Self-Organizing ANNs for Planetary Surface Composition Research *

Erzsébet Merényi †

University of Arizona, Lunar and Planetary Laboratory, PIRL
Tucson, AZ 85721, U.S.A.

Abstract. The mineralogic composition of planetary surfaces is often mapped from remotely sensed spectral images. Advanced hyperspectral sensors today provide more detailed and more voluminous measurements than traditional classification algorithms can efficiently exploit. ANNs, and specifically Self-Organizing Maps, have been used at the Lunar and Planetary Laboratory, University of Arizona, to address these challenges.

1. The challenge in spectral images

Surface composition of planets, including our Earth, is often investigated from remotely sensed spectral images. These are "stacked" images of the same spatial area, each taken at a different wavelength. The individual images are called image bands. Surface reflectance spectroscopy usually includes the visible and near infrared (VIS-NIR) wavelength range, typically from 0.4 to 2.5 μm. *Hyperspectral* sensors, developed in the past 5–10 years, acquire as many as 100–500 image bands simultaneously, contiguously covering a given window of the electromagnetic spectrum at very small wavelength increments. The vector $S^{x,y} = (S_1^{x,y}, ..., S_{NB}^{x,y})$, where $S_k^{x,y}$ is the data value in the kth image band $(k = 1, ..., NB)$ at pixel location (x, y), is called a spectrum. It is a characteristic pattern which provides a clue to the surface material(s) within pixel (x, y). The feature space spanned by VIS-NIR reflectance spectra is $[0, U]^{NB} \subset \Re^{NB}$ where $U > 0$ represents an upper limit of the measured scaled reflectivity. Sections of this space can be very densely populated while other parts may be extremely sparse, depending on the materials in the scene and on the spectral resolution of the sensor. The most advanced hyperspectral imagers such as AVIRIS (NASA/JPL [1]), Hydice, the NIMS and VIMS on the Galileo and Cassini planetary probes, and others (http://polestar.mps.ohio-state.edu/~csatho/wg35.html, under "Hyperspectral Imaging"), can resolve the detailed spectral features that are known to characterize minerals and vegetation, from laboratory measurements. Classification of these intricate spectral patterns poses special challenges because of any combination of the following:

*Work funded by NASA, Applied Information Systems Research Program, NAG54001; Paper in Proceedings of European Symposium on Artificial Neural Networks, ESANN98, Bruges, Belgium, April 22–24, 1998, pp 197–202.

†email:erzsebet@pirl.lpl.arizona.edu for color reprint

- The patterns are high dimensional (dozens $\leq NB \leq$ hundreds);
- The number of data points (image pixels) can be very large, $\simeq$ 1M;
- Given the richness of data, the goal is to separate many cover classes;
- Geologically different materials may be distinguished by very subtle differences in their spectral patterns;
- Very little training data may be available for some classes; and classes may be represented very unevenly.

Additional complications arise from atmospheric distortions, noise, illumination geometry and albedo variations in the scene. These are properly dealt with prior to classification, but are not discussed here.

2. How Self-Organizing Maps help

Dimensionality reduction is often applied in order to accomodate these spectral data to conventional calssifiers, or to Backpropagation (BP) networks. [2] and [3], for example, classified AVIRIS spectral images for geological mapping, both reducing the original 224 bands to 32, to achieve training convergence with BP. This, however, can lead to a loss of information that negates the purpose of sophisticated hyperspectral sensors. Instead, a better approach is to taylor the tool to the sophistication of the data for effective exploitation. Although rigorous proofs on convergence and uniqueness have yet to come forth for high-dimensional data, 1- and 2-D analyses (*e.g.,* [4–6]), and powerful results using SOMs for large and complex data sets [7] give us confidence in our choice of SOM-based approaches. The clustering power of the SOM, its generalization from small amounts of training data and its relative ease of handling high data dimensionality helped us extract more scientific knowledge from Martian, terrestrial and asteroid data than was obtained earlier with conventional methods or BP [8–12]. Two examples will be given below.

