

# Mapping Surface Materials on Mars From Mars Pathfinder Spectral Images With HYPEREYE

Erzsébet Merényi<sup>1</sup>, William H. Farrand<sup>2</sup>, Philip Tracadas<sup>1</sup>

<sup>1</sup>Rice University, Houston, Texas, <sup>2</sup>Space Science Institute, Boulder, Colorado

E-mail: [erzsebet@rice.edu](mailto:erzsebet@rice.edu), [William.Farrand@colorado.edu](mailto:William.Farrand@colorado.edu), [tracadas@rice.edu](mailto:tracadas@rice.edu)

## Abstract

*A comprehensive mapping of spectral variations is presented for one octant of the Imager for Mars Pathfinder SuperPan data set. Both left eye and right eye images are analyzed, and for each, all respective spectral bands are utilized simultaneously. We use a Self-Organizing Map to achieve fine discrimination of over 20 surface units including previously published classes. Agreement with earlier analyses are very good where data are available for comparison. In spite of the separate analysis of the left and right eye data, which cover different spectral windows with little overlap, many classes show very similar spatial distribution in the left and right eye images. The SOM clustering produced refinements within the unit formerly labeled as “black rock”, discovered previously undiscussed units that may be various coatings on rocks, and presented some disagreements with existing units. The clustering tools are part of HYPEREYE, a dedicated research software developed with NASA/OSSA AISRP support.*

*Keywords: Self-Organizing Map, Martian surface, Mars Pathfinder, spectral image analysis*

## 1. Introduction

Self-Organizing Maps (SOMs) have long been recognized as powerful unsupervised clustering algorithms. Invented in its original form by Teuvo Kohonen [1], the Kohonen SOM (KSOM) has matured to spawn many significant applications. We have been using SOMs successfully to improve extraction of scientific information from a number of space science data sets, for example [2,3,4], as well as for terrestrial studies. The past successes inspired more serious theoretical studies and software development, which

are incorporated in our research environment, HYPEREYE, described in section 4. Beyond the original KSOM, we have become very interested in new theoretical augmentations that lend various advantageous properties to the basic KSOM. One such aspect is the so-called map magnification, described in detail in [5], which controls the density of the SOM weight vectors in the input data space and thus effects the designation of larger or smaller number of prototypes to represent certain subsets of the input space. It can force, in principle, several different types of quantization on the reference (weight) vectors of the SOM. Such quantizations are, for example, minimum distortion or maximum entropy quantization. Negative magnification enlarges the representation of small clusters in the neural map (the SOM) — a phenomenon observed in the cerebral cortex — thus can facilitate discoveries of rare events. Control over the magnification of the neural map is one important contribution toward precise clustering. A desired magnification can be achieved by explicit forcing [6], guaranteed by theory for restricted cases of data and SOM dimensions. Of particular interest to us in this work is the magnification which forces maximum entropy quantization and thus the most faithful representation of the unknown probability density of the input data space. Luckily, for this special case the desired maximum entropy quantization can be achieved without explicit forcing of the magnification, using the so-called conscience algorithm [7] instead of [6]. While this algorithm is heuristic, there is experimental evidence that faithful density matching is produced even for hyperspectral images [8]. More details on this particular SOM variant and its applications are found in [8,9,2-4]. We applied this approach to the present study.

Background on SOMs, as noted above, is given, for example, in [5,8,9], all downloadable from URL <http://www.ece.rice.edu/~erzsebet/publications.html>. Here we provide a short summary. An SOM, in

