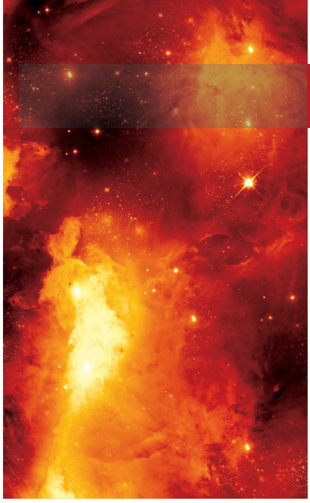

Using Onboard Clustering to Summarize Remotely Sensed Imagery

David S. Hayden, Steve Chien, David R. Thompson, and Rebecca Castano,
Jet Propulsion Laboratory

Many current and future NASA missions are capable of collecting enormous amounts of data, of which only a small portion can be transmitted to Earth. Communications are limited due to distance, visibility constraints, and competing mission downlinks. Long missions and high-resolution, multispectral imaging devices easily produce data exceeding the available bandwidth. As an example, the HiRISE camera aboard the Mars Reconnaissance Orbiter produces images of up to 16.4 Gbits in data volume but downlink bandwidth is limited to 6 Mbits per second (Mbps).

To address this situation, the Jet Propulsion Laboratory has developed computationally efficient algorithms for analyzing science imagery onboard spacecraft. These algorithms autonomously cluster the data into classes of similar imagery. This enables selective downlink of representatives of each class and a map classifying the imaged terrain rather than the full data set, reducing downlinked data volume. This article demonstrates the method on an Earth-based aerial image data set. We examine a range of approaches including k-means clustering using image features based on color, texture, temporal, and spatial arrangement and compare it to the manual clustering of a field expert. In doing so, we demonstrate the potential for such summarization algorithms to enable effective exploratory science despite limited downlink bandwidth.

Future Mission Challenges

Bandwidth issues are particularly poignant for future missions to the outer planets. NASA's Solar System Exploration Strategic Roadmap[1] outlines the role of aerial vehicles in future explorations of the solar system, particularly missions to Venus or Saturn's moon Titan. In the case of Titan, an aerobot would be capable of circumnavigating the moon within six months and remotely collecting sensed data some 8 km above ground level. The Cassini-Huygens mission has shown that Titan contains rich and varied landscapes (see Figure 1), including smooth and rough terrain, sand dunes, ethane lakes, shorelines, craters, and possibly cryovolcanoes. Additionally, there is a significant cloud presence. With such Earth-like diversity, Titan is of great scientific interest.

Data yield for a Titan mission would likely be limited not by the rate of image acquisition, but rather by communications constraints. Communication with Earth incurs latencies exceeding two hours. Downlink bandwidth is expected to be 4,500 bits per second, or 130 Mbits per day assuming an 8 hour transmission window.[2]

Autonomous methods of classifying aerial image data could preselect the most scientifically meaningful data for return to Earth. For example, spacecraft could transmit a representative sample of different image contents or prioritize specific features of interest. Previous work in onboard data understanding has already demonstrated selective data transmission on rover and satellite platforms. The Mars Exploration Rovers (MER) can automatically recognize science targets such as dust devils or clouds.[3] The Earth Observing-1 (EO-1) Satellite can detect hazardous events such as fires, floods, and volcanic activities and downlink pertinent data.[4] In both cases, targets can be detected

1541-1672/10/$26.00 © 2010 IEEE
Published by the IEEE Computer Society

without human direction, allowing for automatic data prioritization and improved science return.

Several challenges influence our design decisions for automatic image analysis. First, onboard processing is limited in spaceflight applications. Avionics computers must satisfy strict radiation and energy constraints, and they share resources with continuous autonomous control and data processing. One modern radiation-hardened processor used in space applications, the RAD750, is clocked at 200 MHz, has 128 Mbytes of RAM, and can achieve 400 million instructions per seconds (MIPS). This is more than a decade behind current desktop processing capabilities, which are approaching 100,000 MIPS.

