Rover Traverse Science for Increased Mission Science Return

Rebecca Castaño, Robert C. Anderson, Tara Estlin, Dennis DeCoste, Forest Fisher, Daniel Gaines, Dominic Mazzoni, Michele Judd

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, CA 91109
(818) 393-5344
Rebecca.Castaño@jpl.nasa.gov

Abstract—Rover traverse distances are increasing at a faster rate than downlink capacity is increasing. As this trend continues, the quantity of data that can be returned to Earth per meter traversed is reduced. The capacity of the rover to collect data, however, remains high. This circumstance leads to an opportunity to increase mission science return by carefully selecting the data with the highest science interest for downlink. We have developed an onboard science analysis technology for increasing science return from missions. Our technology evaluates the geologic data gathered by the rover. This analysis is used to prioritize the data for transmission, so that the data with the highest science value is transmitted to Earth. In addition, the onboard analysis results are used to identify science opportunities. A planning and scheduling component of the system enables the rover to take advantage of the identified science opportunity.

Although our techniques are applicable to a wide range of data modalities, our initial emphasis has focused on image analysis, since images consume a large percentage of downlink bandwidth. We have further focused our foundational work on rocks. Rocks are among the primary features populating the Martian landscape. Characterization and understanding of rocks on the surface is a first step leading towards more complex in situ regional geological assessments by the rover.

TABLE OF CONTENTS

1. INTRODUCTION

Of great importance to NASA's program of exploring the Solar System, and Mars in particular, is the development of intelligent rover control and science analysis systems capable of autonomous operations. Improvements in the acquisition of data with a corresponding increase in the number of planned missions means that the constraints on the volume of data will be governed not just by windows of opportunity or transmission, but by the limits on communication via the Deep Space Network.

We have developed technology for increasing science return from missions by prioritizing geologic data for transmission, leading to data with higher science value being transmitted to Earth. The result is an onboard system that can analyze collected data to identify targets and prioritize the data for transmission. A system diagram of our Onboard Autonomous Science Investigation System, OASIS, is shown in Figure 1.

In this paper, we will describe our methods for the prioritization of geologic data acquired by an in-situ rover. Our techniques are applicable to a wide range of data modalities, however our initial demonstration is focused on image analysis, as images consume a large volume of the downlink bandwidth for such missions.

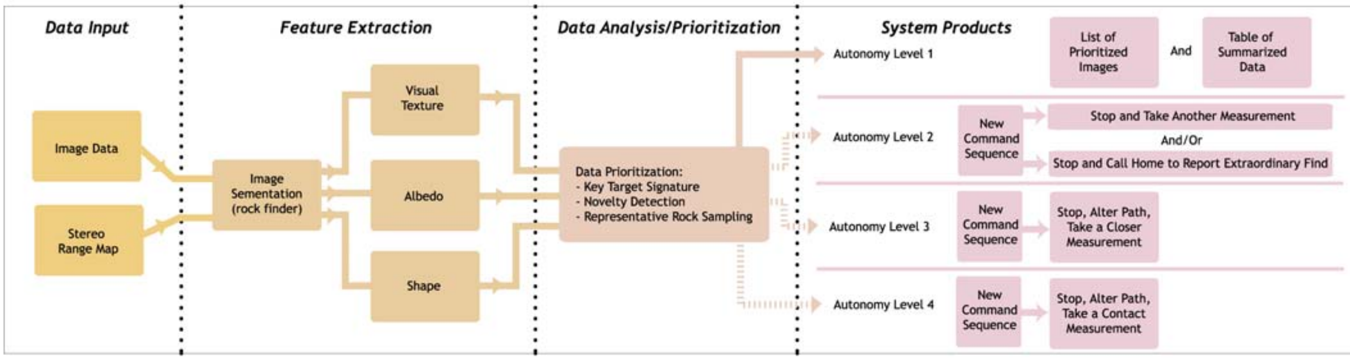


Figure 1. Overview of the Onboard Autonomous Science Investigation System - OASIS. Features such as visual texture, albedo and shape are used to prioritize rocks/images or for downlink and additional data collection.

