Intelligent Understanding of Hyperspectral Images through Self-Organizing Neural Maps

Erzsébet Merényi
Electrical and Computer Engineering, Rice University MS 380
Houston, TX 77005, U.S.A.
e-mail: erzsebet@rice.edu

Abstract

Hyperspectral imagers are present on most current and future Earth-orbiting satellites and space missions, acquiring immense amounts of rich data. Many clustering and classification methods meet their theoretical or practical limitations when confronted with such high-dimensional data. The most popular neural map paradigm, the Self-Organizing Map (SOM) has successfully been used for over twenty years for a vast variety of data mining problems. One attractive feature of SOMs is that they handle high-dimensional data well. In the past several years new theories provided critical evaluation and insight to Kohonen's original SOM revealing suboptimal quantization properties and offering improved ways of data mining with neural maps. This short paper reviews our recent contributions provoked by the new theoretical advances: numerical simulations and new heuristics in support of more intelligent SOM clustering and more effective extraction of the learned cluster boundaries, which are crucial for intelligent, automated understanding of large amounts of intricate high-dimensional data.

Keywords: Intelligent Data Understanding, Self-Organizing Maps, hyperspectral imaging, data mining, classification

I. ADVANCED VARIANTS OF SELF-ORGANIZING MAPS

Unsupervised clustering and supervised classification are both fundamental to the understanding and modeling of the structure of data. High quality clustering can provide strong support for subsequent labeling and supervised classification. In this paper we concentrate on unsupervised clustering by neural learning, highlighting aspects that respond to specific challenges inherent in high-dimensional spectral images.

Learning by Kohonen's original Self-Organizing Map (KSOM) [1], [2] can be summarized as follows: Let $V \subset \Re^d$ and A denote the d-dimensional input data manifold and the regular grid of SOM Processing Elements (PEs, nodes, or neurons), respectively. (A is typically a rectangular lattice.) The SOM PEs are indexed by their (potentially multi-dimensional) grid locations $\mathbf{r}$. The weight attached to node $\mathbf{r} \in A$ is $\mathbf{w_r} \in \Re^d$. For any $\mathbf{x} \in V$ input the KSOM algorithm selects a winner node $\mathbf{s}$ by

$$\mathbf{s} = \underset{\mathbf{r}}{argmin}||\mathbf{w_r} - \mathbf{x}|| \quad (1)$$

and then updates the weights according to

$$\Delta\mathbf{w_r} = \epsilon h_{\mathbf{rs}}(\mathbf{x} - \mathbf{w_r}) \quad (2)$$

where ϵ is constant, the *neighborhood function* $h_{\mathbf{rs}}$, typically a Gaussian function centered over the winner node, defines which weights get updated and to what extent. SOMs perform topology preserving vector quantization, where the optimal placement of the prototypes (the SOM weights) is learned by adaptation during the above iterational learning process. The topology preserving nature of the quantization sets SOMs apart from other clustering techniques. This process results in an ordered set of the learned prototype vectors in A, according to their similarity relations, and forms areas of PEs in the SOM grid that collectively represent groups of similar input vectors.

Let P and Q denote the *pdf* of V and the *pdf* of the converged SOM weights in V, respectively. It has been proven that P and Q are related by the power law

$$Q(\mathbf{w}) = constant * P(\mathbf{w})^{\alpha} \quad (3)$$

where α is the *magnification exponent* [3], [4], and that some values of α carry particular quantization or information theoretical properties. $\alpha = 1$ maximizes information theoretic entropy, therefore such an SOM produces the best (information theroretically optimal) approximation to the *pdf* of the data with the given number of SOM weights. $\alpha = d/(d+p)$ corresponds to minimum mean squared error quantization of d-dimensional data in p-norm [4]. $\alpha < 0$ enlarges SOM response areas for low-frequency inputs, which increases the chance of the detectability of *unknown* rare classes. Kohonen originally assumed that the KSOM produced an $\alpha = 1$ mapping. It was shown later that the inherent property of the KSOM is a map magnification of $\alpha = 2/3$ [5], [6], which is optimal in neither minimum distortion nor maximum entropy sense. This makes the faithfulness of the KSOM uncertain and, especially for large and complicated data with many clusters, decreases confidence in its potential to discover small groups of data that may be the most important targets when sifting through huge data streams. As the need increases for distilling comprehensive data volumes for decision making or for the pursuit of scientific unknown, the demand for more sensitive and precise understanding of the data rises too. This demand brings some of these more recently analyzed SOM properties into focus.