Another challenge is the diversity of surface features the aerobot might encounter. An aerobot would be in constant motion with limited control due to unpredictable atmospheric currents (which makes it challenging to revisit sites of interest). It will also be difficult to schedule image targets in advance or to anticipate features of interest. This favors an "unsupervised" approach that makes few assumptions about image content but instead discovers interesting and representative samples based on the data's intrinsic properties. Scientists have proposed unsupervised methods for selective data return applications in the rover domain.[5,6] Such approaches have also been widely used for image search and retrieval[7] and image-sequence representation.[8,9] Clustering, which classifies a data set into discrete categories of items with similar properties, is one common unsupervised approach. Clustering has been applied to aerial imagery,[10] although not in an online fashion.

At JPL, we are working on unsupervised classification for selective transmission of aerial image data in remote space exploration. We represent images in a metric space to compare their similarities. We identify image feature descriptors to encourage clusters based on semantic content such as presence of horizons, clouds, and water bodies. A broad survey of different image features suggests several that are both computationally efficient for spacecraft computers and relevant to the image categories identified by planetary scientists.

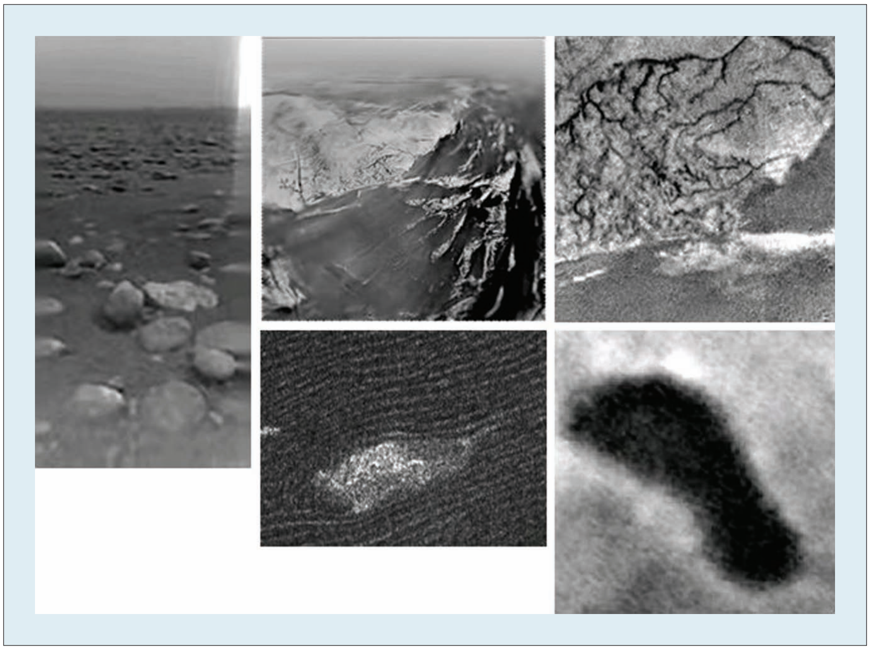

Figure 1. Cassini-Huygens images showing the diversity of Saturn's moon Titan. The top-row images are from the Huygens probe and show a field containing frozen rocks with horizon, a hill etched by hydrocarbon rain, and part of a dried riverbed, respectively. The bottom-row images are from the Cassini mission radar data and show sand dunes and hydrocarbon liquid bodies. (Courtesy of ESA, NASA, JPL, and University of Arizona [top and middle left]; NASA, JPL, ESA, and the University of Arizona [top right]; NASA and JPL [bottom left], and NASA and the JPL Space Science Institute [bottom right].)

To test our approach, we constructed a data set of aerial imagery using a consumer-grade digital camera (Canon PowerShot SD850 IS) with 1,600 × 1,200 pixel resolution. A total of 162 images were collected during a commercial airline flight from New York to Los Angeles. The images primarily contain shots dominated by sky, horizon, or undeveloped land. Some contain clouds, discernable water bodies, developed land, or small portions of the plane's wing or window at one or more edges. The data set has many attributes similar to what we would expect from Titan aerial imagery, including varied terrain, clouds, horizons, water bodies, and artifacts such as an occasional window or wing obstruction.

