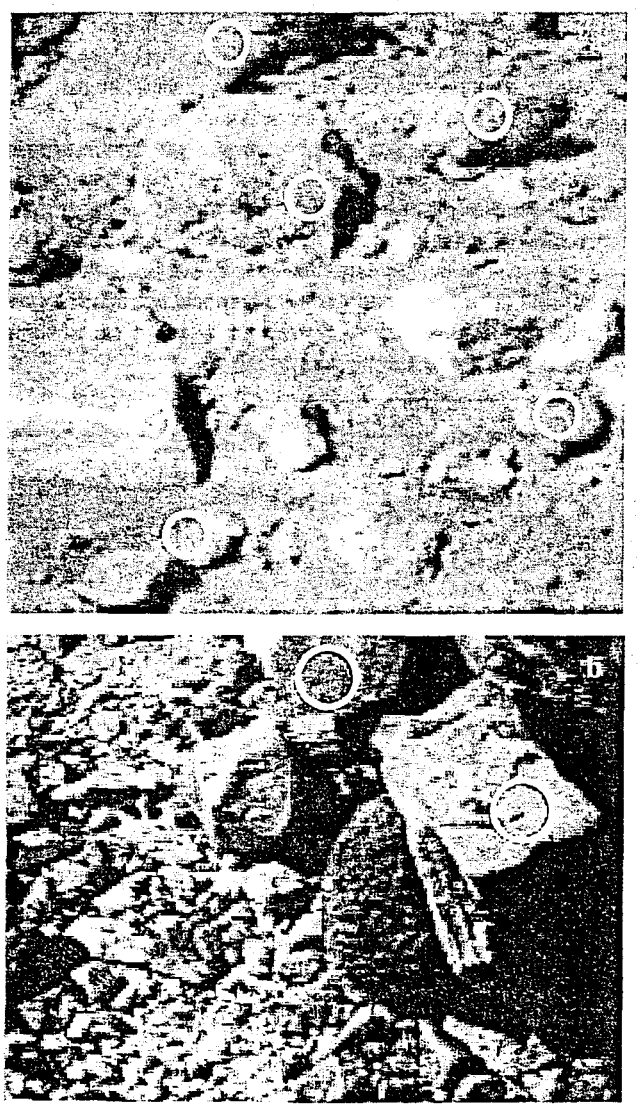

Figure 6. Autonomous selection of spectrometer targets based on terminator detection in images from the Sojourner (a) and Marsokhod (b) rovers. Objects having terminators with length greater than the diameter of the spectrometer's field of view were considered good targets.

Strong assumptions are: 1) illumination can be approximated by a point source at infinity. If not, then the terminators we are searching for may not exist. Some degree of mutual illumination between objects in the scene is acceptable as long as it is negligible compared to the direct illumination from the light source; 2) the terminator on each object in the scene is visible and coincides with a detectable edge in the image; and 3) the field of view of the camera is narrow enough that weak perspective projection is an adequate model for image formation. Weak assumptions are: 1) surfaces in the scene are approximately Lambertian; 2) dramatic changes in image brightness along the expected terminator orientation are caused by shadows; and 3) objects in the scene are approximately spherical. Figure 6 illustrates the application of our algorithm to selection of potential spectrometer targets from extraterrestrial and terrestrial rover image.

If reliable performance is required, images should satisfy all of the assumptions. If occasional errors are acceptable, then the weak assumptions may be violated. Two major failings of this approach have been observed: 1) the inability to reliably distinguish dark objects from shadows due to the difficulty in detecting the edge between a dark rock and its shadow; and 2) the tendency to reject heavily textured objects due to the assumption that variance on either side of an object's terminator is small.

2.5 Line

The stellar pattern detector and the parallelogram detector described below employ the results of line detection. The Hough Transform for line detection [14] was used. Representing the lines in polar coordinates, they are ranked according to total response; allowing partially occluded structures to be accounted for. A predefined number of lines having the highest total response are selected for further analysis in the various detectors described below. In some cases, the total response of a line is normalized by the line length. In order to allow nearly straight lines to be included, the edge map is blurred or the parameter space can be decimated.