Faithful matching of the *pdf* of the input data ($\alpha = 1$), negative, or any other than 2/3 magnification cannot be achieved with KSOM because it has no explicit control over the magnification exponent. A SOM variant called Conscience algorithm [7] was conceived to realize maximum entropy mapping ($\alpha = 1$) through *implicit* control of the magnification by adjusting the winning frequencies of individual PEs with heuristics. Theoretical proof of the achieved map magnification does not exist for the Conscience algorithm, but it works very well in practice. In

recent work we gave numerical verification of its $\alpha = 1$ property [8] on 6-dimensional synthetic data. However, the Conscience algorithm cannot induce any other value of α.

A promising theoretical approach was put forward in [3] for explicit magnification control of SOMs. We will refer to this method as BDH after the initials of the authors. In equation (2) the constant learning rate ϵ is globally defined, *i.e.,* it is the same for all PEs for a given time step and its value is independent of any local properties of the map. The BDH is based on making the learning rate dependent on the local density, that is, effecting

$$\epsilon_{\mathbf{r}} \propto \epsilon_0 P(\mathbf{w_r}) \tag{4}$$

This is achieved by changing ϵ in equation (2) to

$$\epsilon_{\mathbf{s}}(t) = \epsilon_0 \left(\frac{1}{\Delta t_{\mathbf{s}}} \left(\frac{1}{||\mathbf{w_s} - \mathbf{x}||^d} \right) \right)^m \tag{5}$$

where m is a free parameter, t is the time step and $\Delta t_{\mathbf{s}}$ is the time difference since the PE $\mathbf{s}$ won last. ϵ_0 is a constant, and d denotes the effective dimensionality of the receptive field of $\mathbf{w_s}$. (The receptive field of an SOM weight is the subset of data points that are mapped to that weight.) $\epsilon_{\mathbf{s}}(t)$ is locally determined but then applied to *all* weight updates in the current step. By doing so the local property of the map is propagated to the neighbors. With a derivation similar to that in [5] where the $\alpha = 2/3$ property of the KSOM was proven, [3] showed that the learning rate in equation (5) modifies equation (3) to

$$Q(\mathbf{w}) = constant * P(\mathbf{w})^{\alpha'} = P(\mathbf{w})^{\frac{2}{3}(1+m)} \tag{6}$$

where the free parameter m can be used for controlling the value of α'. Therefore, to achieve a desired magnification exponent α', $m = 3/2 * \alpha' - 1$ should be used in equation (5). The reader is referred to [3] for further details.

The theoretical proof of the BDH, unfortunately, is limited to 1-dimensional data and to n-dimensional data whose components are statistically independent (the *pdf* separates into the marginals). Most real data do not meet these conditions, which inspired our recent numerical simulations to chart the behavior of the BDH for "forbidden" data. A systematic assessment in [9] and [8] shows validity up to moderately high-dimensional data, as well as dramatic improvement in the detectability of rare species compared to KSOM. Below we present an example.

Topology preserving mapping of the input manifold V to the output grid A is crucial for correct identification of learned clusters. Topology preservation can be violated for a number of reasons: small SOM grid size, large dimensional mismatch between V and A, too fast learning, etc. It is possible to recognize topology violations after clusters are extracted from the learned SOM but for large and complex data this is costly. A better way is to use a recently proposed measure of topology preservation, the Topographic Function [10]. This measure, which was originally evaluated on moderate-dimensional data (1-4 dimensions), is under our current investigation for high-dimensional data. We show a preliminary application on hyperspectral data below.

Extraction of the boundaries of the learned clusters can be difficult when the data set is rich. To date fully automated tools do not exist. Since the knowledge about cluster separation is largely contained in the learned weights cluster boundary capture is often attempted through various visualizations of the weight distances between PEs. The frequently used approach by [11], [12] works well for smaller data sets with moderate number of clusters, and the manual labor is tolerable. Large size and high complexity complicates this task. A natural approach toward automation is to cluster the SOM weights. With vastly varying cluster statistics this can be a challenge. Partial results were achieved by [13] and several works referenced therein. [13] used an agglomerative multi-phase tree-based clustering that is perhaps the most successful in the recent literature. However, critical examination and repeating their main experiment made it clear to us that 1) different distance metrics are needed in order to accommodate non-spherical clusters and allow singleton weight clusters (for discovery of rare species!); 2) the size of the PE receptive fields should be considered in the partitioning of the tree; 3) the topology of the data should be considered in some decision phases. The SOM topology was considered in [14] in a limited manner. The receptive field size or the local data topology, to our knowledge, has not been used in SOM clustering. We are developing an algorithm based on these observations. It shows more success than [13], at least for moderate dimensional data [15]. A demonstration follows in Section III. We refer the reader to [13] and [15] for a discussion of previous works on this subject.