Image Features

Humans describing natural images refer to abstract concepts such as objects and places rather than primitives such as intensity and texture. These high-level interpretations are difficult to automate; our challenge is to find implementable methods to discriminate between different areas of science interest (such as terrain types).

Image data may be described as a series of measurements that represent the scene in an n-dimensional feature space. For example, images could be completely described by the intensity of each pixel, but this places them in a feature space with a dimensionality

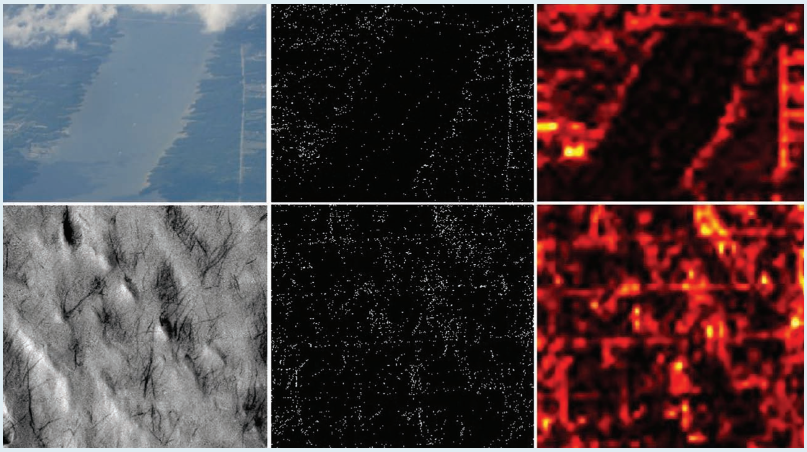

Figure 2. Edge detection (middle) finds areas of abrupt changes in intensity. The density of such edges, as visualized in a heat map (right), provide texture information. Water bodies (top) tend to be homogenous regions containing few edges compared to purely land-based imagery, such as this dust-devil-torn landscape (bottom) in the southern hemisphere of Mars taken by Themis (image ID V07829003[11]). The edge density of the top image is 0.02, and the edge density for the bottom is 50 percent greater at 0.03.

equal to the number of pixels. A modestly sized 1,024 × 1,024 grayscale image will thus be described in more than one million dimensions. Not only is this too large for convenient processing, but distances in this space might not reflect image content. For instance, changing a scene's illumination could shift the image's feature vector radically while leaving content virtually unchanged. However, a more appropriate set of feature measurements could permit meaningful automatic comparisons between images. Our goal is to find computationally efficient features that effectively discriminate images with interesting content.

Basic color information can be gleaned from statistics on each image's intensity histograms. Multispectral imagery contains n such histograms. For imaging in the visible spectra, color can help distinguish basic terrain types such as vegetation from desert or clouds from sky. Without features to describe texture, such basic color features won't be resilient to differing levels of illumination and could be tricked by similarly colored terrain (such as sky and ocean). Color might map to more distinguishing features such as surface temperature or height when using different imaging techniques (including infrared or radar).

Terrain texture can range from smooth to coarse. It might be directional, as in the case of windswept sands, or periodic, as in the case of sand dunes. We can approximate texture coarseness and direction by taking statistics on the gradient of the image intensity function. Edge-detection algorithms such as the Sobel operator approximate this gradient (both magnitude and direction) by convolving the image with specific filters. We can describe texture with statistics such as the density of high contrast changes (edges), mean gradient magnitude, and entropies of the magnitude or orientation responses. Coarse textures will have a higher edge density than smooth textures, as Figure 2 shows. Directional textures will have low entropy orientation responses.

We can also describe texture as a function of horizontal and vertical frequencies. The 2D Fourier transform accomplishes this by representing images as a series of sinusoids. Although difficult to visually interpret for nonsynthetic images, the sinusoids' magnitude can provide information on structure and texture using low-level energy statistics.[12] In particular, along with temporal features, we have found the sums of energy in quadrants 1 and 2 (top right and top left) of the normalized magnitude image to be effective.

