Training of a Crater Detection Algorithm for Mars Crater Imagery

Tatiana Vinogradova, Michael Burl, Eric Mjolsness
Jet Propulsion Laboratory
California Institute of Technology
Pasadena CA 91109
{firstname.lastname}@jpl.nasa.gov

Abstract— Automatic feature identification from orbital imagery would be of wide use in planetary science. For example, the ability to count craters on homogeneous surfaces would enable relative dating of geological processes. The scaling of crater densities and impact rates with crater size is another important issue which could be addressed by automated crater counting. Geological feature cataloging can practically be achieved by hand-labeled imagery only for restricted numbers of features. To handle massive new data sets and higher resolutions such as those arising from Mars Global Surveyor, automated feature identification will be required. Many pattern recognition algorithms could be applied to this problem, but a systematic validation process will be required to select the best method for each scientific application and to determine its reliability for scientific use.

We demonstrate such a validation process applied to a particular trainable feature identification algorithm when used to detect craters in synthetic imagery and in Mars Viking Orbiter imagery. The feature identification algorithm is the Continuously Scalable Template Matching algorithm of (*Burl et al.,* 2001) The validation process involves separate experiments for subpopulations selected from a labeled crater corpus. The subpopulations are defined by crater density. For the selected subpopulations, the validation process includes training the algorithm on some craters and testing its identification accuracy on others. These results can be summarized in terms of statistical efficiency measures. Efficiency results depend on the subpopulation tested. We illustrate algorithm performance on data of several Martian regions of high scientific interest.

TABLE OF CONTENTS

1. INTRODUCTION

Problem: Geological Feature Identification

There is a widespread recognition by Mars planetary scientists of the need to identify large numbers of features such as craters, faults, ridges, and channels, as well as relationships between such features, in order to infer the history of important geological processes on Mars. Processes of major interest include water floods, wind or water erosion, volcanic lava flows, ash falls, and impact cratering, all of which are believed to have contributed to the modification of the Martian surface. A diverse suite of features observed from Viking and Mars Global Surveyor MOC imagery, for example, bear close resemblance to glacial, fluvial, lacustrine, periglacial and mass-movement features of Earth. If such features exist(ed) on the surface of Mars, then processes requiring a relatively dense atmosphere and related transport and precipitation of water are necessary for their formation (Baker, 2000). The anomalous character of these landforms in regard to the very cold, dry present-day Martian conditions have been interpreted to be the result of one or more short-duration (perhaps 10^3 to 10^4 years) episodes of environmental change occurring since about 10^7 to 10^8 years ago (Baker, 2001). Such landforms, coupled with other features such as faults and fractures, pit crater chains, and fluvial channels, many of which cut some of the stratigraphically youngest materials on Mars and that often lack impact craters on their floors, could be representative markers of a recent volcanically, tectonically, and hydrologically active Mars.

Whether Mars has been recently hydrologically, tectonically, and volcanically active is a topic of widespread scientific debate, especially because of its extreme significance to the near-term exploration, investigation, and identification of aqueous, hydrothermal, and/or biologic activity. Although diverse geological information consistently and coherently indicate a recent or perhaps a currently active Mars, the development and validation of state of the art automatic feature identification methods will permit the planetary community to readily map, document, and catalog the geometric shapes and stratigraphic positions among geological features such as impact craters and faults, which will in turn yield improved interpretations of the geological and paleoclimatological histories of Mars at local, region, and global scales. Although the scientific community has performed such analyses since the Mariner and Viking missions, tools used up until now simply assist hand-scanning by knowledgeable experts, so that comprehensive data sets take many years to build. As an

example catalog of all Martian impact craters larger than about 5 km in diameter was produced by N. Barlow in the 1980s using the Viking 1:2,000,000 scale photomosaics (*Barlow*, 1988). The *Catalog of Large Martian Impact Craters* contains information on 42,283 impact craters across the entire Martian surface. However this approach is unlikely to scale up successfully to the full variety of geological features relevant to complex hypothesis, nor does it address possible observer bias. An even more serious problem is that of handling new high-resolution image sets (VO, MOC, MOLA). Furthermore, new refined state of the art feature detection techniques will enable past Herculean mapping efforts, which often times took years for a mapper to complete due to the time-intensive mapping and relative age determination of geologic features and terrains (including determining crater densities), to be much more time efficient and cost effective. That in turn will yield high scientific returns. For these reasons it is becoming increasingly important to provide the scientific community with automated tools which can assist in global analyses on scaled-up data sets and complex feature interactions.