Image prioritization involves two processes: the detection of scene features and the use of these features to assess the scientific value of the scene. The first step in image evaluation is the extraction of features of interest from the scene depicted. Our work has focused on properties of rocks in the scene, and thus we begin by locating rocks in a stereo image pair. Rock properties including albedo and visual texture are then extracted from the rocks identified. The properties extracted from a group of images are then used to prioritize the images. Three prioritization methods are discussed: identification of key target signatures, novelty detection, and sampling by clustering.

Without extensive ground testing and validation, scientists will be extremely reluctant to use autonomous, on-board prioritization of data for downlink. Our first challenge was acquiring representative rover traverse data sets. We then prioritized the data within each set and validated our results. We have developed a robust method for quantifying the correlation between our automated prioritization and a scientist's prioritization of the same data set.

Finally, prioritization can be used for more than just data downlink decisions. It can also be used for opportunistic science. Targets of high science value can be identified for additional instrument measurements.

As NASA ventures into an era of longer mission timetables with increasing requirements on downlink bandwidth for data acquisition, a mechanism for prioritization of data onboard will be critical. A prioritization process that can be used to reduce the volume of data returned, while increasing the science content of the returned data, is a necessity for maintaining an expanding program of missions exploring the Solar System.

2. DATA ANALYSIS

Feature Extraction

The first step in image prioritization is the extraction of features of interest from the scene. Our work has focused on properties of rocks in the scene, and thus we begin by locating rocks in a stereo image pair. Previous methods for locating rocks in an image have used shadows and information about the sun angle [1]. Our technique for locating rocks is based on finding objects above the ground plane.

We begin by determining the ground plane from the stereo range data, which is already calculated for navigation purposes. We then produce a height image, in which the value of each pixel represents the elevation of the point above the ground plane. Level contours in the height image are calculated and then these contours are connected from peaks to the ground plane to identify the rocks [2].

Rock properties including albedo, visual texture and shape, are then extracted from the rocks identified We measure albedo, an indicator of the reflectance properties of a surface, by computing the average gray-scale value of the pixels that comprise the image of the rock. The reflectance properties of a rock provide information about its mineralogical composition. Shadows and sun angle can both affect the gray-scale value of a pixel. Although this can be corrected by using the range data along with knowledge of both the sun angle and the camera orientation, this foundational work does not address these specific issues.

Igneous rock Metamorphic rock

(a)

(b)

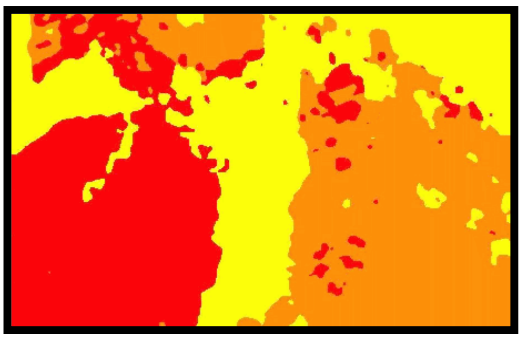


Figure 2. Examples of visual texture providing information about the geologic texture of rocks. (a) original image (b) image segmented based on texture.

The second rock property extracted is visual texture. Visual texture can provide valuable clues to both the mineral composition and geological history of a rock (see Figure 2).

These igneous and metamorphic rocks have the same mineral composition but have undergone different geological processes. Using visual texture, we are able to distinguish between the two different rocks.

Visual texture can be described by gray-scale intensity variations at different orientations and spatial frequencies within the image. We measure texture using a bank of Gabor filters [3, 4]. Gabor filters are scale and orientation specific, properties that make them successful in discriminating between different textures. Due to onboard computational constraints, a compromise must be made between the number of filters used and the discriminatory power of the filter bank. We have found that 12 filters (four orientations and three scales) are effective without being computationally prohibitive.

Another important and geologically useful feature of rocks is their inherent shape. For example, a rock that is highly rounded may have undergone fluvial processing and traveled far from its source while a rock that is highly angular is likely to be close to its source and has undergone minimal secondary processing. The shape of a rock is a complex property which is oftentimes difficult to describe precisely.

Our system describes the shape of a rock in an image using three parameters which capture how close to circular and how angular the rock is [5, 6]. We begin by fitting an ellipse to the boundary points of the identified rock in the image. Our first shape measure is the eccentricity of this ellipse. Our second measure is the error between the boundary points and the ellipse. The third and final measure is angularity, which is measured as the standard deviation of the angle of the edge at each boundary point. Image analysis is two dimensional, however rocks are three dimensional objects. From the range data, we have information about the 3D shape of the rocks. We have developed methods for estimating the sphericity and 3D angularity, and we are in the process of incorporating these into our system.