Image-segmentation methods can find areas of interest that are each described by a different set of features but are likely too expensive for spacecraft computers. Giving up rotational invariance, images can instead be split into $n \times n$ equal-sized subimages. All features can be collected in each subimage and appended to a single feature vector for each image. Our results suggest that this provides small improvements for feature combinations that don't collect data on an image's Fourier transform.

Finally, in an aerial imaging scenario, we would expect image similarity to correspond with temporal proximity. We can encode this feature as the acquisition order of an image.

Clustering

Clustering algorithms group data into disjoint sets based on similarities in their feature vectors. One fast and common method, Lloyd's algorithm,[13] begins by placing k additional data points, or centroids, in the feature space. Assigning each data point to the closest centroid forms clusters. An updated centroid is calculated for each cluster and all data are again assigned to the closest centroid. This process continues until cluster membership no longer changes significantly. In practice, convergence is fast and we can stop the algorithm early for a less optimal solution.

The primary challenges with this algorithm are in choosing the number of clusters and the centroids' initial locations.

Once images are clustered in a metric space, we can perform selective data return. Images closest to each cluster centroid serve as exemplars for that category. Downlinking the n closest images to each centroid provides a broad overview of the types of data collected. If some clusters prove to be consistently interesting, then all images from those clusters can be prioritized for downlink. Finally, scientists can also opt to downlink images most dissimilar to the others—outliers or anomalies, which will be located far away from cluster centroids. Thus, clustering enables selective data return based on a representative sample, a biased sample, or outliers. These options provide compelling alternatives to conventional approaches such returning data collected at periodic intervals.

Cluster Comparisons

Ideal clusterings contain compact clusters that are spread far apart from one another. However, satisfying these qualities provides no guarantee that a given clustering will support a particular science objective. In our work, we use empirical tests to compare algorithmic clusterings with manual clusterings performed by experts. This provides insight into the match between each artificial feature space and the physical scene features of interest to the expert.

To evaluate performance, we compare the overlap between automatic feature-based clustering and time-based clustering (based solely on acquisition order) with the expert clustering. Automatic feature-based methods that outperform time-based clustering as measured by correspondence with experts are good candidates for selective data return. We used adjusted mutual information (AMI)[14,15] to compare the similarity between two partitionings of a data set.

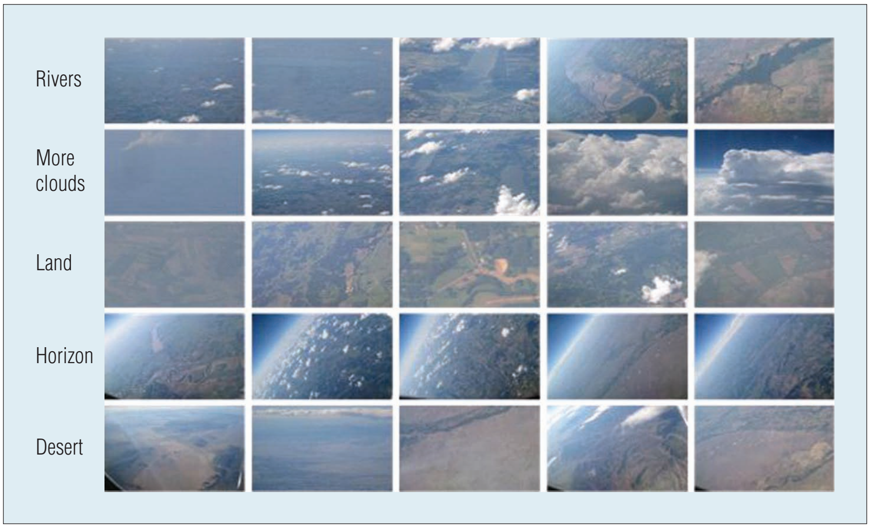

Figure 3. Expert cluster examples. Five randomly chosen images from each of five clusters created and named by a planetary volcanologist.