The classical approach to detect geological features involved the recognition and definition by planetary scientists of their need to identify large numbers of objects such as craters for surface dating studies in large volumes of image data; this problem definition was followed by computational research in applying and adapting standard pattern recognition methods to the identification of such objects; finally planetary scientists could qualitatively evaluate the algorithm performance. In this paradigm, the accumulation of successively higher volume data sets from robotic missions required either heroic efforts at human examination of imagery for geological content (*Kanefsky, B., Barlow, N.G. and Gulick*, 2001), or new automation tools, or both. Examples of the first approach include Barlow's crater catalog (*Barlow, N.G.*, 1988, 2000), Anderson's fault catalog (*Anderson et al.*, 2000), and systematic mapping of a wider variety of feature types and terrains.

Automated approaches to geological feature identification have not yet gone into production use because they haven't been validated scientifically. Scientific validation requires that a method actually have sufficient accuracy for some well-defined scientific use, that it be computationally feasible, and that its accuracy be dependably quantified as a function of the difficulty of the feature identification problems it must solve within the context of a particular scientific application.

The significance of the research direction pursued here is that it will greatly increase the scientific return from orbital missions to Mars, and eventually landed ones, by greatly improving the derivation of geomorphology from global image and elevation datasets.

Previous Work

Pattern recognition algorithms taken from other application areas such as optical character recognition, particle physics, and medical imaging have been applied to problems identified in planetary science. Types of features which can be identified include circles, ellipses, lines, edges, ridges, junctions and image templates. For example line segments can be detected in images by the Hough Transform (HT). The principal idea of the method is to compute a mapping between a parameter space that defines the shape, one needs to detect, and an image space (the original HT method was invented for pattern-recognition in high energy physics (*Hough*, 1959); one of the most complete summaries of the HT method can be found in (*Olson*, 1996, 2000)). This led to development of generalized HT, which were used to detect craters in planetary images. A crater detection algorithm based on GHT for circle detection was applied and examined (*Honda et al.*, 1999) for selected lunar images (Clementine, LDIM); the method was not validated for scientific use. An ellipse/edge detection method based on GHT was also used for crater detection on asteroids for spacecraft position estimation and choosing a landing site (*Johnson et. al.*, 2001; *Johnson*, 2000, *Leroy et. al.*, 1999). The efficiency of the method (defined as a ratio of detected craters to true craters) was about 20%, which was acceptable for addressing navigational/landing problems.

A second general approach begins with template matching. Image pixel arrays are rotated, translated or otherwise transformed to match pieces of an image. This establishes an image-based paradigm, which may be extended to matched spatial filters, principal components methods and artificial neural networks (*Brunelli and Poggio*, 1995). For example, Venusian volcanoes were identified from Magellan radar data in (*Burl, Asker et al.*, 1998) which provides ROC curves for several related detectors compared to several hand-labelings. A subset of the data was deposited with the UC Irvine machine learning benchmarks repository.

The Continuously Scalable Template Matching (CSTM) algorithm is an image-matching algorithm which was implemented and tested on selected lunar crater images as reported in (*Burl et al.* 2001), complete with ROC curve. The algorithm uses examples (templates) provided by a scientist to generate a model for target detection in a user-specified continuous range of scales. The statistical efficiency of the implemented algorithm on the regions of the *Lunar Maria* (provided by Clementine) was about 80% for craters larger than four pixels in diameter with a 12% false alarm rate. However, complexity in the background terrain for the crater detection (images of Europa, for example) caused a reduction in algorithm performance.