general, is a neighborhood preserving, adaptive vector quantizer that maps a high-D space non-linearly onto a 2-D lattice of artificial neurons such that the similarity relationships of the input data are reflected in the proximity relationships of the neurons representing the data. (2-D SOM lattice is most popular because high-D clusters will be visualized in 2-D. 1-D lattices are obviously more limited, and 3-D is more complicated to render.) Each neuron in the (usually) rectangular lattice has a weight vector of the same dimensionality as the input data, say  $n$ . The SOM learns the structure of the  $n$ -D data space by cycling through the following steps many times: 1) Competition. Pick an  $n$ -D input pattern,  $\mathbf{x}$ , randomly, find a “winner” neuron with index  $c$  such that  $c = \operatorname{argmax}_i \|\mathbf{x} - \mathbf{w}_i\|$  over all neuron indices  $i$  in the lattice. (The neuron whose weight vector is most similar to the input pattern wins the representation of that pattern.) 2) Cooperation. The winning neuron activates its neighbors. The neighborhood is defined by a neighborhood function,  $h(c(\mathbf{x}), i)$  which can be different in different SOM paradigms. In the conscience algorithm [7], the immediate adjacent neurons form the neighborhood. In the Kohonen SOM (KSOM), the neighborhood is defined by a Gaussian function and is decreased over time. 3) Synaptic adaptation. All neurons in the neighborhood, including the winner, adjust their weight vectors to become more similar to the input pattern. The winner typically moves its weight toward the input pattern to the greatest extent, while the neighbors adjust according to their distance to the winner (which is expressed in the neighborhood function).  $\mathbf{w}_j^{\text{new}} = \mathbf{w}_j^{\text{old}} + a * h(c(\mathbf{x}), i) (\mathbf{x} - \mathbf{w}_j^{\text{old}})$ , where  $a$  is a learning rate decreasing in time, and  $j \in h(c(\mathbf{x}), i)$ . The difference between the weight vectors of adjacent neurons in the converged SOM lattice is indicative of the similarity of the data points mapped to those neurons. Presetting of desired or maximum number of clusters is not needed for SOM formation. Cluster boundary detection is based on the analysis of the converged weights. See [3,4,5,8,9] for examples of real planetary cases.

## 2. Previous analyses of Mars Pathfinder IMP spectral images

The highly successful 1997 Mars Pathfinder mission landed on the Ares Vallis floodplain. The pre-landing expectation was that Mars Pathfinder would encounter a variety of rock types carried from the southern highlands by outflow channel floods. Instead, the initial interpretation of data from the Alpha Proton X-Ray Spectrometer (APXS) instrument and

the multispectral Imager for Mars Pathfinder (IMP) indicated that there was only a single rock type with various levels or types of coatings [10]. The fine grained materials making up the immobile and mobile surface layer (indurated soil and wind-blown dust) were also characterized on the basis of spectral characteristics [11]. Murchie et al. [12,13] and Bell et al. [14] recognized at least one other primary rock type besides the previously recognized “gray rock” spectral class. Where the “gray rock” spectral class had a shallow absorption at or beyond 1000 nm (the end of the IMP spectral coverage), this second rock spectral class, “black rock” was initially described by [12] as having a deep absorption centered on 930 nm. [14] noted some variability in the position of the band minimum for different occurrences of “black rock”. Murchie et al. in [13] suggest that the IMP data indicate a greater abundance of this “black rock” class in the far field, especially in the vicinity of the distant Twin Peaks and Big Crater that lie on the horizon.

The IMP consisted of two cameras that were physically separated in order to return stereo imagery. The spectral coverage of the system ranged from 440 to 1000 nm with FWHM (Full Width at Half Maximum) band widths ranging from 19 to 40 nm. The IMP collected 15 channels from the two cameras (with overlap in the 440 and 670 nm bands). Recent studies of IMP data have primarily made use of the “SuperPan” dataset. The Super Pan was acquired in eight separate image mosaics (“octants” labeled S0181 through S0188), with each mosaic consisting of many separate but contiguous camera azimuth and elevation pointing positions. The segments making up each octant were collected at disparate times leading to subtle differences between the segments in terms of lighting geometry and, potentially, calibration. [14] analyzed a version of the SuperPan calibrated by the U.S. Geological Survey in Flagstaff [15]. Residual differences between the image segments making up each octant were empirically corrected as described by [13] and it is this latest version of the SuperPan that was used in this study.

Previous analyses of the IMP SuperPan dataset have focused on spectral parameterization of the data [10], [11], [13], spectral mixture analysis of the data [14], and analysis of rock morphology in association with spectral parameters [13]. The SuperPan data set has not been subjected to a comprehensive classification. In the present work, we are seeking to classify the SuperPan data set through the use of a Self Organizing Map Artificial Neural Network architecture. In so doing we hope to address outstanding questions related to how much variability there is of primary rock spectral classes, and to classify

rock coatings and soils and see if there is any link to past climate regimes.

### 3. Spectral analyses of IMP images with Self-Organizing Neural Maps

The starting point of our analyses was the calibrated left and right eye SuperPan images described above. This study was done on octant S0184 [13]. The last, 967 nm, band of the right eye image of his particular octant had to be removed because of its poor quality. A brightness normalization (dividing each spectrum by its Euclidean vector norm, see, for example, [2] for more explanation) was performed prior to clustering in order to eliminate linear illumination geometry effects. While this also removes the geometric albedo, the relative spectral relations among the bands are preserved, resulting in much more uniform spectral patterns within each spectral unit. We also applied a mask of “no data” pixels to eliminate unnecessary processing of useless data. The S0184 images each contain in the order of half million multidimensional pixels. The wavelength suite for the left eye data is (443, 671, 802, 858, 898, 931, 968, 1003) nm, and (443, 480, 531, 600, 671, 752, 967) nm for the right eye image, (967 removed for the octant analyzed here).