Rock Prioritization

The features extracted from a group of images are then used to rank the images using the three distinct prioritization algorithms described in this section.

Prioritization: Key Target Signature

Scientists have studied areas extensively and have an idea of what they expect to see or encounter during an in situ mission. On a mission, the instruments have all been carefully selected to collect information that will provide valuable insight into the history or current conditions on the planet. They have specific clues that they are looking for. Examples of what they are looking for include any signs of life past or present and signs of water, past or present. Each of the instruments will have a specific signal signature indicating the presence of key evidence. Thus, when only limited data can be sent to Earth, it is very important to scientists that any data containing these signatures is among the data that is returned.

We have implemented a method for enabling scientists to efficiently and easily stipulate the value and importance to give each feature. Rocks are then prioritized as a function of the distance of their extracted feature vector from the specified weighted feature vector. Scientists can either manually specify a feature vector, or they may select a rock from among the set of rocks already identified and rank the rocks as a function of the distance of their feature vectors from the feature vector of the selected rock (see Figure 3 for an example).

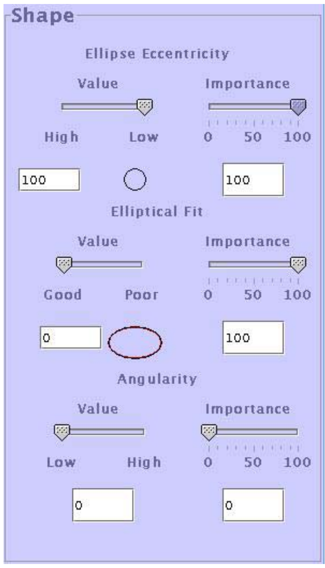


Figure 3. a) Target signature GUI for selecting round rocks, i.e. low eccentricity and good fit to an ellipse

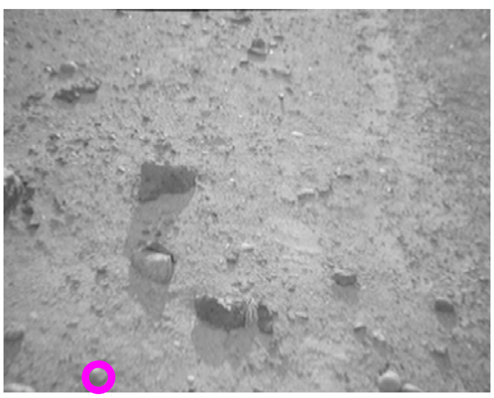


Figure 3. b) Image containing one of the best fitting (round) rocks to the designated target signature.

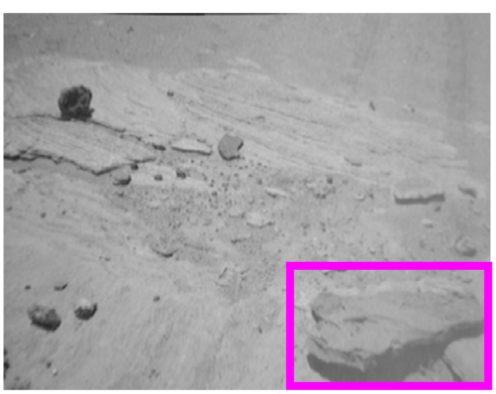


Figure 3. c) Image containing one of the worst fitting (not round) rocks to the specified target signature.

Prioritization: Novelty Detection

We have developed three methods for detecting and prioritizing novel rocks. They will have general utility for other novelty detection tasks as well, but are specifically designed with onboard computing constraints and the possibility of large feature spaces in mind.

First, we have a distance-based k-means clustering approach, in which a set of rocks are clustered. A new rock that is a great distance from any of the cluster centers is considered novel. In the second method, the probability density over the feature space for a set of rocks is approximated using a Gaussian mixture model. The novelty of a new rock is inversely proportional to the probability of that rock being generated from the learned mixture model.

The third novelty detection method uses a discrimination-based kernel one-class classifier approach. In this approach we treat all previous rock data as the "positive class" and learn the discriminant boundary that encloses all that data in feature space. Future rocks with features falling outside the boundary are considered novel.