2.6 Parallelogram (see Figure 2B)

Every pair of parallel lines P is compared with every other distinct pair of parallel lines R that are not parallel to P. Using these four lines, four intersections are computed. Then a box is traced from one intersection to the next and compared with the edge map. If enough of this box has an edge response that is above some predefined threshold, then the tracing is labeled as a parallelogram. In order to deal with affine transformed parallelograms, the image must either be rectified [15] or a slack term is used to allow the parallel criteria to be approximate.

2.7 Stellar and Radial Pattern

We define a stellar pattern as a shape having appendages that can be decomposed into a set of elongated patterns, or appendages, that converge in a center. The polar lines from the Hough transform [14] are faintly "drawn" across a blank image. Intersections are pixels upon which multiple

distinct lines have been drawn. The intersections that result from the most lines will have the highest intensity.
In cases of objects with wide arms, the intersections created through the edge map will not correspond to the intersections created by the arms of the stellar pattern. Performance, in this case, is improved by decimating the (Hough) parameter space. A decimated parameter space will map neighboring intersections to the same parameter vector. Decimating the parameter space combines neighboring intersections. The thicker the appendages, the greater the decimation needed. Performance, can be enhanced by preprocessing the image, for example segmenting the original image then thinning.

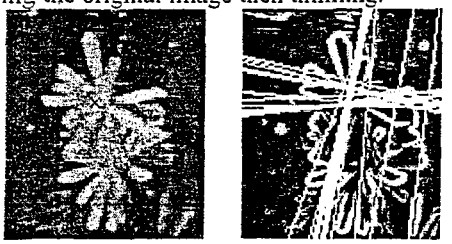

Figure 7. An image of a stellar pattern with the detected center (left), and the 40 strongest lines detected drawn across the edge map (right). The lines are drawn across the entire image to account for occlusion.

2.8 Ellipse (Figure 2D)

The Hough transform is also used. The parameter space is defined using minor axis, major axis, center x, center y, and rotation of the ellipse in radians. Shapes that are nearly ellipses also return high responses if we *collapse* the rotation of the ellipse and scan for a small range of rotations.
Since an ellipse with major axis j and minor axis n, rotated by pi/2 is equivalent to an ellipse with a major axis of n and a minor axis of j, the search space is truncated to include only ellipses with minor axis greater than or equal to its major axis.

3. Spectral Analysis

In geology, mineral identification is key to classification of rock types and hence, in addressing what geological processes have been active within a given locale. Distinct mineral identification typically requires analytical laboratory techniques. Here we rely upon remote spectral observations as a proxy for the types of triage analyses that might be used in rover missions to help select what targets to further investigate using more resource intensive tools. We require a label for a spectrum that indicates a specific mineral presence or absence. In our case, given a spectrum, the result is classified as a carbonate or not.

3.1 Bayesian Method

Simple implementations of Bayes Belief Networks [16, hereafter BBNs], such as the Naï ve Bayes Net (NBN), commonly use a two-tiered model. The NBN in particular uses one response variable with independent variables as child nodes [16]; an approach that reduces the number of samples needed to train the network. Even so, the NBN requires a training set that represents the minerals in the proportions in which they occur in the evaluation environment.
In a mineral identification problem, even this restraint still requires an inordinate amount of data because of the large number of minerals. By using a *hidden (dummy)* variable of *mineral class* (see Fig. 8), with a probability distribution defined by a scientist, only the conditional probabilities for each mineral given its class need to be computed.
A mineral class, e.g. carbonate, is simply a group of minerals, e.g. calcite, dolomite, *et cetra*. The likelihood of a mineral, given its class, is either assumed to be the same for all minerals in that class or defined explicitly. The mineral class need not correspond to groupings of geologic significance. It is only important that it be a partition of all minerals defined by the mineral node. However, classes that do correspond to groupings of geologic significance facilitate the assigning of prior probabilities of each class and the collection of the necessary spectra to create the training data set.
As illustrated in Figure 8, The network used has a three-tiered topology. The response node is defined over the space of chosen mineral classes and has the *mineral* variable as a child node. There is an *other* value for the mineral class variable to try to account for unexpected minerals. In practice however, trying to account for the unexpected is futile.
The mineral variable is defined over the space of selected minerals. This mineral variable has *features* as child variables, the set of which is called a *feature vector*. Each feature is a summary statistic derived from the spectrum.