Spatially invariant versions of current Support Vector Machine (SVM) algorithms were reported in (*DeCoste and Scholkopf*, 2001) where they were demonstrated to improve the world record on the MNIST database of 10,000 digit

recognition problems. They were also shown to improve the previous best ROC curves for the Magellan volcanoes. Speedup techniques, which make SVM competitive with trained feed forward neural nets in recognition speed are discussed in (*Decoste* ,2001). Both CSTM and ISVM algorithms can, at increased expense, be made to check for model instances at a variety of orientations and sun angles in cases where those parameters are not known or predictable.

In this paper, we will use the CSTM algorithm and subject it to systematic tests. Much of the validation methodology was introduced in (*Burl et al.,* 2001). The novel applications here are to synthetic imagery with arbitrarily scalable ground truth, and to real Mars imagery. The use of synthetic imagery is expected to be an important in scaling up scientific validation of algorithm performance to large Mars datasets for which hand-labeled data is too expensive to obtain. In addition, quantification of algorithm performance is extended here due to the availability of synthetic imagery and ground truth.

2. EXPERIMENTAL METHODS

For any feature identification system there are important quantitative measures of success. These include the probability of finding a target that is really there (detection efficiency, or "recall" in information retrieval), and the probability that an output detection actually corresponds to a real target (detection purity, or "precision"). If detection includes parametric outputs like position or scale, the accuracies of these outputs also can be measured. In that case, one may need to set a numerical criterion for sufficient accuracy for a "successful" detection before reporting either efficiency or purity. After such a criterion is defined, efficiency of the algorithm for feature identification can be calculated as a ratio of number of detected features to the number of "ground truth" features.

To define the "ground truth" for any surface we require adequate amounts of labeled image data. The most valuable such data is expert-labeled imagery of Mars, such as Viking Orbiter imagery with co-referenced feature label catalogs in Barlow's catalog of 42,283 Mars craters of size more than 4km (*Barlow*, 1988). However, expert-labeled data do not cover small size features (such as craters of less than 4 km in diameter, for example) which are very important for the scientific application, nor does it solve the problem of observer bias. The problem of reliability and biases of expert labeling was addressed for Venus volcano labeling (*Smyth. et al.*, 1995) and for MOC crater labeling (*Kanefsky et al.,* 2001).

Another approach for "ground truth" number estimation is the use of simulated terrain, where neither a problem of human biases, nor a size problem exist since features of any size (in the pixel space) with a known map location can be generated. Although Monte Carlo simulations almost always have some systematic and important differences from real data, they provide a volume of labeled data that is inexpensive, inexhaustible, correctly labeled, and can be similar enough to real data that trainable algorithms can be retrained to real data once they are functioning on synthetic data. We will use an updated version of R. Gaskell's terrain simulator (*Gaskell*, 1993, *Gaskell* 2001), which can support our goals as an analytical and predictive tool. The Martian terrain simulator has been used for rover navigability studies. It includes a number of simulations, each representing a specific geological process including impact cratering. Since it was created specifically for Mars, it includes Martian geological history up to the present knowledge with an appropriate parameterization. For the efficiency study we will use simulated terrain of the varying complexity and account for the potential problems in simulated vs. real terrain by scaling algorithms (the approach is discussed in detail in (*Vinogradova*, 1999)). Figure 1 shows simulated terrains of a different crater density with a generated crater size distribution.

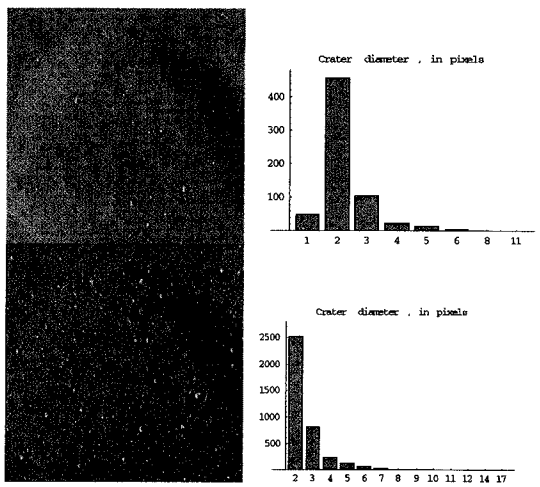


Figure 1. Example of simulated terrain with different crater densities; histograms show the distribution of crater diameters in pixels corresponding to each terrain.