#### 3.1. Self-Organizing Neural Maps in this study

For the specific analyses presented in this paper, a 40 x 40 SOM was used for each of the left and right eye clustering. An SOM does not require an estimated number of clusters, however, the combination of its size and the complexity of the data determines whether all possible clusters can be resolved, therefore we allowed generous space for the clusters to evolve. We chose the “conscience algorithm” SOM variant [7] because of its maximum entropy quantization property. A neighborhood of half-width 1 works well with this paradigm, and although it is more prone to partial organization (thus topology violations) than the KSOM, the size of the images as well as the size of the SOM is small enough that we can catch errors fairly easily with our tools. In return, the small constant neighborhood, speeds up the processing considerably. This algorithm is realized by the **ann-SOMconsc** module of HYPEREYE, and it can be run in a highly interactive mode with many visualization controls, or in batch mode. Approximately 7 million learning steps were performed for each image. After SOM convergence the **remap** semi-automated HYPEREYE module is used for cluster boundary extraction and color coating the clusters draped over a reference

image, for example a grey-scale band of the original image cube. (The reference image can also be a simple blank image.) The cluster map then serves as a mask to collect statistics such as mean, standard deviation, band extremes, etc. for each cluster, to examine the spectral characteristics of each unit.

#### 3.2. Results and Discussion

At the most basic level, the martian landscape observed by Pathfinder consisted of rocks and soils. As noted above, the current interpretation of the data recognizes two basic lithologies: gray rock and black rock. Gray rock's spectral (as measured by IMP) and chemical characteristics (as measured by the Pathfinder's APXS instrument) are consistent with an andesite or basaltic andesite [10]. The spectral characteristics of black rock are consistent with a more mafic basalt [11,13]. Rocks in the scene are also coated to varying extents and the soils consist of mobile, fine grained, soils (“bright red soil”); indurated soils (“brown soil”) and soils that appear to have some admixture of more mafic components (“dark soil”) including some that appear to contain mechanically weathered debris from rocks (“rock soil”). The maps produced in this work consist of a number of spectral classes which can be related to previously recognized classes, others which can be interpreted as spectral mixtures of some classes, and a few classes of special interest which are tentatively identified as previously unrecognized spectral classes. In both the right eye and left eye class maps, bright red dust is accurately mapped out as class H (orange). The “dark soil” class [11] is also well mapped in the left eye as class “b” (cyan). The rock soil class is well mapped in the right eye data by class Y (maroon) and generally well mapped in the left eye data by the same class. The gray rock class is very well mapped in the right eye data as class D, E and Z (yellow, pale yellow and green-yellow), mapping of gray rock in the left eye data is complicated by data dropouts and subtle calibration differences between component segments of the octant. An interesting class apparent in both maps is class M (purple) which, in the left eye data has its most apparent expression on the flanks of the west Twin Peak (on the horizon, upper right). This class was interpreted by Murchie et al. [13] as part of the black rock class, but the spectra in Figure 3 indicate that while both classes have an absorption near 930 nm the band depth of the black rock class is significantly deeper than that of class M (purple). The black rock class itself is mapped here as two classes. The first, class O, corresponds most closely to the original

description of black rock with an absorption centered on 930 nm which could be consistent with a predominantly orthopyroxene mineral assemblage. The second, class R, is different from the original description of the black rock class in that it has a deep absorption at or beyond 1000 nm which would be more consistent with a predominantly clinopyroxene, or with a mixed clinopyroxene and olivine mineral assemblage. These black rocks with a longer band center were also noted in Bell et al. [11].

The two extremely small clusters, O and R, contain some spurious pixels that contaminate the average class characteristics (more so for O than for R). This may be due to insufficient size of the SOM that can be helped in follow-up processing. However, we note that several contiguous patches highlighted as O are indeed close spectral matches to the prototype. It is also noteworthy that the class R black rock is also well mapped in the right eye data.