NeuralWare's implementation of Kohonen's SOM was used [13]. It is based on [14–16], and is algorithmically similar to [17], including a choice of (a rectangular or a diamond shaped 'bubble' type) neighbourhood topology. The neighbourhood size is constant in time; however, a "conscience" mechanism [15] adjusts the winning frequencies of the PEs so as to ensure proper spreading (compaction) of information that is dense (sparse) in the input feature space. Update of the weight vectors, in [17] notation, is given by $m_i^{t+1} = m_i^t + \alpha^t(S - m_i^t)$, for $i \in N_c$ and $m_i^{t+1} = m_i$ otherwise, where c is the index of the winning PE_c, N_c denotes the index set of the neighbourhood of PE_c. $c = arg\,\min_i(\|S - m_i\| - B_i)$ where the bias term $B_i = \gamma(1/M - F_i^t)$ adjusts the distance between the incoming pattern S and PE_i based on the historic winning frequency F_i of PE_i. M is the number of PEs in the Kohonen lattice. F_i is updated along with the weights, according to $F_i^{t+1} = F_i^t + \beta(\delta_{N_c} - F_i^t)$ where δ_{N_c} is 1 for $i \in N_c$, 0 otherwise. α, β and γ are user controlled parameters decreasing in time.

A single categorization layer with Widrow-Hoff learning is coupled with the SOM, for optional supervised training after SOM convergence, with the SOM output as its input. The preformed clusters prevent learning of inconsistent class labels. This is extremely helpful in the selection/verification of consistent training samples which is difficult for many classes with subtle differences.

2.1. Discovery of new taxonomical groups of asteroids

Here we were presented with a small set (130 spectra) of diverse ($\sim$ 15) types of asteroids, observed in the 0.4–1.0 and 0.8–2.5 μm spectral windows. From the earlier, 0.4–1.0 μm data a compositional taxonomy was established by minimum tree clustering. When the 0.8–2.5 μm data became available no further structure was found, using PCA. S (silicate) type asteroids had long been believed to have olivine- and pyroxene-rich subgroups (indicators of primary evolutionary states), which are characterized by the shapes of the 1- and 2-μm absorptions in the spectra. Yet formal identification eluded the conventional approaches. We clustered the joined, 0.4–2.5 μm spectra with an SOM and found the S type asteroids falling consistently into 3 subgroups. The two end groups were identified as olivine- and pyroxene-rich [8] [11], So and Sp in Figure 1.

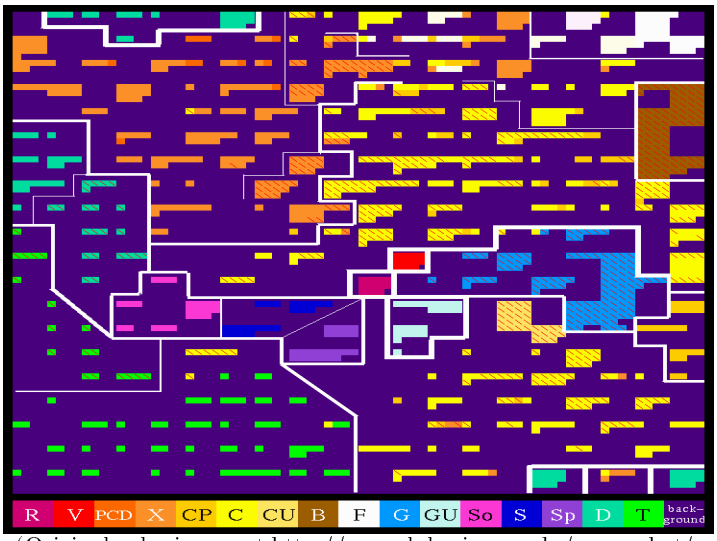

Figure 1. (Original color image at http://www.lpl.arizona.edu/~erzsebet/emast.html) SOM of 13-D asteroid spectra, from [11]. Cluster boundaries were extracted by the u-matrix method [18], and hand-drawn for clarity. Line widths indicate separation strengths. The S asteroids, just southwest of the center enclosed by a thick line, comprise the So, S (intermediate), and Sp subclasses separated by thinner lines. Each small square represents one asteroid spectrum. Max. 4 x 4 spectra are visualized for any of the 20 x 20 Kohonen PEs. For this clustering, 10 additional, equally possible spectra were generated for each original one, based on the measurement uncertainties.

After finding these and other interesting new clusters, the SOM-based hybrid ANN, described above, was used to "clean up" the previous taxonomical designations. Given the small, diverse data set, "jackknifing" was employed for this

classification, which resulted in an extremely low number of training samples in each round. Recall success rate based on training with our adjusted class labels increased to 90%, while forcing the old labels resulted in 70% success [8]. In comparision, a 3-layer BP network (with no SOM layer) "memorized" the individual training samples and no consistent prediction could be achieved.

2.2. Hyperspectral image exploitation

Figure 2 shows a classification map of 23 surface covers for a volcanic field in the Nevada desert, from an AVIRIS image, produced by the above-described hybrid ANN architecture, with 194 input, 40 x 40 SOM, and 23 output PEs. The full spectral resolution was used except for elimination of the saturated atmospheric water-bands, thus retaining 194 of the original 224 bands. This 512 x 614 pixel image of a 10 x 12 km^2 area, comprising 140 Mb of data, is considered moderate size. About 30K unsupervised and 20K supervised steps sufficed to train our hybrid network to achieve very good classification accuracy [12]. Coherent unclassified regions indicate possible additional classes.