Typically a detector is also parameterized by likelihood cutoffs and other parameters that affect the tradeoff between detection efficiency and purity; for each value of allowed impurity (1 - purity, also known as the false alarm rate) there is some optimal parameter set with maximal efficiency and we can plot that maximal efficiency as a function of allowed impurity. This is the well known Receiver Operator Characteristic or ROC curve, which shows the tradeoff between efficiency and purity.

Individual feature detectors are able to detect their targets in the image with some probability (efficiency) and false alarm rate (1-purity) as summarized in an ROC curve. Each detector is equipped with a crucial numerical readout: a measure of the likelihood of the detected match. If we sweep through possible values of a cutoff threshold for this

likelihood, eliminating detections in the order of increasing likelihood, we can make a series of measurements of purity and efficiency to be plotted against one another in the ROC curve.

Thus

$$\theta = \text{variable likelihood threshold, } \in [0,1]$$

$$\hat{D} = \text{fixed detection distance threshold}$$

$$n(\theta;\hat{D}) = \begin{array}{c}\text{number of features detected, with} \\ \text{likelihood } p \geq \theta \text{ and distance } D < \hat{D}\end{array}$$

$$N = \text{ total number of features}$$

$$M(\theta;\hat{D}) = \text{ total number of detections with } D < \hat{D}$$

$$\varepsilon(\theta) = \text{ efficiency } = n(\theta;\hat{D})/\,N$$

$$\pi(\theta) = \text{ purity } = \text{ precision } = n(\theta;\hat{D})/\,M(\theta;\hat{D}).$$

This function should in principle be monotonically increasing in θ, hence invertible, but may not be for actual detectors. The function $M(\theta;\hat{D})$ is, however, monotonically decreasing in θ and hence invertible.

Finally the translation- and scale-invariant "distance" function between a detected feature at (x,y) with radius r, and a true feature at (x',y') with radius r', is taken to be:

$$D(x,y,r;x',y',r') = \frac{(x-x')^2+(y-y')^2}{rr'} + \frac{(r-r')^2}{rr'}$$

which has a visually acceptable threshold of about 2 or 3. This distance function differs from the one used for the same purpose in (Burl et al., 2001), which measured instead relative area of crater overlap.

Using these definitions, the ROC curve at a given distance threshold $\hat{D}$ is determined by

$$\varepsilon(\pi) = \varepsilon\left(\theta(\pi;\hat{D})\hat{D}\right) = \pi M\left(\theta(\pi;\hat{D})\hat{D}\right)/\,N\,.$$

This expression only makes sense if $\pi(\theta)$ is invertible. Otherwise, the definition may be taken to be:

$$\varepsilon(\pi) = \max_{\{\theta|\pi(\theta)\geq\pi\}} \varepsilon(\theta;\hat{D})$$

which is monotonically nonincreasing. Conventionally the ROC curve is actually $\varepsilon(\phi = 1-\pi)$ where ϕ is the false alarm rate.

3. Results and Analysis

Figure 2 shows a moderately cratered synthetic crater image with detections and ground truth superimposed. Figure 3 shows histograms of crater sizes in pixels and in meters for the moderately and heavily cratered images used here. Figure 4 shows a histogram of the distances D for all detected craters (with likelihood threshold θ=0.5) to their nearest true crater. Figure 5 shows histograms of the pixel errors in detected crater locations for moderately and heavily cratered terrains.