II. THE CHALLENGES OF EXTRACTING INFORMATION FROM SPECTRAL IMAGES

Spectral images are "stacked" images of the same spatial area, each taken at a different wavelength. The individual images are called image bands. *Hyperspectral* sensors 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 repeatable, unique pattern identifying the surface material(s) within pixel (x, y). The feature space spanned by VIS-NIR spectra is $[0, U]^{NB} \subset \Re^{NB}$ where $U > 0$ is an upper limit of the measured reflectivity. Sections of this space can be very dense while other parts may be extremely sparse, depending on the materials in the scene. Detailed clustering and classification of hyperspectral imagery can provide a great wealth of information. However, the intricate spectra pose unique challenges due to high dimensionality, subtle pattern differences, and other factors, discussed in more detail in [16]. Many favorite classifiers and clustering algorithms (such as Maximum Likelihood, Parallel Piped, Mahalonobis Distance, K-Means or Isodata) have difficulty handling the high-dimensional data vectors. Prior dimensionality reduction is frequently performed by PCA or wavelets, or by selection of important image bands by domain experts. We found undesirable loss of class distinction with all of these methods [17], [18], [16]. Non-linear feature extraction such as by [19] may retain more of the relevant information but systematic studies do not exist to show their general power

for many classes with subtle spectral shape differences, and that the discovery potential of small "interesting" groups of data is preserved. Another, fundamental aspect is that linear approaches such as PCA and/or methods based on second order (Gaussian) statistics only, may not detect some of the most interesting features because hyperspectral images are characterized by higher order statistics [20]. Among the few methods that have successfully been developed to suit hyperspectral data, the rule based supervised classifier Tetracorder (formerly Tricorder) [21], linear mixture modeling [22] (works well up to a handful of mixing endmembers) and the hierarchical segmentation HSEG [23], [24] are mature and richly textured procedures with many years of experience.

III. DATA ANALYSES

First we illustrate the dramatic effect of negative magnification for the discovery of rare species in a Mars Pathfinder spectral image taken at the landing site. A full clustering of the entire ($\approx$ 900 x 1000 pixel) image with a Conscience ($alpha = 1$) SOM was published in [25] along with some discussion of the geologic units. Here, Figure 1 compares $alpha = 1$ and $alpha < 0$ mapping showing the differences for a selected small image area.