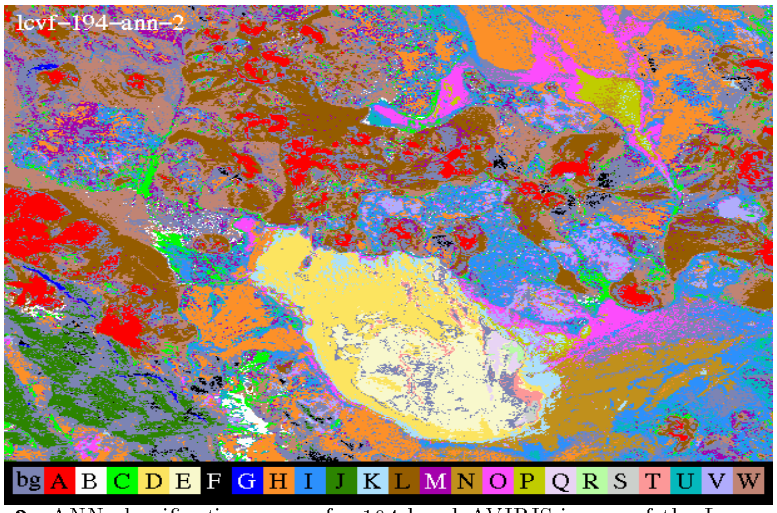

Figure 2. ANN classification map of a 194-band AVIRIS image of the Lunar Crater Volcanic Field, Nevada, U.S.A., into 23 geologic cover types, from [12]. Original color image at http://www.lpl.arizona.edu/~erzsebet/emearth.html. No dimensionality reduction was done before classification. The spectral detail in the data allows distinction of many relevant classes in the scene, including volcanic cinders (A) and weathered derivatives (L, W), young and old basalts (F, G, I), clay-containing dry playa and outwash materials (E, D, Q – T) and evaporites (N – P). Classes are described in [12].

Conventional classifiers, notably covariance based ones such as Maximum Likelihood (ML) may do a superb job on Landsat images or other low-dimensional data. However, in remote sensing, finding $NB + 1$ reliable training samples (as needed by ML), for a large number of classes, NC, is an irrealistic requirement. Some relevant geologic classes may not contain NC pixels; and identification of $NC * (NB + 1)$ samples can be prohibitively expensive, or impossible. For this example, 23 x 195 = 4485 field locations would need to be verified to obtain

training material for ML classification. A BP network may be too difficult to train with 194 input and 23 output nodes, and it may not generalize well from small amounts of training pixels. For the classes B, Q, R, S, T, less than 10 training samples could be identified confidently. The entire training set for the 23 classes comprises only $\sim$ 900 samples. [12] shows that reduction of the above 194 bands to 32 (or less) in order to apply a ML classifier, eliminates much of the subtle discriminating features (Figure 3), resulting in a loss of 7 classes, and severe under- or overestimation of the rest of the classes. (This exercise was done with a strategic de-selection of bands to retain as much as possible of the compositional representation of the relevant materials.) The accuracy of the ANN classification allows examination of mineralogic properties from average (less noisy) class spectra rather than from single samples which is a current practice. This was especially valuable for rare Martian data [9–10].