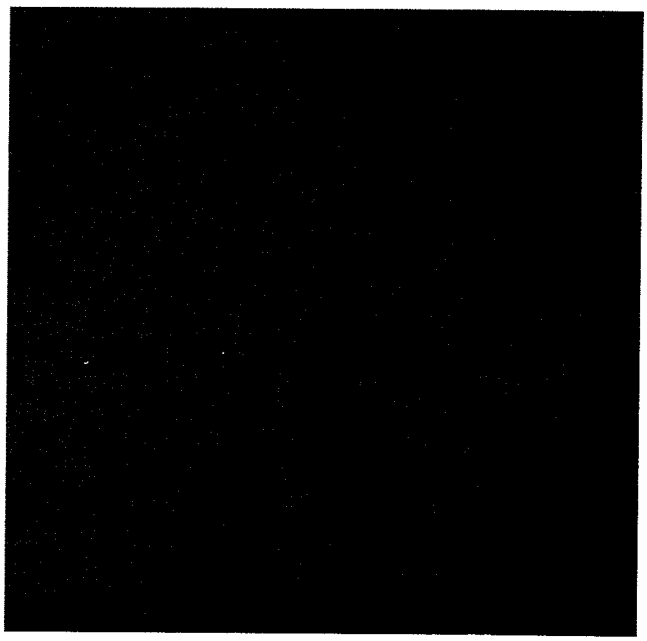
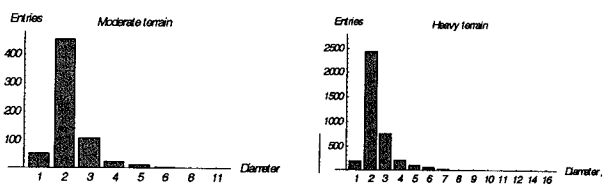


Figure 2. Simulated terrain image with detected craters (blue) and true craters (red) superimposed. Almost all craters (88%) are correctly detected. No spurious craters are detected.



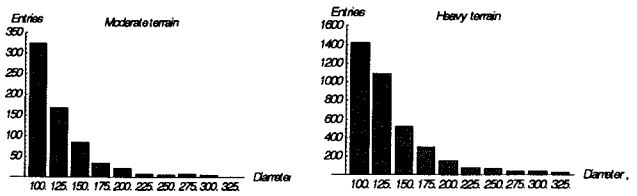


Figure 3. Histogram of crater diameters in two terrains, in pixels and in meters.

Figure 4. Distance histogram for crater-crater matches in moderately and heavily cratered terrain.

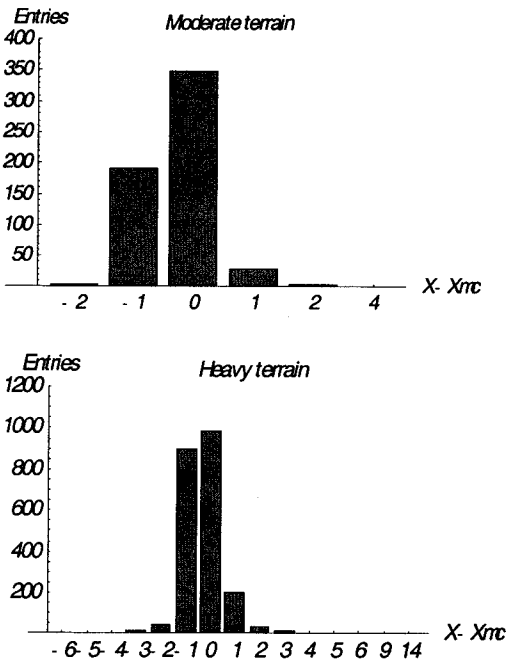


Figure 5. Histogram of pixel errors in x coordinate of detected craters shows very low error (0 or 1 pixel displacement) in recovered crater center. (a) Moderately cratered synthetic terrain. (b) Heavily cratered synthetic terrain.

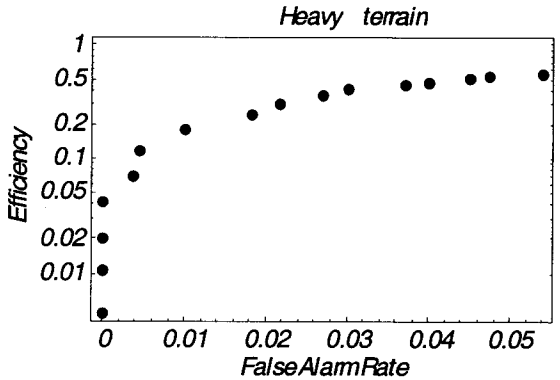


Figure 6. ROC curve for detection efficiency and purity in heavy terrain. Note very low false alarm rate. Curve shows all points available for crater detections with likelihood parameter $p > \theta = 0.5$. Lower values of θ will be required to explore the rest of the ROC curve. For moderate terrain, up to 88% of craters were detected with zero false alarms.