#### **4. HYPEREYE, the supporting software**

The HYPEREYE data management and programming system for analysis of hyperspectral imagery with Self-Organizing Maps and SOM-hybrid neural nets is the specialized tool our group has developed for both algorithm research and data analysis. The system has three different facets. It is first a collection of analysis tools for viewing and reducing hyperspectral imagery and the corresponding SOM products. The tools are both modular, and designed to allow easy transition to platform independence, planned for the near future. It is also a data organization system for storing the imagery and neural networks and tracing the history of their analysis. Lastly, it is also a programming environment for algorithm research by our group.

The analysis tools comprise a variety of home-grown applications and commercial software libraries. The Khoros data storage, manipulation, and visualization libraries manage the large (multi-megabyte) dataset reading, writing, and display interaction. The NeuralWorks Designer Pack libraries store, manipulate, and perform learning on the SOM neural networks we construct. Both libraries are available on a variety of UNIX and PC platforms. HYPEREYE adds its own set of libraries for creating graphical displays of the SOM networks and for creating and using worksheet files that keep the history of each tool's manipulation of the data. Together, the system is a data-analyst-friendly graphical window environment that receives raw spectral data and delivers statistical results and map and plot-based data.

The data organization design uses a "project" metaphor with database-like rules on top of the UNIX file system's permissions rules. The initiation and signing-in to a "project" creates a directory structure, file naming convention, and access rules that allow one user consistent control over the "project's" data analysis. Working in the "project" produces worksheet files and data products (which are reproducible from the worksheets). Data, worksheet, and output files can all be accessed by the user with conventional methods (i.e. text editor or graphic viewer) as well. But while access to the "database elements" directly (by-passing the "project") through the file system is allowed, built-in safeguards ensure that it is not detrimental to the "project's" integrity.

Neural network algorithm design is an important facet of our research pursuits and the HYPEREYE application and library environment facilitates this development among students, staff, and faculty. Each part of the network analysis system has hooks for standard C (or Fortran) functions which isolate the researcher from the mundane data handling and graphics presentation. For instance, the SOM learn algorithm by [6] was recently implemented for two-dimensional SOM's by one graduate student. The code is three functions and less than 120 lines due, in part, to the use of NeuralWorks' library network-handling. Another example is the easy implementation of cluster boundary functions for helping the analyst identify clusters in the SOM. A future goal is to make these development additions separate plug-ins to further simplify the maintenance of the HYPEREYE environment versus expanding its functionality through various research efforts.

#### **5. Conclusion and Next Steps**

We have presented a comprehensive spectral classification of the S0184 octant of the Imager for Mars Pathfinder SuperPan data, for both the left and the right eye images, using unsupervised neural clustering technique. The results are generally consistent with earlier analyses, and also produced refinements that may be variations in mineralogy or rock coatings.

This data set appears extremely varied, possibly in part due to noise and calibration issues. A larger size SOM could be used to accommodate more of the spurious spectra instead of being forced to assigning them to neurons that map valid clusters. Specifically, for clusters O and R we will repeat the analysis with a larger SOM, and also with forced negative magnification, to make the detection of these rare

classes cleaner. We also plan to look more closely at class M in order to examine its relation to class M pixels in the near field and, potentially, to other classes with similar spectra. The effect of atmospheric scattering on the observed reflectance of these far field regions also needs to be considered more fully.