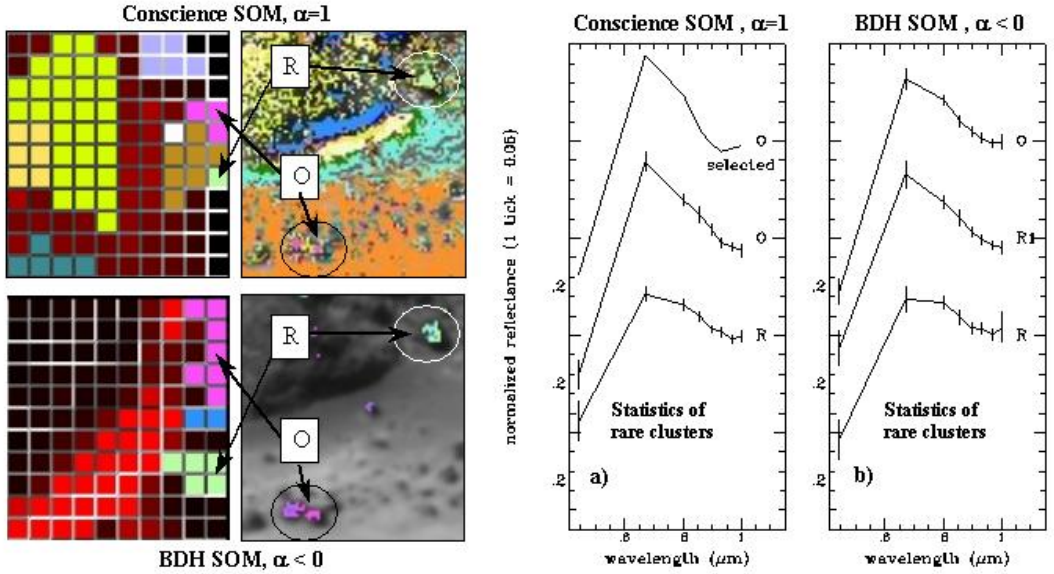

Fig. 1. Finding rare species in octant S0184 of the Imager for Mars Pathfinder SuperPan panorama, taken by the left eye of IMP. **Left panel:** Enlarged small details of the 40 x 40 SOMs (left) and the clustered images (right). *Upper pair:* Left: Conscience SOM detail, showing the representation areas of cluster O (hot pink) and R (light green). Right: The prototypical image area where "black" rock units were found by previous research. Labels and arrows point out the tiny, 15 - 25 pixel areas. O and R were both categorized as relatively pristine (olivine or pyroxene rich) material before our clustering separated the two subtypes, consistent with predominantly ortho- and clinopyroxene compositions. *Lower pair:* SOM and image detail, produced by BDH magnification with $\alpha < 0$. Left: O is now represented by 7, R by 5 PEs, in contrast to the Conscience SOM representation of 3 and 1 PEs, respectively, and the separation by the white fences is more obvious. In the image detail only the rare species are shown here superimposed on a grey scale image band. **Right panel:** The spectral signatures of the rare clusters, offset for clarity. Left: Statistics from the Conscience SOM clustering. The upper graph is an average of the prototypical O spectra (selected few spectra of confirmed black rock pixels). The middle graph is the mean of the O cluster with standard deviations indicated for each spectral channel. It is apparent that although the Conscience SOM O cluster preserves some main characteristics of the prototype, the pronounced absorption band toward 0.9 μm is lost, which means that many non-O spectra were captured in this cluster. The cluster statistics of R shows marked differences from O (our discovery in [25].) Right: Statistics from the BDH SOM clustering. On the upper graph, the mean of the (unfiltered) O cluster shows very good resemblance to the mean of the selected O prototypes on the left. This means that negative magnification not only made the O cluster more easily detectable but also more consistent. Even though this O cluster generated by BDH contains 818 pixels compared to 283 pixels in O generated by the Conscience SOM, it is a better match to the prototype than its Conscience SOM counterpart in the left middle graph. The BDH R (194 pixels) versus Conscience R (79 pixels) exhibit very similar general characteristics. A third, in-between cluster, labeled R1 (112 pixels) and colored blue in the BDH SOM is clearly separated from both O and R, and possesses mixed traits. Several blue pixels occur in both the O and R prototypical areas in the left panel, lower right image.

The Conscience SOM ($\alpha \approx 1$, verified for synthetic imagery in [8]) not only produces faithful *pdf* matching but has the additional advantage of only updating the immediate neighbors during learning, therefore it is fast. This makes it well suited for hyperspectral imagery. Figure 2 shows the result of a semi-manual cluster capture from a Conscience SOM that learned the structure of an AVIRIS image. The cluster identification was done with the "remap" tool of our HYPEREYE environment. HYPEREYE is an algorithm research, software development and data analysis environment, some details of which are posted at http://ece-old.rice.edu/HYPEREYE . In "remap", various knowledge representation layers (receptive field density, weight distances between PEs visualized as proportionally bright fences, etc.) can be overlain, and visually detected weight clusters outlined with a cursor, then the outlined PE area and the corresponding image pixels highlighted in a selected color. Visualization parameters can be controled on the fly for proper scrutinization of particular areas.