We have also experimented with real Mars terrain of varying complexity based on Viking Gazetteer Classification : Memnonia Fossae (Latitude (-23.67, -15.51), Longitude(140.33,164.21)) and Juventae Dorsa (Latitude (-1.30,3.96), Longitude (71.64,73.57)). We used the Barlow catalog for ground truth labelings of these images. The results appear promising for crater-counting applications, though they are not as strong as the synthetic data set results presented above. These experiments will be fully reported in the final paper.

4. Conclusions

The use of synthetic terrain allows for systematic validation of crater feature detection algorithms, which can then be extended to real Mars images for which sufficient ground truth labeling data is difficult or impossible to obtain. Preliminary results with the Continuously Scalable Template Matching algorithm are promising in both domains.

ACKNOWLEDGEMENTS

We wish to thank Robert Gaskell for providing access to his program for generating synthetic crater images, along with ground truth labelings. We also wish to acknowledge James Dohm and Robert Anderson for very useful discussions. This work was supported in part by the NASA Applied Information Systems Research Program (AISRP), and by the Automated Reasoning and Intelligent Data Understanding elements of the NASA Intelligent Systems program.

REFERENCES

1. Anderson, R.C., Golombek, M.P., Franklin, B. J.,

Tanaka, K. L., Dohm, J. M. , Lias, J. and B. Peer, "Primary centers and secondary concentrations of tectonic activity through time in the western hemisphere of mars", *In press, JGR, (2000)*

2. Barlow, N.G., (1988), *Icarus, 75, 285-305 (1988)*

3. Barlow, N.G., "Updates to the Catalog of Large Martian Impact Caters ", *Proc. Lunar and Planetary Science, 2000*

4. Barlow, N.G., R. A. Allen, and F. Vilas (1999). Mercurian impact craters: Implications for Polar Ground Ice., *Icarus 141, 194-204 (1999)*

5. Brunelli, R and Poggio, T., "Template Matching: Matched Spatial Filters and beyond," Artificial Intelligence Laboratory, MIT, A.I. Memo no. 1549, 14 pages, 1995.

6. Burl, M.C., Weber, M. and Perona, P., "A Probabilistic Approach to Object Recognition Using Local Photometry and Global Geometry", In *European Conf. on Computer Vision*, June 1998.

7. Burl, M.C., Asker, L. , Smyth, P., Fayyad, U., Perona, P., Aubele, J., and Crumpler, L., "Learning to Recognize Volcanoes on Venus", Machine Learning , April 1998.

8. Burl, M.C., C. Fowlkes, J. Roden, A. Stechert, S. Mukhtar, Diamond Eye: A Distributed Architecture for Image Data Mining, in SPIE AeroSense Conference on Data Mining and Knowledge Discovery, (Orlando, FL), April 1999.

9. Burl, M.C., Stough, T., Colwell, W., Bierhaus, E.B., Merline, W.J., Chapman, C., "Automated Detection of Craters and Other Geological Features", *Proc. I-SAIRAS-01, Monreal, Canada, June 2001*

10. Davies, A., Greeley, R., Williams, K., Baker V., Dohm, J., Burl, M., Mjolsness, E.,. Castano, R , Stough, T., Roden, J., Chien, S., Sherwood, R., "Autonomous Sciencecraft Constellation Science Study Report," Jet Propulsion Laboratory, http://asc.jpl.nasa.gov, August 2001.

11. DeCoste, D., Scholkopf, B., "Training Invariant Support Vector Machines", *Machine Learning Journal,* in press, Kluwer Academic Publishers.. 2001.

12. Gaskell, R., "Martian Surface Simulation", *Jour. Geophysical Research-Planets, 98(66), pp. 11099-11103, 1993.*

13. Gaskell, R., Collier, J., Husman L. and Chen, R., "Synthetic Terrain Enviroments for Simulated Missions", *JPL Internal Report, 2001.*

14. Honda, R., Konishi, O., Azuma, R., Yokogawa, H., Yamanaka S. and Iijima, Yu. "Data Mining System for Planetary Images – Crater Detection and Categorization", 1999.