These three approaches represent the three dominant flavors of machine learning approaches to novelty detection: distance-based, probability-based (i.e. "generative"), and discriminative. Considering all three types in one hybrid approach allows us to tradeoff their respective advantages and disadvantages.

Prioritization: Representative Sampling

One of the objectives for rover traverse science is to gain an understanding of the region being traversed. As such, it is desirable to have information on representative rocks, not just potentially very interesting unusual rocks, returned to Earth. A region is likely to be populated by several types of rocks with each type having a different abundance. A uniform sampling will be biased towards the dominant class of rock present and may result in smaller classes not being represented at all in the downlinked data.

To provide an understanding of the typical characteristics of a region, rocks are clustered into groups with similar properties and the data is then prioritized to ensure that representative rocks from each class are sampled. The rocks are clustered into groups based on the features extracted from the image data for each rock. To determine the classes, the property values are connected together in a series to form a feature vector, and a weight is assigned to the importance of each property. Different weight assignments can be used as a function of the particular properties that are of interest. For example, albedo and texture are typically used to distinguish types of rocks, but rock size may be used if sorting is of interest. Unsupervised clustering is then used to separate the feature vectors into similar classes, as shown in Figure 4. We currently employ

K-means due to its relatively low computational requirements, although any unsupervised method could be used. For each class of rocks we find the most representative rock in the class, i.e., the single rock in any image that is closest to the mean of the set. We give a high priority to the image containing this rock. The optimal number of classes can be determined using cross-validation techniques [7].

Figure 4. Image with two distinct types of rocks.

In the future we will use the spatial location of the rocks in addition to their property values to enable expanded analyses, including characterizing local surface regions and sorting which requires size and location information.

3. PRIORITIZATION FOR DOWNLINK

Image Prioritization

The initial approach to image prioritization has been to assign image priorities based on the rock rankings, i.e., the image containing the highest ranked rock is given the highest priority. In the near future this will be expanded so that all rocks in an image are taken into consideration in assigning the priority of the image.

Expert Validation

One of our primary concerns in this project is to develop techniques for validating the results of our autonomous prioritization algorithms. In particular, we would like to ensure that the prioritizations that our algorithms produce are comparable to those made by planetary geologists. We want a quantitative measurement so that we can gauge how closely our algorithms match the priorities of experts, and so that we can track progress that we make during development.

Our approach for validation is to gather sample prioritizations from expert planetary geologists on various collections of images. We use statistical methods to combine these expert prioritizations and compare them with the prioritizations produced by our algorithms [8]. The major advantage of this approach is that it is based on accepted statistical methods for combining and comparing rankings, and it provides a quantitative measure of how well our algorithms are matching expert rankings.

The implementation of our approach includes a web-based application to enable experts to prioritize images and add annotations for their decisions. The implementation also includes a set of statistical methods for comparing consistency across experts and measuring how well our rankings match expert rankings. An example ranking by an expert for a set of images is shown in Figure 5. The display includes the rank and score of each image. In this case, the score is the number of times it was preferred by the expert in a pair comparison. An alternate statistic is used when combining results from multiple experts.

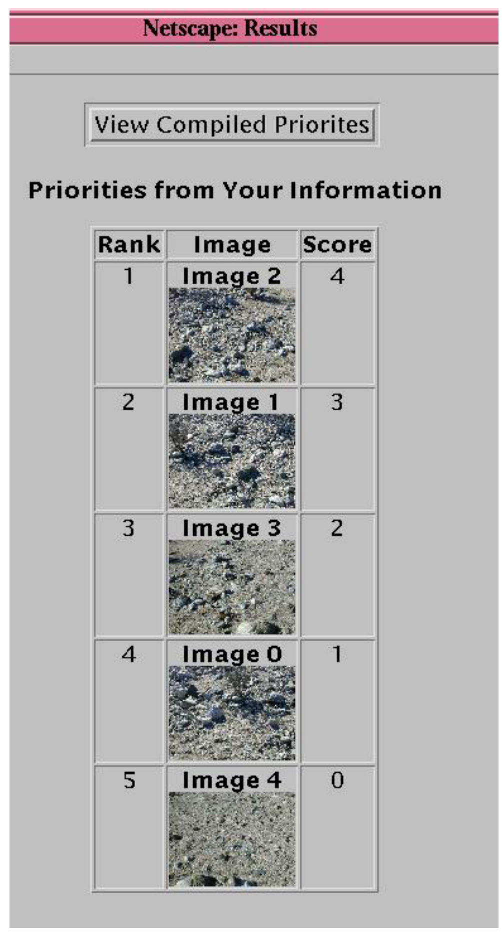


Figure 5. Example results for expert prioritization of a set of five images.