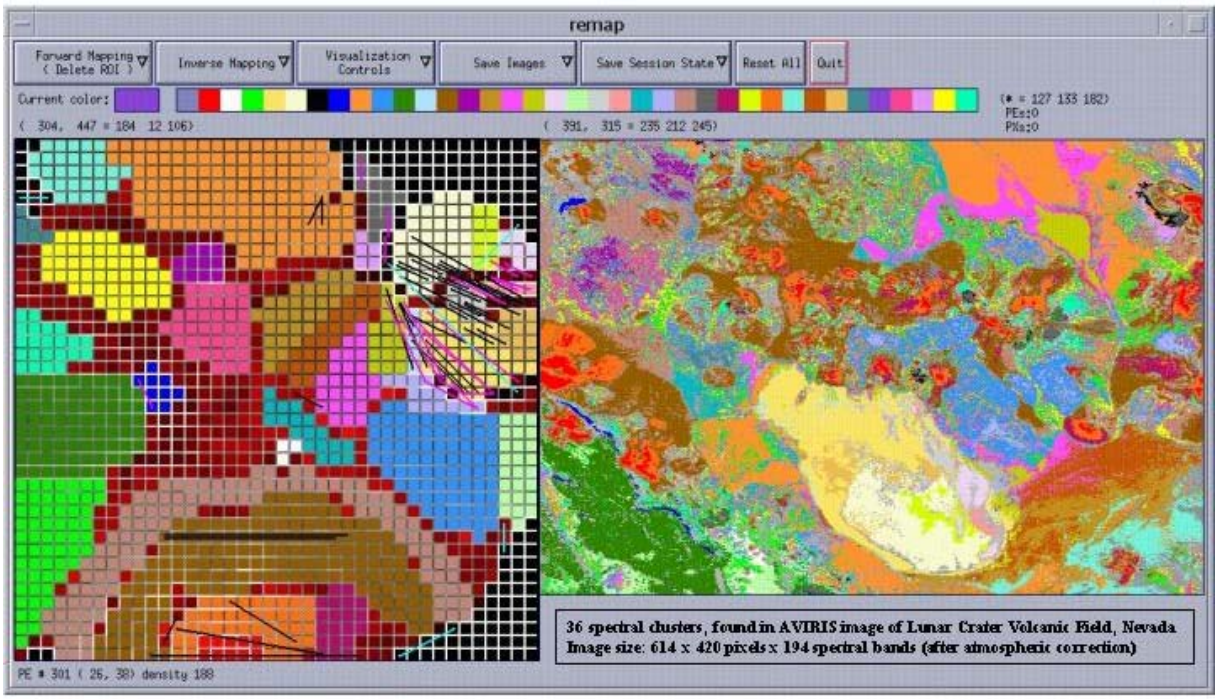

Fig. 2. Clustering a hyperspectral image with a Conscience SOM. The source data is a 194-band, 512 x 614 pixel AVIRIS image of a 10 x 12 square km area over the Lunar Crater Volcanic Field, Nevada. The image comprises 140 Mb of data. It contains, among other materials, volcanic cinder cones (class A, red) and weathered derivatives thereof such as ferric oxide rich soils (L, brown; W, mauve), basalt flows of various ages (black and medium blue classes, F, I), a dry lake divided into two halves of sandy (D, dark yellow) and clayey composition (E, light yellow); a small rhyolitic outcrop at the bottom (B, white); and some vegetation at the lower left corner (J, dark green). A long linear feature (G, dark blue) bordering the vegetated area is a scarp exposing a particular type of basalt. The labels are tied to spectral signatures in Figure 3, right. **Left:** The 40 x 40 SOM with the captured clusters. The underlying representation of the SOM knowledge includes the weight distances between PEs (black-to-white fences, for low-to-high weight differences); and the density (the size of the receptive field) for each PE, represented by proportionally bright red color. The corners contain a number of inactive PEs with 0 density. In addition, we indicate the extent of topology violations in this SOM by drawing lines from each PE to all those PEs which should be its neighbors based on data space topology, but they are not neighbors. This is determined by the Topographic Function. In short, there are few violations, and they are mostly connecting PEs within the same cluster (no violation on the cluster level). Several lines connect to unevaluated areas or to PEs in other clusters. These are helpful in scrutinizing and improving the semi-manual cluster determination. **Right:** The clusters shown remapped into image space. 36 clusters were found, each corresponding to a meaningful geologic cover type. 24 of these cover types were previously studied through supervised classifications. The clusters corresponding to the previously classified areas match those remarkably accurately. A more detailed description and comparison is given in [16].

Figure 3, left, shows a cluster-level verification of the $\alpha \approx 1$ property of the SOM in Figure 2. If the SOM is a maximum entropy map, all active PEs should have about the same number of data points mapped to them. On the cluster level this translates to the number of PEs representing a cluster being proportional to the size of the cluster. This is nearly true according to Figure 3. The deviations can be due to the number of unaccounted PEs that were simply not evaluated through the semi-manual cluster capturing, and it can also be due to less than ideal learning.

IV. DISCUSSION AND FUTURE TASKS

The above considerations were made in the context of fixed size SOMs. Size and other configuration aspects (such as shape and periodicity) of the SOM grid are important issues wrt the success of data modeling, which we omitted for space limitations. We mention, however, that the Growing Self-Organizing Map (GSOM) [26] can automatically determine the ideal configuration, at extra cost.

Our systematic verification of the BDH magnification algorithm for higher-dimensional data is work in progress. We have not yet fully evaluated its behaviour, especially the finer tuning of the controls, for higher than the 8-dimensional cases in Figure 1 and in [8]. The preliminary results shown here are extremely encouraging, but a lot more work is needed to make this a production tool. Meanwhile, using the Conscience SOM for hyperspectral data appears to reliably generate maximum entropy quantization. We point out that at the same time it also approximates minimum distortion quantization for high-dimensional data because $\alpha = 1 \approx d/(d+2)$ as d increases.

The number of unevaluated PEs between clusters shows, however, that this semi-manual approach (however good the results) is too laborious to get to all details.

The third example, explained in the captions of Figures 4 and 5, demonstrates the current capabilities of our automated SOM weight clustering on an 8-band urban remote sensing spectral image of Ocean City. Former supervised classification of this image and semi-manual SOM cluster extraction [8] serve as benchmark.

Algorithmic problems in our SOM weight clustering (Figure 5) stem partly from widely varied cluster shapes and sizes, and from definition of thresholds that we aim to derive from data and SOM characteristics. However, the difficulty of this increases with spectral dimension. We