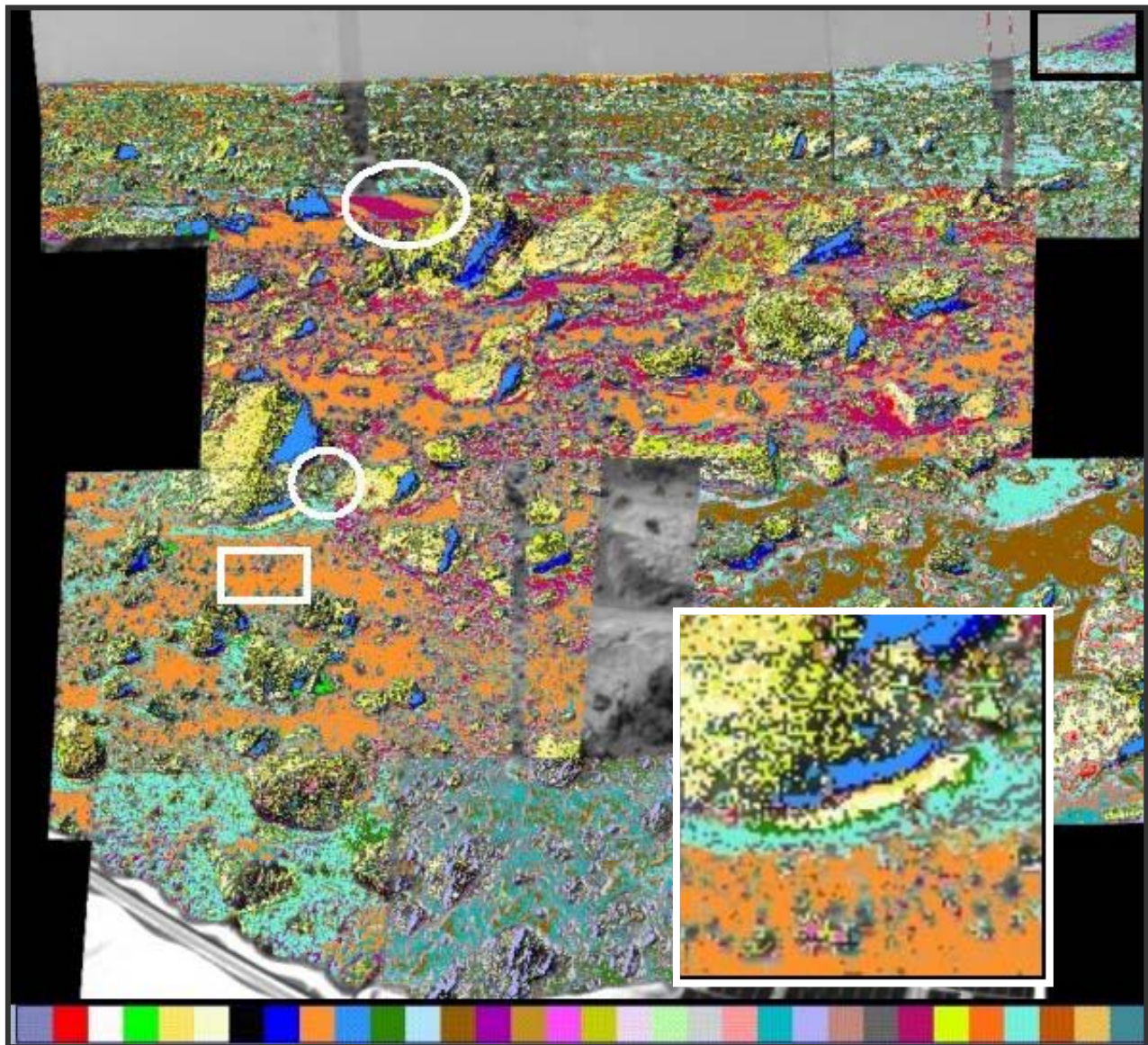
Lastly, mosaic segments with imperfect calibration in the left eye octant become highlighted because the same soil type appears as spectrally different across those mosaics. The statistics on these spectral anomalies may help track down the nature of the calibration problem. Additional octants in the SuperPan will also be analyzed.

## 6. Acknowledgements

This work is partially supported by the Applied Information Systems Research Program, NASA/OSSA NAG5-9405 and NAG9-10432 for SOM/HYPEREYE research and by NASA/OSS Mars Data Analysis Program NAG5-13294 for work on the IMP SuperPan data. We thank Dr. Scott Murchie of the Johns Hopkins Applied Physics Lab for providing the SuperPan imagery.