4. OPPORTUNISTIC SCIENCE

In addition to prioritization for downlink, the feature extraction and rock ranking can be used to identify science opportunities along a traverse. In particular, additional data may be collected on rocks that conform to key target signatures or that appear to be novel.

Once the data analysis software has identified a set of new science targets, these targets can be passed to other onboard autonomy software that will modify the onboard command sequence in order to collect the new science data. This capability is currently provided in our system by the CASPER planning and scheduling system [9, 10], which can:

1) Autonomously evaluate whether new science targets can be achieved given the state of the rover,
2) Modify the current command sequence to incorporate new targets, and
3) Monitor execution of that sequence in case further adjustments are necessary.

By integrating data analysis and planning capabilities, the resulting system can operate in a closed-loop fashion. This framework enables new science targets to be addressed onboard with little or no communication with Earth. An important contribution of this work is closing the loop between the sensor data collection, science goal selection, and activity planning and scheduling. Current approaches require human analysis to determine goals and to manually convert the set of high-level science goals into low-level rover command sequences. By integrating these components onboard, we enable a rover to function autonomously, as if a scientist were always in communication. This type of capability should dramatically increase the science return of future rover missions.

The CASPER planning and scheduling system operates by evaluating an input set of goals and the rover's current state and resource levels. Based on this information, CASPER generates a new sequence of activities that satisfies as many of the new goals as possible while obeying any rover resource and operation constraints. Goals are evaluated based on their priority (as assigned by the data analysis software), and if limited resources are available, then only the highest priority goals may be included in the new plan (i.e., command sequence). New plans are produced by using an "iterative repair" algorithm, which classifies conflicts and resolves them individually by performing one or more plan modifications. Once a new plan is produced, commands are sent to the rover's low-level control software for execution. During execution, status updates are relayed back from the control software where they are monitored by CASPER. As information is acquired regarding command status and actual resource utilization, the planner can update the current plan. New problems may often arise, requiring the planner to replan in order to accommodate the unexpected events.

The CASPER planning and scheduling system is also interfaced to the CLARAty robotic architecture [11, 12], which is being developed to support autonomous rover operations at the Jet Propulsion Laboratory and other NASA sites. CLARAty provides high-level decision-making software with a flexible interface to basic rover functionality. This functionality can range from low-level control of a motor or sensor to system level operations such as traversing a rover to a target location using obstacle avoidance. Currently, CLARAty provides a direct interface for the use of planning and scheduling software onboard the rover, and this interface has been tested in controlling two different rovers. We are currently extending the architecture to accommodate data analysis methods and enable these methods access to data as well as the capability to communicate with planning and other relevant software systems.

Testing and Validation

The system will be tested using rover data taken under controlled conditions in which the science opportunities that are to be discovered are known a priori. The testing sequence will begin using data acquired from the rover, as described above, where the rover is in the JPL Mars Yard with hand-placed science targets. As the system becomes more advanced, it will be tested onboard the rover with the rover taking appropriate action upon identifying a new science opportunity. Testing will be performed in situations with well-defined science opportunities as well as situations in which there are no special science opportunities during the traverse. The performance of the system will be evaluated by comparing the false alarm rate to the science opportunity detection rate. This ratio will be evaluated for science opportunities of varying detection difficulties.

5. CONCLUSIONS AND FUTURE WORK

The Deep Space Network will remain a constraining resource for future deep space missions as the number of high bandwidth missions increases. Traditional data compression can provide a valuable mechanism for increasing the amount of useful data returned; however, a limited amount of compression is possible before distortion levels become intolerably high. Science return can be maximized by returning the data with the highest science content possible. This requires a measure of science interest that can be evaluated onboard and a prioritization mechanism to rank the data for downlink based on the measured science interest. We have described a system that implements this functionality including extracting properties from image data and three methods of prioritizing the data that encompass the primary scientific exploratory objectives. We also described how the system can be validated by scientist and field tested. The use of onboard analysis to select the data with the highest scientific interest will be a critical functionality to maximize science return on future deep space missions with high data volume instruments.