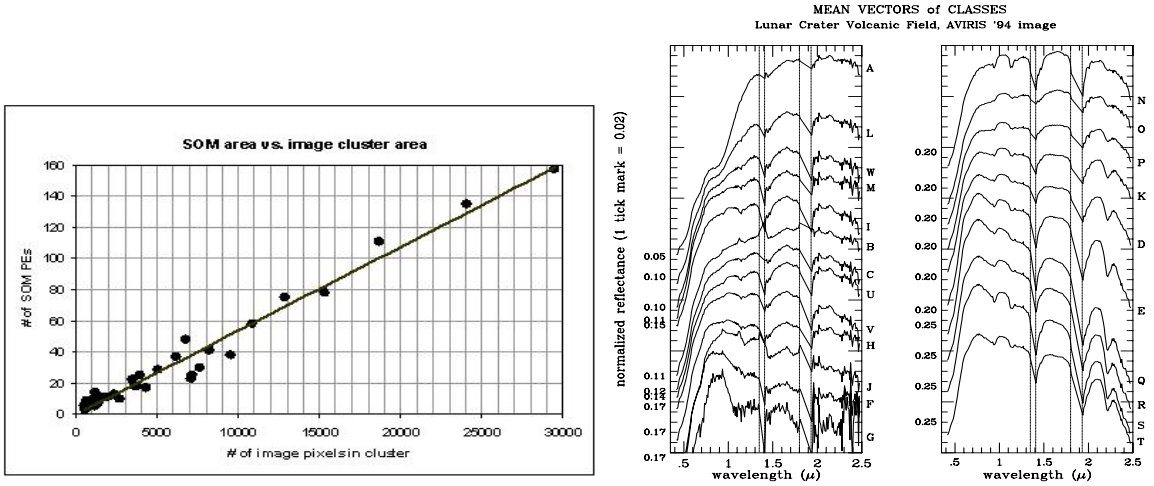

Fig. 3. **Left:** Maximum entropy property of the mapping with Conscience SOM ($\alpha = 1$) in Figure 2. The PE areas representing the various clusters are very nearly linearly proportional to the size of the image clusters. The source of imperfection can be the many unevaluated PEs in the SOM. The entropy measured directly from the SOM, normalized to the maximum possible entropy value for a 40 x 40 SOM, is 0.97. **Right:** Mean spectral signatures of 23 of the classes in Figure 2 illustrate some of the subtle differences among verified geological cover types. Classes E, K, N - P, Q - T (playa and outflow materials), for example, form a series with increasing amounts of water and clay (more details in [16]). Spectra are offset for clarity. The dotted vertical lines indicate the atmospheric water bands.

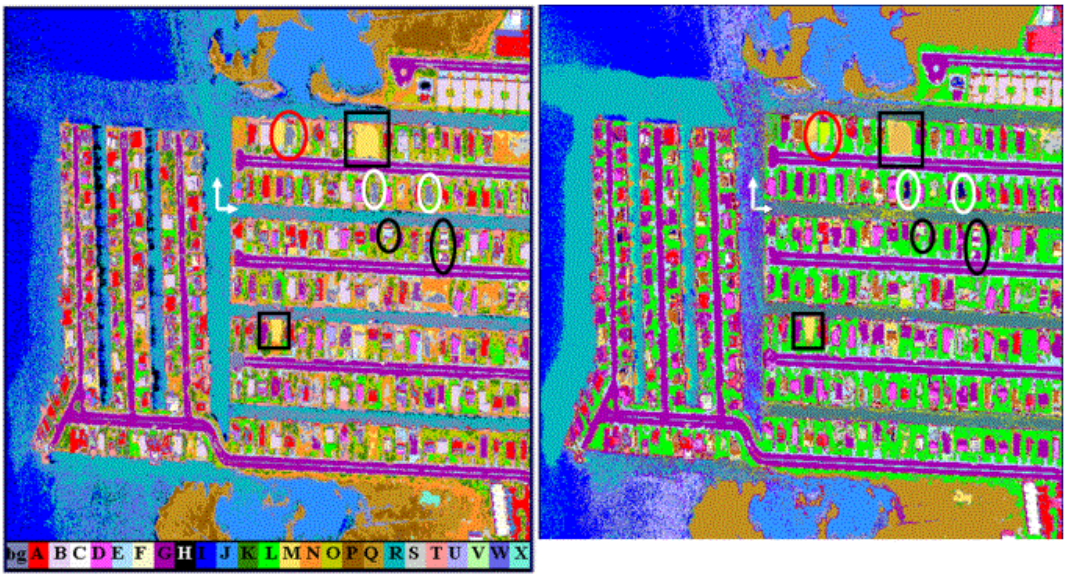

Fig. 4. **Left:** An earlier supervised classification of the Ocean City image, mapping 24 known cover types. Red and white ovals show unclassified shapes of buildings (the colour of the background, 'bg'). The black ovals show one rare class that was included in the supervised classification. Semi-manual clustering produced a cluster map strikingly similar to this supervised class map [8]. **Right:** Clusters detected by our automated SOM cluster extraction algorithm [15]. The agreement between the cluster map and supervised class map is very good. While the supervised class map has many unclassified pixels, here the vast majority of the pixels are assigned to clusters, which produces more appearances of some colours such as green and turquoise. The unclassified gray spots (in red and white ovals on the left) are now filled exactly, and with colours (greenish-yellow, label a, in the red oval, and dark blue, Z in the white ovals) different from the existing 24 colours. See Figure 5 for their labels and spectral signatures, distinct from the rest. The new-found rare clusters only occur at these locations. We also found subclusters. One example is in the black rectangles (the supervised class M is split into M and e). Arrows indicate two subclusters of the former turquoise water class: R (turquoise), water in the canals open to the ocean and f (dark turquoise), water in canals farther into the city.

consider our algorithm very preliminary at this stage and continue to research and improve it.