15. Hough, P.V.C., Machine Analysis of Bubble Chamber Pictures, International Conference on High Energy Accelerators and Instrumentation, CERN, 1959.

16. Johnson, A., Klumpp, A., Collier J. and Wolf, A., "Lidar-based Hazard Avoidance for Safe Landing on Mars," *AAS/AIAA Space Flight Mechanics Meeting, February 2001.*

17. Johnson, A.E., "Surface Landmark Selection and Matching in Natural Terrain," *Proc. IEEE Computer Vision and Pattern Recognition (CVPR '00),* June 2000 .

18. Johnson, A.E., Matthies, L.E., " Precise Image-Based Motion Estimation for Autonomous Small Body Exploration", *Proc. 5 Int'l Symposium on Artificial Intelligence, Robotics and Automation in Space, (iSAIRAS'99), 1999, pp. 627-634.*

19. Kanefsky, B., Barlow, N.G., Gulick, V.G., "Can Distributed Volunteers Accomplish Massive Data Analysis tasks?", *In Lunar and Planetary Sciense Conference, 2001.*

20. Leroy, B., Medioni, G., Johnson, A.E., Matthies, L.H, "Crater Detection for Autonomous Landing on Asteroids", *Workshop on Perception for Mobile Agents IEEE Conf. Computer Vision and Pattern Recognition (CVPR'99), June 1999.*

21. Lindeberg, T., "Feature Detection with Automatic Scale Selection", *Int. J. of Computer Vision, vol. 30, number 2, 1998.*

22. Mjolsness, E., DeCoste, D., "Machine Learning for Science: State of the Art and Future Prospects", Science 293, pp. 2051-2055, 14 September 2001.

23. Pappu, S., Gold, S., and Rangarajan, A., "A framework for non-rigid matching and correspondence", *Advances in Neural Information Processing Systems 8*, pp. 795-801, 1996.

24. Nayar, S., Baker, S., Murase, H., "Parametric feature detection", *In Proc. Of IEEE Comp. Society Conf. On Computer Vision and Pattern recognition, pp. 471-477, 1996*

25. Olson, C. D., "Constrained Hough Transforms for Curve Detection", *Computer Vision and Image Understanding*, 73(3): 329-345, March 1999.

26. Olson, C. D., "Decomposition of the Hough Transform: Curve Detection with Efficient Error Propagation." *In Proceedings of the European Conference on Computer Vision, pages 263-272, 1996.*

27. Scott, D. H., and Tanaka, K. L., "Geologic Map of the Western Equatorial Region of Mars (U.S. Geological Survey map I-1802-A)", USGS 1986.

28. Smyth P., Fayyad U., Burl M., Perona P., Baldi P., "Inferring Ground Truth from Subjective Labeling of Venus Images", 1995.

29. Vinogradova T., (1999), Chapter "General statistical analysis principals" in the PhD thesis "Search for muon-neutrino to tau-neutrino oscillations with the NOMAD detector at CERN", *UMI-99-17231-mc (microfiche), 1999*

30. Vinogradova T., Mjolsness E. and Burl M., *2001 IEEE abstract submitted*, "Training of a crater detection algorithm on Mars crater imagery", Big Sky, Montana.

Tatiana Vinogradova *is a Member of the Staff in the Observational Systems Division of the Jet Propulsion Laboratory, California Institute of Technology. Dr. Vinogradova received a PhD in experimental elementary particle physics from UCLA in 1999.*

Michael Burl *was a Senior Member of the Staff in the Exploration Systems Autonomy section of the Jet Propulsion Laboratory, California Institute of Technology. He currently works at Idealabs in Pasadena CA. Dr. Burl received a PhD in computer vision from the California Institute of Technology in 1997.*

Eric Mjolsness *is a Principal in the Exploration Systems Autonomy section of the Jet Propulsion Laboratory, California Institute of Technology. He has held faculty appointments at Yale University, University of California San Diego, and the California Institute of Technology. Dr. Mjolsness received a PhD in physics from Caltech in 1986.*