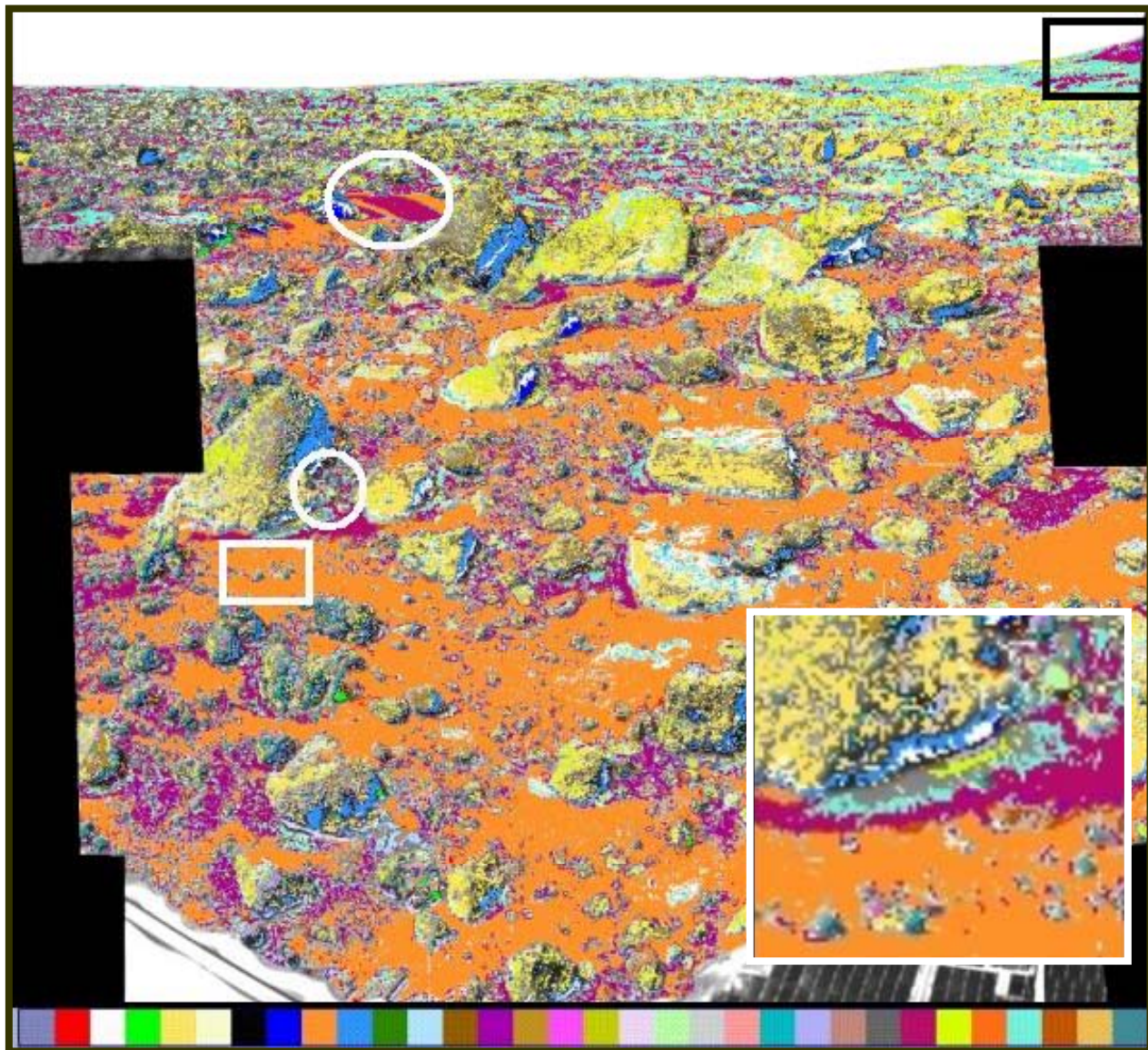
## 7. References

- [1] Kohonen, T., *Self-Organizing Maps*. Springer Series in Information Sciences, **30**, Springer, Berlin, Heidelberg, New York, 1995, 1997.
- [2] Merényi, E., Singer, R.B., Miller, J.S. "Mapping of Spectral Variations On the Surface of Mars From High Spectral Resolution Telescopic Images", *Icarus* **124**, 1996, pp. 280-295.
- [3] Howell, E.S., Merényi, E., Lebofsky, L.A., "Classification of Asteroid Spectra Using a Neural Network" *Jour. Geophys. Res.* **99**, 1994, pp. 10,847--10,865.
- [4] Merényi, E., Howell, E.S., et al., Prediction of Water In Asteroids from Spectral Data Shortward of 3 Microns". *Icarus* **129**, 1997, pp. 421—439.
- [5] T. Villmann and E. Merényi, "Extensions and Modifications of the Kohonen Som and Applications in Remote Sensing Image Analysis". In *Self-Organizing Maps: Recent Advances and Applications*, (U.Seiffert and L.C. Jain Eds.), Springer-Verlag, 2001, pp. 121-145.
- [6] Bauer, H.U., Der. R., Herrmann, M. "Controlling the Magnification Factor of Self-Organizing Feature Maps". *Neural Computation* **8**, 1996, pp. 757-771.
- [7] DeSieno, D., "Adding a Conscience to Competitive Learning". *Proc. ICNN*, New York, July, 1988, **I**, pp. 117-124.
- [8] Merényi, E., "Self-Organizing ANNs for Planetary Surface Composition Research", *Proc. 6th European Symposium on Artificial Neural Networks, ESANN'98*, Bruges, Belgium, April 22-24, 1998, pp. 197-202.
- [9] Merényi, E. "Precision Mining of High-Dimensional Patterns With Self-Organizing Maps". In *Quo Vadis Computational Intelligence: New Trends and Approaches in Computational Intelligence, Studies in Fuzziness and Soft Computing*, Vol. 54, Peter Sincak, Jan Vascak, Eds, Physica-Verlag, 2000.
- [10] McSween, H.Y. Jr. et al. "Chemical, multispectral and textural constraints on the composition and origin of rocks at the Mars Pathfinder landing site". *J. Geophys. Res.* **104**, 1999, pp. 8679-8716.
- [11] Bell, J.F. III et al., "Mineralogic and compositional properties of Martian soil and dust: Results from Mars Pathfinder". *J. Geophys. Res.* **105**, 1999, 2000, pp. 1721-1756.
- [12] Murchie, S.L., O. Barnouin-Jha, J. Johnson, J.F. Bell III; H. McSween, and R.V. Morris, "Diverse Rock Types at the Mars Pathfinder Landing Site", 31st Annual Lunar and Planetary Science Conference, March 13-17, 2000, Houston, Texas, 126 (abstract).
- [13] Murchie, S.L. et al., "New Insights into the Geology of the Mars Pathfinder Landing Site from Spectral and Morphologic Analysis of the IMP 12-color Super Panorama", submitted to *Icarus*, June 2003.
- [14] Bell, J.F., W.H. Farrand, J.R. Johnson, and R.V. Morris "Low abundance materials at the Mars Pathfinder Landing Site: An investigation using spectral mixture analysis and related techniques". *Icarus*, **158**, 2002, pp. 56-71.