Acknowledgments

This research was partially supported by grants NAG9-10432 and NNG05GA94G from the Applied Information Systems Research Program, and by grant NAG5-13294 from the Mars Data Analysis Program, both under NASA's Science Mission Directorate. Collaborations with Dr. Thomas Villmann, University of Leipzig, Germany, Dr. William H. Farrand, Space Science Institute, Boulder, CO, are gratefully acknowledged. Mr. Kadim Tasdemir, Ms. Lily Zhang, graduate students at Rice University performed parts of the simulations. Thanks to Mr. Philip Tracadas, Rice University, for software development, to Dr. Bea Csathó, Byrd Polar Research Institute, Ohio State University, for Ocean City data, and to Dr. Scott Murchie, Johns Hopkins APL, for the Mars imagery.

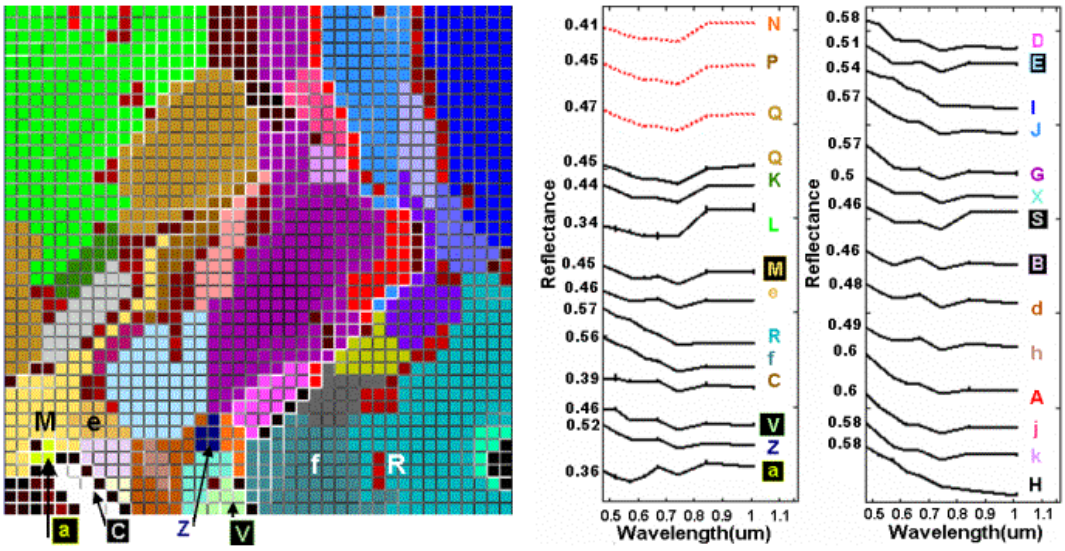

Fig. 5. **Left:** The 32 clusters identified by our automated SOM weight clustering in the SOM of the Ocean City image discussed in Figure 4. Arrows show rare clusters. Parent cluster M in Figure 4, Left, and in earlier semi-manual clustering [8] splits here into M and e. Similarly, former cluster R splits into R and f. **Right:** The mean spectral signatures of 24 of the 32 clusters, offset for clarity. The standard deviations are small, an indication of the clustering quality. The signatures of the rare clusters (C, V, Z, a) are distinct, the subclusters M, e, R, f show slight but consistent spectral differences, justifying the algorithm's identification of these clusters. Dotted lines at the top of the left panel are signatures of earlier supervised classes P, Q, and N. These are grouped into one supercluster, Q (two sparate brown areas), by our clustering, indicating an algorithmic imperfection.

References

[1] T. Kohonen. *Self-Organization and Associative Memory*. Springer-Verlag, New York, 1988.

[2] T. Kohonen. *Self-Organizing Maps*. Springer-Verlag, Berlin Heidelberg New York, 1997.

[3] H.-U. Bauer, R. Der, and M. Herrmann. Controlling the magnification factor of self–organizing feature maps. *Neural Computation*, 8(4):757–771, 1996.

[4] P.L. Zador. Asymptotic quantization error of continuous signals and the quantization dimension. *IEEE Trans. on Information Theory*, 28(2):139–149, March 1982.

[5] H. Ritter and K. Schulten. On the stationary state of Kohonen's self-organizing sensory mapping. *Biol. Cyb.*, 54:99–106, 1986.