Expert Labeling

With the appropriate prompt, experts can manually cluster data to serve as a ground-truth standard for evaluation. For the terrestrial data set, we elicited a manual clustering from a planetary volcanologist. We provided the expert with this written prompt:

> Suppose that the following aerial images were taken of an environment for which we have little knowledge or data. Furthermore, suppose that you might not be able to receive all images. Please sort these images into five groups in such a way that if you could only receive a small number of images from each group you could reasonably infer the content of the remaining images in that group.

We did not allow the expert to choose the number of clusters because this could create confusion about the clusters' scope: a dozen clusters could incorporate rare classes like "solitary clouds" but a triumvirate classification could only capture broad distinctions like the presence of a horizon. Upon interview, the expert felt that five was a mostly adequate number; six would have been ideal so that an outlier group of "must return" images could have been established. Figure 3 displays examples from each expert cluster.

Experiments

Our planetary volcanologist chose to sort the aerial images by semantic distinctions: rivers, more clouds, land, horizon, and desert. Figure 3 shows representative images selected at random from each. In terms of low-level features, images in the horizon category contain a line separating the image into two regions of color and texture. Above the line the sky contains nearly uniform texture and a gentle gradient change from light to dark blue, while below the line both texture and color vary significantly. The land images largely contain patches of ground in shades of brown and green, while the desert images are dominated with brown patches. Texture in all the images generally becomes smoother as the altitude increases. The clouds in the more clouds group contain patches different from neighboring areas in color and texture. The river images appear to be the most difficult to describe in image primitives. One observation is that the