- [15] Johnson, J. R., T. Sucharski, and R.J. Reid, "Implementation of the Image for Mars Pathfinder Calibration Algorithms in ISIS", 32nd Annual Lunar and Planetary Science Conference, March 12-16, 2001, Houston, Texas. 2062 (abstract).
- [16] Rasure, J., and M. Young, 1992, "An Open Environment for Image Processing Software Development", *Proc. Of the SPIE/IS&T Symposium in Electronic Imaging, February 14, 1992*, vol. 1659.
- [17] NeuralWare, Inc. *Neural Computing*, A Technology Handbook; and Designer Pack, for NeuralWorks Professional II/Plus. NeuralWare, Inc., Pittsburgh, 2001.



bg A B C D E F G H I J K L M N O P Q R S T U V W X Y Z a b c d e

Figure 1. Spectral clusters identified with a Self-Organizing Map in IMP SuperPan octant S0184, left eye image. The color-coded clusters are draped over a grey-scale reference image. The unmapped grey scale patches (such as the rectangle spanning vertically in the center and the black fringes framing the image) are areas where data fallout occurs in one or more of the eight left-eye spectral bands. Some of the mosaics (most notably the two center bottom mosaics) show calibration problems. The most interesting clusters, M, Y, D, O, and R are discussed in the text. Clusters H and b generally map the known distribution of bright red dust (H) and dark soil (b). E, D and Z correspond to rock surfaces and may indicate variations in the coatings of the rocks. An occurrence of each of the black rock classes O and R is highlighted by the small white circle and the white rectangle, respectively. The inset contains both areas with the O black rock showing at the bottom, center, the R patch showing close to the upper right on the diagonal. The large oval and the black rectangle point out an example of the "rock soil", and the class M soil, mostly covering Twin Peaks. (Color copy downloadable at [www.ece.rice.edu/~erzsebet](http://www.ece.rice.edu/~erzsebet) Publications, or request one in email.)



bg A B C D E F G H I J K L M N O P Q R S T U V W X Y Z a b c d e

Figure 2. Spectral clusters identified with a Self-Organizing Neural Map in IMP SuperPan octant S0184, right eye image. The unmapped frame indicates no-data regions. The spatial distribution of clusters in this right-eye map is remarkably similar to the left eye map in Figure 1 in spite of the fact that the left and right eye wavelength sets have only two points (443 and 671 nm) in common, and that most left eye bandpasses are between 671 – 1003 nm while most right eye bandpasses fall in the 443 – 671 nm range. This indicates that a number of characteristic mineralogies manifest in both the VIS and the NIR wavelength regions. Ovals and rectangles highlight the same surface types as in Figure 1. The inset likewise is an enlargement of the area that contains an occurrence of each of the black rock units O and R. (Color copy downloadable at [www.ece.rice.edu/~erzsebet](http://www.ece.rice.edu/~erzsebet) Publications, or request one in email.)