[6] D.R. Dersch and P. Tavan. Asymptotic level density in topological feature maps. *IEEE Trans. on Neural Networks*, 6(1):230–236, January 1995.

[7] D. DeSieno. Adding a conscience to competitive learning. In *Proc. ICNN, July 1988* **I**, pages 117–124, New York, 1988.

[8] E. Merényi and A. Jain. Forbidden magnification? II. In *Proc. of European Symposium on Artificial Neural Networks (ESANN'04)*, pages 57–62, Bruges, Belgium, April 28-30, 2004. D facto publications.

[9] A. Jain and E. Merényi. Forbidden magnification? I. In *Proc. of European Symposium on Artificial Neural Networks (ESANN'04)*, pages 51–56, Bruges, Belgium, April 28-30, 2004. D facto publications.

[10] Th. Villmann, R. Der, M. Herrmann, and Th. Martinetz. Topology Preservation in Self–Organizing Feature Maps: Exact Definition and Measurement. *IEEE Transactions on Neural Networks*, 8(2):256–266, 1997.

[11] A. Ultsch and H. P. Simeon. Kohonen's self organizing feature map for exploratory data analysis. In *Proc. INNC-90-PARIS* **I**, pages 305–308, Paris, 1990.

[12] A. Ultsch. Self-organizing neural networks for visualization and classification. In R. Klar O. Opitz, B. Lausen, editor, *Information an dClassification — Concepts, Methods and Applications*, pages 307–313. Springer Verlag, Berlin, 1993.

[13] J. Vesanto and E. Alhoniemi. Clustering of the self-organizing map. *IEEE Transactions on Neural Networks*, 11(3):586–600, May 2000.

[14] F. Murthag. Interpreting the Kohonen self-organizing map using contiguity-constrained clustering. *Patt. Recognit. Lett.*, 16:399–408, 1995.

[15] K. Tasdemir and E. Merényi. Considering topology in the clustering of Self-Organizing Maps. In *Proc. Workshop on Self-Organizing Maps (WSOM 2005)*, page Submitted, Paris, 2005.

[16] E. Merényi. Precision mining of high-dimensional patterns with self-organizing maps: Interpretation of hyperspectral images. In *Quo Vadis Computational Intelligence: New Trends and Approaches in Computational Intelligence (Studies in Fuzziness and Soft Computing, Vol 54, P. Sincak and J. Vascak Eds.)*. Physica Verlag, 2000.

[17] E. S. Howell, E. Merényi, and L. A. Lebofsky. Classification of asteroid spectra using a neural network. *Jour. Geophys. Res.*, 99(E5):10,847–10,865, 1994.

[18] T. Moon and E. Merényi. Classification of hyperspectral images using wavelet transforms and neural networks. In *Proc. Annual SPIE Conf.*, page 2569, San diego, CA, 1995.

[19] J. A. Benediktsson J. R. Sveinsson and et al. Classification of very-high-dimensional data with geological applications. In *Proc. MAC Europe 91*, pages 13–18, Lenggries, Germany, 1994.

[20] E. Merényi and T. Villmann. Self-Organizing Neural Net Approaches For Hyperspectral Images. In *Proc. Int'l Conf. on Intelligent Computing and Infromation Systems*, Ain Shams University, Cairo, Egypt, June 24–26 2002.

[21] R. N. Clark and G. A. Swazey. Mapping minerals, amorphous materials, environmental materials, vegetation, water, ice and snow, and other materials: The usgs tricorder algorithm. In *Proc. 5th AVIRIS Earth Science and Applications Workshop, JPL Publication 95–1*, volume 1, pages 39–40, Pasadena, CA, 1995.

[22] J. B. Adams, M. O. Smith, and A. R. Gillespie. Imaging spectroscopy: Interpretation based on spectral mixture analysis. In C.M. Peters and P.A.J. Englert, editors, *Remote Geochemical Analysis: Elemental and Mineralogical Composition*, pages 145–166. Cambridge University Press, New York, 1993.

[23] J. Tilton. Analysis of hierarchically related image segmentations. In *Proceedings of the IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, pages 60–69, Greenbelt, MD, USA, Oct. 27-28, 2003.

[24] J. A. Gualtieri and J. C. Tilton. Hierarchical segmentation of hyperspectral data. In *Proceedings of AVIRIS Earth Sciencs and applications Workshop*, Pasadena, CA, USA, 2002.

[25] E. Merényi, W.H. Farrand, and P. Tracadas. Mapping surface materials on Mars from Mars Pathfinder spectral images with HYPEREYE. In *Proc. International Conference on Information Technology (ITCC 2004)*, pages 607–614, Las Vegas, Nevada, 2004. IEEE.

[26] H.-U. Bauer and Th. Villmann. Growing a Hypercubical Output Space in a Self–Organizing Feature Map. *IEEE Transactions on Neural Networks*, 8(2):218–226, 1997.