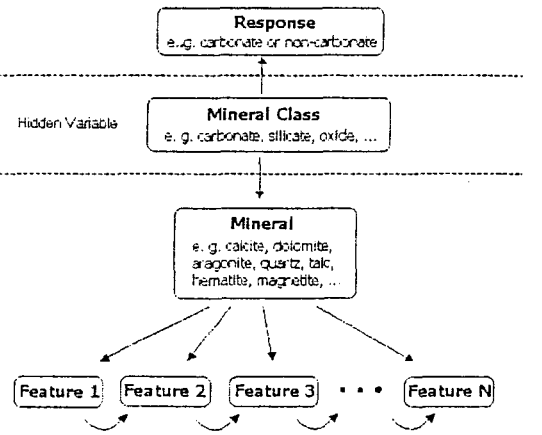

Figure 8. Illustration of the three-tiered approach where features are allowed to be dependent upon adjacent features.

For carbonate detection, carbonate is simply one of the mineral classes and if it is the most likely class for a given spectrum, that spectrum is labeled carbonate.

A scientist must choose the relevant mineral types, partition them into mineral classes and set the prior probability that an arbitary mineral belongs to each of these classes. Mixtures of minerals giving rise to spectral features not exhibited by their component minerals should be separated into unique types to account for the emergent property.

The feature vector can be any set of summary statistics derived from the spectrum. The features derived here are a result of correlating chosen templates with mineral absorption bands. Several spectral ranges that contain features are believed to be helpful in identifying minerals. Since spectral analysis often involves looking for maxima and minima, a Gaussian shaped template is correlated with the spectrum to extract peaks and troughs. The coefficient is then normalized to prevent dominance of features. This coefficient is computed as:

$$\lim_{\sigma \to \infty} \frac{C_{ST}(\sigma + 1)}{\sigma_T \sqrt{\sigma_T^2 + \sigma^2}}$$

where C_{ST} is the correlation coefficient of a spectral range and a template. σ_T and σ_S are the standard deviations of the template and spectrum respectively. Other features include overall and average intensity of selected bands. Features are assumed to only depend on adjacent features and thus, considered independent of all other features given their neighbors and source mineral.

Given an environment (for example, a location on Mars) scinetists select the minerals that may be present making sure to include minerals that are most likely to be confused with the response mineral(s). For the Mars example, there will be orbital data that can be used to help define the compositional materials to be encountered and assist in defining minerals that might cause confusion. Then, the minerals are partitioned into groups of similar minerals and the probability of the appearance of each group is estimated.

To summarize, the steps needed to prepare to train the BBN are: 1) select the relevant minerals; 2) define the prior probability of the appearance of each mineral; 3) define the probability of each mineral given its class; 4) define templates for feature extraction; amd 5) collect relevant samples of each chosen mineral.

The BBN should be considered as a method for enabling scientific expertise to be used in a robust manner. The efficacy of this topology is highly dependent on the accuracy of these preparatory steps. In this way, the BNN method is similar to an expert system in that it attempts to reason autonomously with information set by experts. The network is a model of the reasoning process of experts and has flexibility as an advantage. It is also able to make use of relationships that may be missed by experts if those relationships are contained within the training data. However, one limitation is that it remains incumbent upon a scientist to speculate on the needed minerals for the mineral variable, estimate the probabilities for the dummy variable and construct the feature extraction templates.

The current implementation uses the Netica software package provided by Norsys. Future work will continue using an alternate package (Tetrad) provided by [17].

3.2 Expert System

Carbonate_Identifier is a rule-based system for the identification of specific mineral target (carbonate) from reflectance spectra. It is written in C++ and implemented under a Coupled Layer Archicteture for Rover Autonomy (CLARAty) as a single class. This system consists of a hierarchy of components that preprocess a spectrum, identify and extract lists of features, apply a rule-based system to classify the spectra based upon the characteristics of these features, and forward the final results for any further uses. The overall structure of is shown in Figure 9.

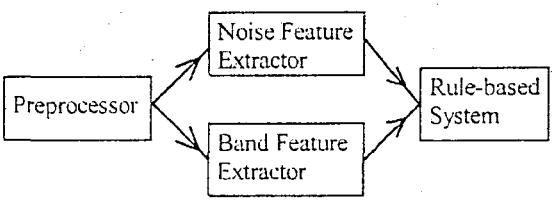

Figure 9. Structure of the expert system carbonate identifier.