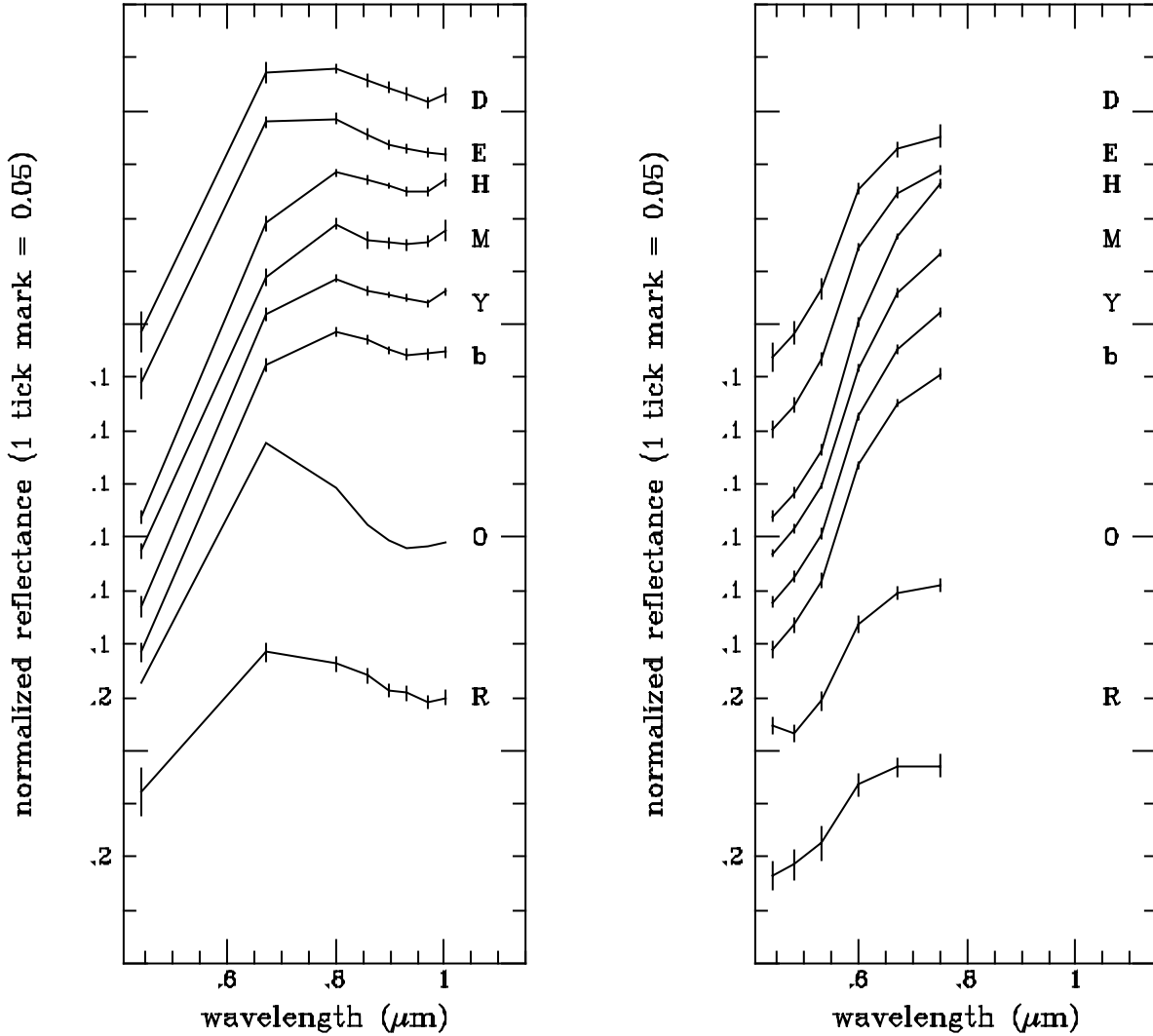


Figure 3. Average spectral characteristics of several of the clusters identified with a Self-Organizing Neural Map in SuperPan octant S0184 and displayed in Figures 1 and 2. The left frame here shows the left eye spectra and the right frame shows the spectra of corresponding spatial units in right eye image. The standard deviation of each cluster is indicated by the vertical bars. The spectra are offset for clarity and plotted on an arbitrary but consistent vertical scale. Left and right eye data were roughly scaled together for this viewing. Precise scaling for the entire data set is not possible presently because of calibration uncertainties. Class O contained some spurious pixels that may be due to insufficient SOM space, therefore we show only the average of the successfully identified "O" type black rock spectra.