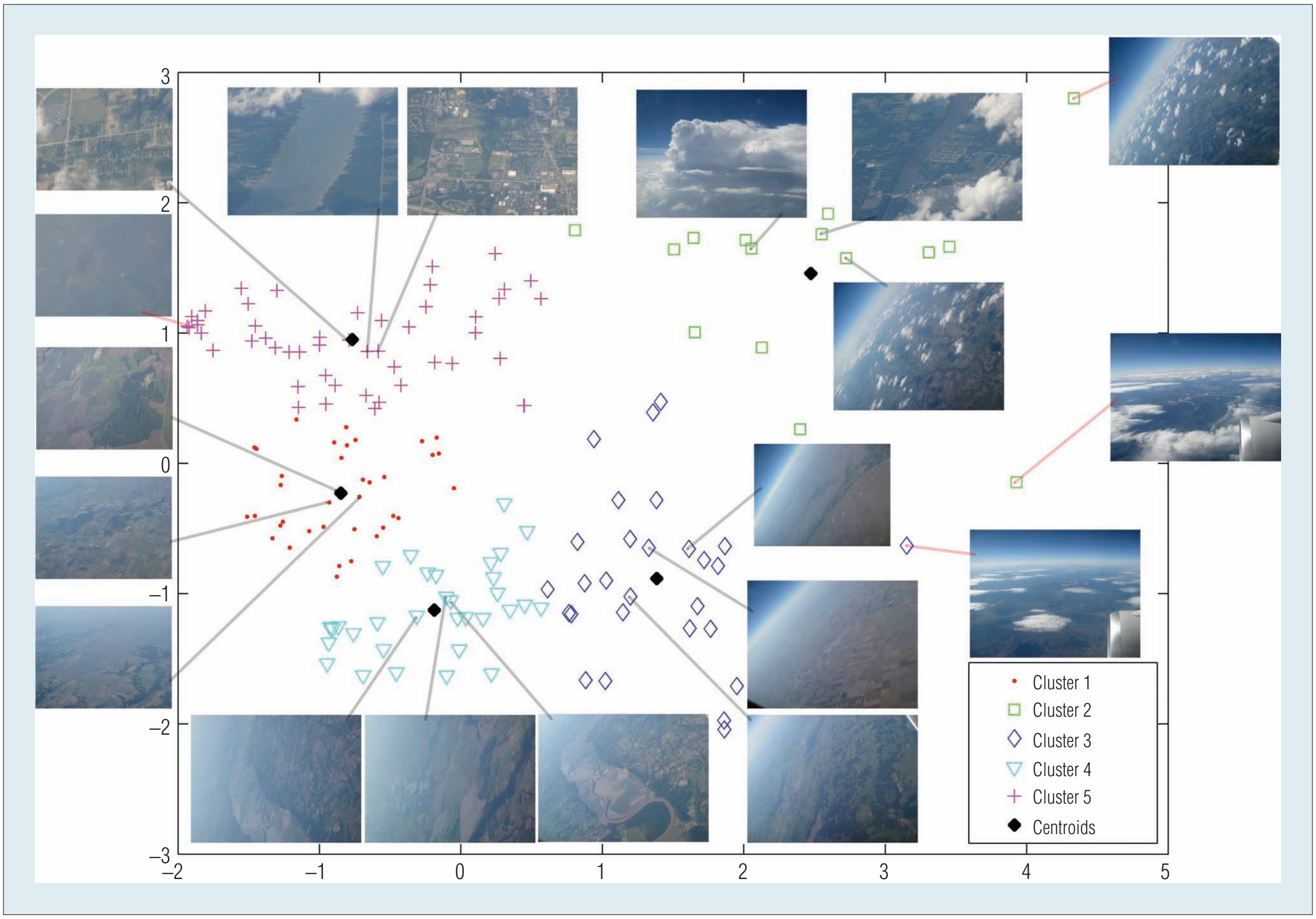

Figure 4. Example algorithmic clustering of a terrestrial data set using FT features. The gray lines link three representative images closest to each cluster centroid. The red lines indicate outliers.

rivers are meandering polylines of widely varying thickness and are typically of a different color and smoother texture than their surroundings.

Our tests suggest that automatic image features can reproduce the expert clustering better than either random or acquisition-order-based grouping. A combination of two frequency-space statistics along with acquisition order features performed best overall; we call these the FT features. Performance was quantified with the AMI score to measure the similarity between each automatic clustering and the target expert clustering. The FT feature set produced an improvement of 33 percent relative to periodic sampling by this metric.

Figure 4 displays an FT features-based clustering in two dimensions with near-centroid and outlier images. Most striking is the separation between images containing horizon and clouds from those containing ground-based images. Cluster 2 favors images with clouds while cluster 3 favors images with horizon. Both horizon and clouds are abundant in the outlier images in cluster 2. Both of these clusters contain images with significantly lower energy in the power spectrum of their Fourier transforms when compared to the rest of the images. This is likely due to the contribution of strong, lower-frequency signals from homogenously textured clouds and sky, and the weakly periodic nature of images speckled with clouds. No clear distinction exits between clusters 1, 4, and 5 in terms of the expert categories, but the representative images in cluster 5 have highly detailed textures and are taken at a relatively low altitude.

The features we have described are relatively fast and simple to collect. For the FT features, the most demanding operation is the $n \log n$ Fourier transform.[16] With a fixed number of iterations, clustering can be completed in linear time. Therefore, the total algorithmic complexity is subpolynomial in image size and linear in the number of images.

Although our aerial data was constructed from visible spectra, these techniques apply to all remote sensed imagery (such as radar). Because the features we used are general, they would likely perform well for scenarios other than aerial imagery such as land or underwater traversals.

Future directions include testing on additional data sets, finding domain-specific features in support of specific science tasks, careful initialization for clustering,[17] and automatic rejection of images marred by sensor artifacts or noise.