REFERENCES

[1] 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, Vol. 106, No. E4, pg 7745-7764, 2001.

[2] V. Gor, V. Gor, R. Castano, R. Manduchi, R. Anderson, and E. Mjolsness, "Autonomous Rock Detection for Mars Terrain", *Space 2001, American Institute of Aeronautics and Astronautics*, Aug. 2001.

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

[4] M. Gilmore, R. Castano, T. Mann, R. C. Anderson, E. Mjolsness, R. Manduchi, and R. S. Saunders, "Strategies for autonomous rovers at Mars", in *J. of Geophysical Res*., Vol. 105, No. E12, Dec. 2000 pp. 29223-29237.

[5] J. Fox, R. Castano, and R. C. Anderson, "Onboard autonomous rock shape analysis for Mars rovers," *IEEE Aerospace Conference*, Big Sky, Montana, Mar. 2002.

[6] R. Castano, R.C. Anderson, J. Fox, J.M. Dohm, A.F.C. Haldemann, and W. Fink, "Automating shape analysis of rocks on Mars," Lunar and Planetary Science Conference, March 2002.

[7] P. Smyth, "Clustering using Monte-Carlo cross-validation", *Proc. 2nd Int. Conf. on Knowledge Discovery and Data Mining*, AAAI Press, 1996.

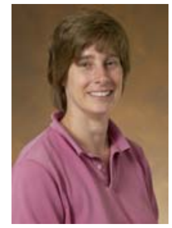
[8] E. L. Lehmann, *Nonparametrics: Statistical Methods Based on Ranks*, 1975.

[9] T. Estlin, A. Gray, T. Mann, G. Rabideau, R. Castano, S. Chien, and E. Mjolsness, "An Integrated System for Multi-Rover Scientific Exploration." *Proceedings of the Sixteenth National Conference on Artificial Intelligence,* Orlando, FL, August 1999.

[10] Chien, S., Knight, R., Stechert, S., Sherwood, R., and Rabideau, G., "Using Iterative Repair to Improve Responsiveness of Planning and Scheduling," *Proceedings of the 5th Int'l Conference on Artificial Intelligence Planning and Scheduling*, Breckenridge, CO, April 2000.

[11] Estlin, T., Volpe, R., Nesnas, I., Mutz, D., Fisher, F., Engelhardt, B., and Chien, S. "Decision-Making in a Robotic Architecture for Autonomy," *Proceedings of the Intl Symposium, on AI, Robotics and Automation for Space*, Montreal, Canada, June 2001.

[12] Volpe, R., Nesnas, I., Estlin, T., Mutz, D., Petras, R., Das, H., "The CLARAty Architecture for Robotic Autonomy," *Proceedings of the 2001 IEEE Aerospace Conference*, Big Sky, Montana, March 2001.

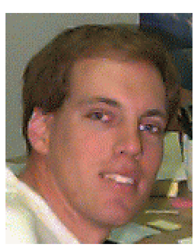


***Dr. Rebecca Castaño**, Supervisor, Machine Learning Systems group at JPL, OASIS Technical team lead. She received her Ph.D. in Electrical Engineering from the University of Illinois with her dissertion in the area of computer vision. Dr. Castaño has been advancing the state of the art in onboard science analysis methods for the past five years and has been lead author on numerous publications in this field. From 1999 - 2001, Dr. Castano served as the application lead for phenomenological computational field geology efforts in the MLS Group. She is currently the technology lead for science data processing for the Autonomous Sciencecraft Experiment on the New Millenium Program's ST6 project. Dr. Castaño is also the Team Lead of the Onboard Autonomous Science Investigation System (OASIS) project and the subtask lead for Data Analysis on the automated Multi-rover Integrated Science Understanding System (MISUS). Her research interests include machine learning, computer vision and pattern recognition.*