The preprocessor renormalizes the spectrum to some predefined average albedo. This is done to avoid difficulties inherent in obtaining absolute albedos given that the spectra of the samples and the reference standard would likely be collected at widely different natural illuminations. In an operational scenario, the preprocessor will also perform such steps as discarding portions of the spectrum that may be contaminated by atmospheric features and compensating for any instrumental effects.

There are two feature extractors: 1) a noise feature extractor and 2) a spectral band identifier. The noise feature extractor calculates the standard deviation of the reflectance values between 2.0 and 2.5 μm to evaluate the amount of noise in the spectrum at this wavelength range. It is used to reject data with insufficient signal-to-noise characteristics. The spectral band identifier consists of a set of steps loosely based on the algorithms described by *Grove et al.* [18]. These steps apply a boxcar average to smooth the spectrum, subtract a hull fit if desired to remove the spectral continuum, and finally, search for inflection points and local minima to identify possible absorption bands. The operation of these feature extractors is controlled by a set of control parameters. The noise value and the positions, depth, and other characteristics of these spectral bands determined by these feature extractors are passed to the rule-based system as features.

The rule-based system, is a conventional forward-chaining expert system that applies a set of rules to a list of facts to generate new facts in an iterative fashion until no new facts can be obtained. It is based on the well-known algorithm described by *Winston and Horn* [19] and does not contain any refinements such as the Rete algorithm [20] to improve performance. This design was chosen for reasons of simplicity, reliability, clarity, ease of modification, and speed. In particular, the code was designed to be as compact as possible to conform to the memory limitations of typical spacecraft CPUs. It is well-suited for the fast and efficient identification of a small number of possible minerals.

This approach has several advantages that make it well-suited for autonomous spacecraft operation. In particular, it is compact (on the order of 1-200 kB) and extremely fast. It has also proved reliable. During the 1999 Marsokhod Field Tests [21,22], a prototype of this system produced success rates and false positive rates comparable to those of a human expert [23].

This approach also has a number of significant disadvantages. In its current form, it is unable to learn new rules on its own. While this limitation could be addressed by the incorporation of rule-learning schemes, it is unclear that this would offer any significant advantages over simply uploading new rules to the spacecraft as they might be required.

The system may also be poorly suited for the identification of large numbers of different minerals. In its current form, performance scales as N^2, where N is the number of rules, and while this limitation could potentially be addressed by use of the Rete [20] algorithm, we have no plans to implement the relevant changes.

4. SUMMARY

- Existing time, communication, and computational resource constraints currently associated with operating rovers on Mars suggest on-board science evaluation of sensor data can contribute to decreasing human-directed operational planning, optimizing returned data volumes, and recognition of unique or novel data. All of which can act to increase the scientific return from a mission.
- Many different levels of science autonomy exist. Each impacts the data collected and returned by, and activities of, rovers
- Several computational algorithms, designed to recognize objects of interest to geologists and biologists, are representative of various functions that can produce scientific opinions.
- Several scenarios illustrate how these opinions can be used, but realistic testing and comparison to human performance remain to be evaluated.
- Future efforts to develop additional methods for the detection of geologically significant patterns in images and spectra are clearly needed.

REFERENCES

[1] mars.jpl.nasa.gov/mer/mission/spacecraft_rover_brains.html

[2] mars.jpl.nasa.gov/mer/mission/comm_data.html

[3] mars.jpl.nasa.gov/missions/future/2005-plus.html

[4] Roush, T.L., V. Gulick, R. Morris, P. Gazis, G. Benedix, C. Glymour, J. Ramsey, L. Pedersen, M. Ruzon, W. Buntine, and J. Oliver, "Autonomous science decisions for Mars sample return", *Lunar Planet. Sci. Conf. XXX*, Abstract #1586, Lunar & Planetary Inst., Houston (CD-ROM), 1999.

[5] Castaño, R. L., T. Mann, and E. Mjolsness, "Texture analysis for Mars rover images", *Proc. Applications of Digital Image Processing XXII, SPIE Vol. 3808*, 1999.

[6] Castaño, R.L., R. C. Anderson, T. Estlin, D. DeCoste, F. Fisher, D. Gaines, D. Mazzoni, M. Judd, "Rover Traverse Science for Increased Mission Science Return", *Proc. IEEE Aerospace Conference 2003*, Big Sky, Montana, 2003.