References

1. "2006 Solar System Exploration Roadmap," tech. report, NASA, 2006.
2. J. Hall et al., "An Aerobot for Global In Situ Exploration of Titan," *Advances in Space Research*, vol. 37, no. 11, 2006, pp. 2108–2119.
3. R. Castano et al., "Opportunistic Rover Science: Finding and Reacting to Rocks, Clouds and Dust Devils," *Proc. IEEE Aerospace Conf.*, IEEE Press, 2006, p. 16.
4. S. Chien et al., "Using Autonomy Flight Software to Improve Science Return on Earth Observing One," *J. Aerospace Computing, Information, and Communication*, AIAA, 2005, pp. 196–216.
5. R. Castano et al., "Oasis: Onboard Autonomous Science Investigation System for Opportunistic Rover Science: Research Articles," *J. Field Robotics*, vol. 24, no. 5, 2007, pp. 379–397.
6. D. Thompson, T. Smith, and D. Wettergreen, "Information-Optimal Selective Data Return for Autonomous Rover Traverse Science and Survey," *Proc. IEEE Int'l Conf. Robotics and Automation* (ICRA 08), 2008, pp. 968–973.
7. Y. Chen, J.Z. Wang, and R. Krovetz, "CLUE: Cluster-Based Retrieval of Images by Unsupervised Learning," *IEEE Trans. Image Processing*, vol. 14, no. 8, 2004, pp. 1187–1201.
8. L.S. Kennedy and M. Naaman, "Generating Diverse and Representative Image Search Results for Landmarks," *Proc. 17th Int'l Conf. World Wide Web*, ACM Press, 2008, pp. 297–306.
9. S. Krishnamachari and M. Abdel-Mottaleb, "Image Browsing Using Hierarchical Clustering," *Proc. 4th IEEE Symp. Computers and Communications* (ISCC), IEEE CS Press, 1999, pp. 301–309.
10. K.R. Harvey and G.J.E. Hill, "Vegetation Mapping of a Tropical Freshwater Swamp in the Northern Territory, Australia: A Comparison of Aerial Photography, Landsat TM, and SPOT Satellite Imagery," *Int'l J. Remote Sensing*, vol. 22, no. 15, 2001, pp. 2911–2925.
11. P.R. Christensen et al., "Planetary Data System Node," THEMIS public data releases, Arizona State Univ., http://themis-data.asu.edu.
12. S. Liu and M.E. Jernigan, "Texture Analysis and Discrimination in Additive Noise," *Computer Vision, Graphics, and Image Processing*, vol. 49, no. 1, 1990, pp. 52–67.
13. S. Lloyd, "Least Squares Quantization in PCM," *IEEE Trans. Information Theory*, vol. 28, no. 2, 1982, pp. 129–137.
14. N.X. Vinh, J. Epps, and J. Bailey, "Information Theoretic Measures for Clusterings Comparison: Is a Correction for Chance Necessary?" *Proc. 26th Ann. Int'l Conf. Machine Learning*, ACM Press, 2009, pp. 1073–1080.
15. L. Hubert and P. Arabie, "Comparing partitions," *J. Classification*, vol. 2, Dec. 1985, pp. 193–218.
16. J.W. Cooley and J.W. Tukey, "An Algorithm for the Machine Calculation of Complex Fourier Series," *Mathematics of Computation*, vol. 19, Apr. 1965, pp. 297–301.
17. D. Arthur and S. Vassilvitskii, "k-Means++: The Advantages of Careful Seeding," *Proc. 18th Ann. ACM-SIAM Symp. Discrete Algorithms* (SODA 07), Soc. for Industrial and Applied Mathematics, 2007, pp. 1035, 1027.

David S. Hayden is an affiliate of the Artificial Intelligence Group at the California Institute of Technology Jet Propulsion Laboratory (JPL) and a student at Arizona State University. Contact him at dshayden@asu.edu.

Steve Chien is a senior research scientist and head of the Artificial Intelligence Group at JPL. Contact him at Steve.Chien@jpl.nasa.gov.

David R. Thompson is a researcher in the Machine Learning and Instrument Autonomy Group at JPL. Contact him at david.r.thompson@jpl.nasa.gov.

Rebecca Castano is an assistant section manager for technology development and infusion in the Instrument Software and Science Data Systems Section and the ACOS science algorithms manager at JPL. Contact her at Rebecca.castano@jpl.nasa.gov.

cn *Selected CS articles and columns are also available for free at http://ComputingNow.computer.org.*