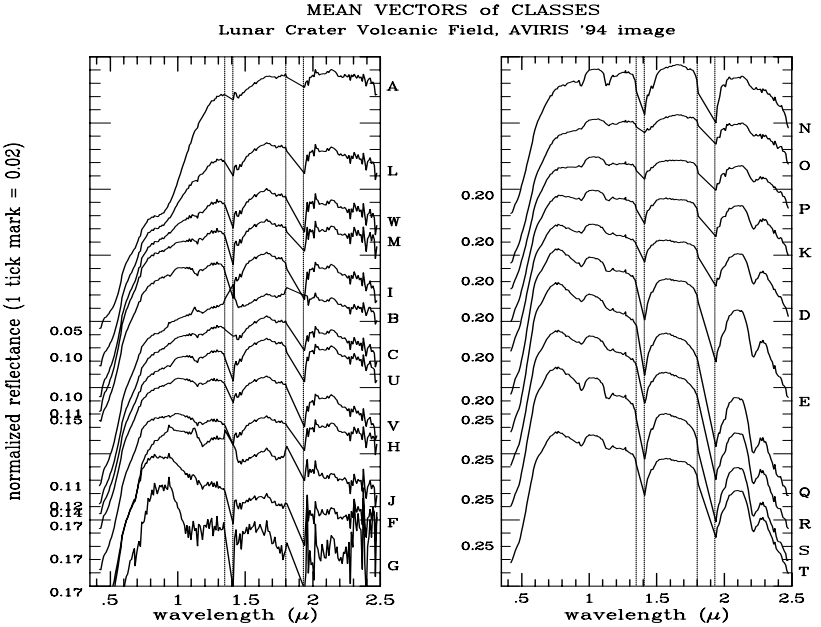

Figure 3. Mean spectral signatures of the classes in Figure 2. illustrate some of the subtle differences among meaningful and spatially coherent geological cover types. Classes E, Q–T, for example, contain increasing amounts of water and clay. Spectra are offset for clarity. The dotted vertical lines indicate data fallout in the water bands.

3. Future work

Automation of detection/visualization/interpretation of clusters, similarly to [17] but with a focus on the specifics of large (>100 Mbytes) hyperspectral images, is underway. This will improve our capabilities to investigate the useful information content of, and SOM behaviour for, (hyper)spectral images, discover new knowledge more efficiently. The ultimate goal is parallel hardware implementation for envisioned on-board, real-time (and fast, on-ground) analyses.

References

1. Green, R.O., Ed.: *Summaries of 6th Annual JPL Airborne Geoscience Workshop, Vol. 1. AVIRIS Workshop.* Pasadena, CA, March 4–6, 1996 (1996)
2. Benediktsson, J.A., Sveinsson, J.R., Arnason, K.: Classification of Very-High-Dimensional data With Geological Applications. *Proc. MAC Europe 91, Lenggreis, Germany, 4–6 October, 1994*, 13–18 (1994)
3. Paola, J.D.: *Neural Network Classification of Multispectral Imagery.* M.Sc. Thesis, University of Arizona (1994)
4. Cottrell, M., and Fort, J.C.: A Stochastic Model of Retinotopy: A Self Organizing Process. *Biol. Cybern.,* 53, 405–411 (1986)
5. Cottrell, M., Fort, J.C., and Pagès, G.: Comment about 'analysis of the convergence properties of topology preserving neural networks'. *IEEE Trans. on Neural Networks,* 6(3), 797–799 (1995)
6. Flanagan, J.A.: Analysing a Self-organising Algorithm. *Neural Networks,* 10(5), 875–883 (1997)
7. Kohonen, T.: Exploration of Very Large Databases by Self-Organizing Maps. *Proc. IEEE ICNN'97,* I, PL1–6 (1997)
8. Howell, E.S., Merényi, E., Lebofsky, L.A.: Classification of Asteroid Spectra Using a Neural Network. *J. Geophys. Res.,* 99, 10,847–10,865 (1994)
9. 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, 280–295 (1996a)
10. Merényi, E., Edgett, K.S., and Singer, R.B.: Deucalionis Regio, Mars: Evidence For a New Type of Immobile Weathered Soil Unit. *ICARUS,* 124, 296–307 (1996b)
11. Merényi, E., E.S. Howell, L.A. Lebofsky, A.S. Rivkin: Prediction of Water In Asteroids From Spectral Data Shortward of 3 Microns. *ICARUS,* 129, 421–439 (1997a)
12. Merényi, E., Taranik, J.V., Minor, T.B., and Farrand, W.H.: Quantitative Comparison of Neural Network and Conventional Classifiers For Hyperspectral Imagery. Submitted to *Remote Sens. Environ.* (1997b)
13. NeuralWare, Inc.: *Neural Computing, A Technology Handbook for NeuralWorks Professional II/Plus.* NeuralWare, Inc., Pittsburgh (1993)
14. Kohonen, T.: *Self-Organization and Associative Memory.* Springer-Verlag, New York (1988)
15. DeSieno, D.: Adding a Conscience to Competitive Learning. *Proc. ICNN, New York, July 1988,* I, 117–124 (1988)
16. Willshaw, D.J., von der Marlsburg, C.: How Patterned Neural Connections Can Be Set Up by Self-Organization. *Proc. R. Soc., London* , 194, 432–445 (1976)
17. Kohonen, T., Hynninen, J., Kangas, J., and Laaksonen, J.: SOM-PAK: The Self-Organizing Map Program Package. *Helsinki University of Technology,* Report A31, 1–25 (1996)
18. Ultsch, A., and H. P. Simeon: Kohonen's Self Organizing Feature Map for Exploratory Data Analysis. *Proc. INNC-90-PARIS,* I, 305–308 (1990)