[7] Castaño, R.L., M. Judd , R. C. Anderson, and T. Estlin, "Machine learning challenges in Mars rover traverse science", *Intl. Conf. Machine Learning 2003*, Washington, DC, 2003.

[8] Manduchi, R., S. Dolinar, F Pollara, and A. Matache, "Onboard science processing and buffer management for intelligent deep space communications", *Proc. IEEE Aerospace Conference 2000*, Big Sky, Montana, 2000.

[9] Duda, R.O., & Hart, P.E. *Pattern Classification and Scene Analysis*. New York: John Wiley & Sons, 1973.

[10] J. F. Canny. "A computational approach to edge detection" *IEEE Trans. Pattern Analysis and Machine Intelligence*, 679-698, 1986.

[11] V. Gulick, R. Morris, M. Ruzon, and T.L. Roush, "Autonomous image analyses during the 1999 Marsokhod rover field tests", *J. Geophys. Res, 106*, 7745-7764, 2001.

[12] Williams, D. and M. Shah, "A Fast Algorithm for Active Contours and Curvature Estimation", *CVGIP: Image Understanding. 55*, 14-26, 1992.

[13] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models", *Proc. First Intl Conf. on Computer Vision*, London, 259-269, 1987.

[14] Hough, P.V.C. (1962). "Method and means for recognising complex patterns". US Patent 3069654.

[15] Hartley, R. and A. Zisserman, *Multiple View Geometry*, Cambridge University Press, 2000.

[16] Pearl, J., *Probabilistic reasoning in intelligent systems: networks of plausible inference*, (Series in representation and reasoning). San Fransisco, CA, Morgan Kaufmann, 1998.

[17] Glymour, C. *Computation Causation and Discovery*, Boston, MA: MIT Press, 1999.

[18] Grove, C.I., S.J. Hook, and E.D. Paylor II, Laboratory "Reflectance spectra of 160 minerals", Jet Propul. Lab., Pasadena, California, 1992.
[19] Winston, P.H. and B.K.P. Horn, *Lisp*, 1st ed., Addison-Wesley, Reading Mass., 1981.
[20] Forgy, C.L., Rete: "A Fast Algorithm for the Many pattern/many object pattern match problem", *Artificial Intelligence*, 19, 17-37, 1982.
[21] Stoker, C.R., and 34 others, "The 1999 Marsokhod rover mission simulation at Silver Lake, California: Mission overview, data sets, and summary of results", *J. Geophys. Res., 106*, 7639-7664, 2001.
[22] Gazis, P.R., and T.L. Roush, "Autonomous identification of carbonates using near-IR reflectance spectra during the February 1999 Marsokhod field tests", *J. Geophys. Res., 106*, 7765-7773, 2001.
[23] Ramsey, J., P. Gazis, T. Roush, P. Spirtes, and C. Glymour, "Automated remote sensing with near infrared reflectance spectra: Carbonate recognition", *Data Mining and Knowledge Discovery, 6*, 277-293, 2002.

BIOGRAPHIES

Ted L. Roush *is a space scientist at NASA Ames Research Center. He has been involved in surface compositional interpretation of solid objects throughout our solar system. He has a BS in Geology from Univ. of Washington & a MS and PhD in Geology & Geophysics from Univ. of Hawaii.*

Mark Shipman *is software engineer at Raytheon Inc.. His efforts focus on image analysis, machine learning, & statistical inference. He has a MS in Logic and Computation from Carnegie Mellon Univ. and a BA in Philosophy from Univ. of Dallas.*

Robert Morris *is a software engineer at the SETI Institute. His efforts focus on image understanding, machine learning, & statistical inference. His BS degrees are in Physics, Susquehanna Univ. (BS), and Electrical & Computer Engineering, New Mexico State Univ. (MS).*

Paul Gazis *is a space scientist at San Jose State Univ. Foundation. His efforts focus on spectral analysis, machine learning, and statistical inference. His MS degree in Physics is from CalTech and his PhD degree is from MIT.*

Liam Pedersen *is a roboticist at QSS Inc. His efforts focus on statistical inference. He has degrees in Physics, Mathematics and Computer Science (BS, Rhodes Univ., South Africa), Electrical Engineering (MEE, George Washington Univ.), and Robotics (PhD, Carnegie Mellon Univ.).*