***Dr. Robert C. Anderson**, OASIS Science team lead, senior member technical staff. Dr. Anderson attended Old Dominion University in Norfolk, Virginia, where he received his Bachelor of Science degree in geology in 1979. In 1985, he received a Master of Science from Old Dominion University in geology with an emphasis on structural geology and mapping tectonic features surrounding the Tharsis region of Mars. In 1995, he received a Doctor of Philosophy from the University of Pittsburgh in geology with an emphasis on visible and near infra-red remote sensing. His Ph.D. research was centering on mapping Quaternary surfaces and soils around the Whipple Mountains of southwestern Arizona. Dr. Anderson worked on the successful Mars Pathfinder Project as science support for the Mineralogy and Geochemistry Science Operations Group. Currently Dr. Anderson is the Investigation Scientist for the Rock Abrasion Tool (RAT) and science support for Mission Operations on the Mars 03 mission. He works closely with the FIDO rover team and is presently the Science Team lead on the Onboard Autonomous Science Investigation System (OASIS) project. Dr. Anderson's research is centered on unraveling the geologic history of Mars.*

Dr. Tara Estlin*, OASIS Planning and Scheduling team lead, senior member technical staff. Dr. Estlin is a member of the Artificial Intelligence Group at the Jet Propulsion Laboratory where she performs research and development of planning and scheduling systems for rover automation and multi-rover coordination. Dr. Estlin is currently the PI of the "An Onboard Scientist for Multi-Rover Scientific Exploration" and the "Integrated Resource and Path Planning" projects. These efforts are developing capabilities for onboard rover-command generation, resource planning and scheduling, and data analysis for single and multiple rovers. She is also a member of the "Robotic Autonomy Architecture" project and is a past member of the Long-Range Science Rover (LRSR) project, which developed the Rocky 7 and Rocky 8 rovers. Dr. Estlin has been at JPL since 1998. She received a B.S. in computer science in 1992 from Tulane University, an M.S. in computer science in 1994 and a Ph.D. in computer science in 1997, both from the University of Texas at Austin.*

Dr. Dennis M. DeCoste *is a Senior Member of Technical Staff in the Machine Learning Systems Group at the Jet Propulsion Laboratory. He received his Ph.D. in Computer Science / Artificial Intelligence from the University of Illinois at Urbana-Champaign in 1994. Dr. DeCoste's research interests are in data mining and machine learning of time-series and image data, for both engineering and science domains. Since 1994 he has lead several research projects at JPL in these areas, including automated modeling and fault detection for a variety of NASA domains, including Deep Space 1, Space Shuttle, Rocky 7 Mars Rover, IPEX-II SIM, and MISR. His fault detection software flew on the recent DS1 mission, as part of the onboard Beacon / data summarization experiment.*



Forest Fisher *is a senior member of the technical staff, Artificial Intelligence Group of JPL, California Institute of Technology. He holds a B.S. in Computer Science from the University of Texas at Austin, and is currently completing a M.S. in Computer Science at the University of Southern California. Mr. Fisher performs research and development of automated planning, scheduling and execution systems for autonomous rover operations and ground communication station automation. He is the PI of the "Unified Planning and Execution" project. He is also a member of the "Onboard Autonomous Science Investigation System" (OASIS) and the "Multi-Rover Integrated Science Understanding System" (MISUS) projects.*

Dr. Daniel Gaines *is a senior member of the Artificial Intelligence Group at JPL. His research interests are in integrated planning and execution and in machine learning to improve planning. Dr. Gaines received a Ph.D. in Computer Science from the University of Illinois at Urbana-Champaign. Before coming to JPL, Dr. Gaines was an Assistant Professor in Computer Science at Vanderbilt University. His work at JPL is primarily focuses on planning and execution techniques for planetary exploration rovers.*

Dominic Mazzoni*, OASIS Research Programmer, associate staff, Machine Learning Systems. Mr. Mazzoni received his Bachelor of Science degree in Mathematics from Harvey Mudd College in Claremont, CA in 1999 and expects to receive his Master's Degree in Computer Science from Carnegie Mellon University in Pittsburgh, PA in early 2003. His Master's work focused on Music Information Retrieval, resulting in a working system for retrieving songs from a database of music using hummed queries. His work at JPL has focused on extending Machine Learning algorithms to work on massive implicitly-defined data sets.*

Michele Judd, P.E*., OASIS project manager, senior member technical staff. Ms. Judd received her B.S. in Petroleum Engineering from Stanford University and her M.L.A. in Organizational Development from Southern Methodist University. A nationally certified Quality Manager and licensed professional engineer, Ms. Judd has 14 years of experience in project management, both within JPL and for Mobil Oil.*