SELF-ORGANIZING NEURAL NETWORK APPROACHES FOR HYPERSPECTRAL IMAGES[1]

Erzsébet Merényi
Rice University, Electrical and Computer Engineering Department, Houston, TX 77005, U.S.A.
Ph: 1-713-348-3595 Fax: 1-713-348-6196 e-mail: erzsebet@rice.edu

and

Thomas Villmann
University of Leipzig, Clinic of Psychotherapy, Karl-Tauchnitz Str. 25, 04107 Leipzig, Germany
Ph: +49 (0)341 9718868 Fax: +49 (0)341 2131257 e-mail: villmann@informatik.uni-leipzig.de

Abstract

Utilization of remote sensing multi- and hyperspectral imagery has shown a rapid increase in many areas of economic and scientific significance over the past ten years. Hyperspectral sensors, in particular, are capable of capturing the detailed spectral signatures that uniquely characterize a great number of diverse surface materials. Interpretation of these very high-dimensional signatures, however, has proved an insurmountable challenge for many traditional classification, clustering and visualization methods. We present Artificial Neural Network approaches, Self-Organizing Maps in particular, for spectral image interpretations, utilizing the full spectral resolution. Recent extensions to the original Kohonen SOM that emerged in the last few years and lend theoretical support for principled analyses of high-dimensional data spaces, are reviewed. The necessity of faithful topological mapping for correct interpretation is emphasized. The effectiveness of the presented approaches is demonstrated through case studies on real hyperspectral imagery.

Keywords: Neural Networks, Self-Organizing Maps, hyperspectral images, clustering, classification, discovery in high-dimensional data

1 Introduction

Hyperspectral imagery is one important example of recent, modern high-dimensional data. Acquired remotely by airborne or spaceborne sensors it is comprised of hundreds of spatially co-registered image bands. Each image band is obtained at a different wavelength through a narrow bandpass. Fig. 1.a illustrates the concept. Associated with every image pixel there is an n-dimensional feature vector across the wavelength dimension, which is called the spectrum of the material in that pixel. These n-dimensional vectors are the patterns to cluster or to classify for the identification of surface materials. In contrast to traditional multi-spectral imagery, such as, *e.g.,* Landsat TM, which uses a small number of broad bandpasses with large gaps over a given wavelength window, hyperpsectral sensors sample the same window contiguously and at regularly spaced wavelengths. The typical wavelength range for passive optical imaging is $0.4 - 2.5$ μm (Visible to Near-Infrared), measuring reflected sunlight, or $3 - 12$ μm (Thermal Infrared), measuring emitted heat. At hyperspectral resolution ($10 - 20$ nanometers) sufficient detail is recorded to capture the unique compositional fingerprints of surface materials. Fig. 1.b shows an example of the spectral detail that multi-spectral and hyperspectral resolutions provide. See, *e.g.,* [1] for a review. The history of hyperspectral remote imaging is about 15 years old. In the past 8–10 years this technology has matured, and presently a handful of sensors is offered for data collection, operated by commercial companies or government agencies. The best known are AVIRIS, developed by NASA/JPL [2], followed by HYDICE (Naval Research Lab [3]), Probe-1 and HYMAP (Earth Search Science, USA and HyVista, Australia), Hyperion (by NASA, on the Terra satellite,) DAIS (DLR) in Germany, and CASI in Canada. Their numbers have been growing in the past 3–5 years. The wealth of information provided by these advanced sensors proved spectacularly promising for a wide variety of large scale real applications. Environmental monitoring, exploration geology, precision farming, ecosystem monitoring, assessment of natural hazards, of water and air quality, battlefield characterization, planetary research, and many other areas eagerly seek to utilize hyperspectral imagery.

However, the dramatically increased spectral resolution makes the exploitation of hyperspetcral images a substantial challenge due to the high dimensionality of the data. The difficulties caused at the same time by huge data volumes, will not be discussed here. We assume that computing power and/or parallel implementations will respond to that question. The focus of this paper is the clustering and classification of hyperspectral images at full spectral resolution, *i.e.,* the analysis of several hundred dimensional data spaces.

[1] Download color copy from http://www.ece.rice.edu/~erzsebet/emann.html, ICICIS'02 paper.

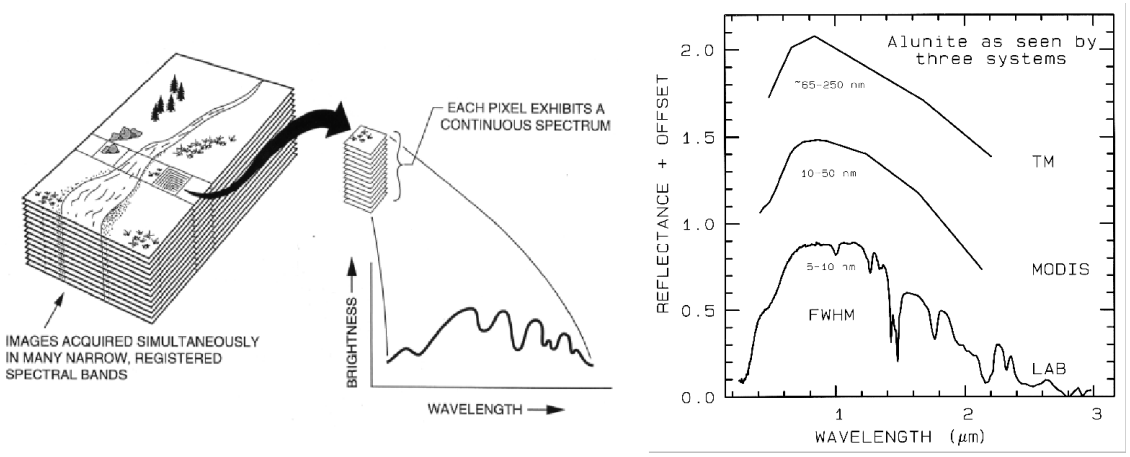

Figure 1: a) Left: The concept of hyperspectral imaging. Figure from Campbell, 1996. b) Right: The spectral signature of the mineral alunite as seen through the 6 broad bands of Landsat TM, as seen by the moderate spectral resolution sensor MODIS (20 bands in this region), and as measured in laboratory. Hyperspectral sensors such as AVIRIS of NASA/JPL produce spectral details comparable to laboratory measurements.

Why do we have to deal with such high dimensionality? Why not to reduce it to a few dozens or less dimensions, prior to classification? The answer may be surprising: hyperspectral signatures do not seem to lend themselves to significant reduction. While for classifications involving a relatively small number of specifically targeted surface units substantial dimensionality reduction may be possible, applications that seek to exploit the sophisticated sensor information for the discrimination of many surface covers, or which seek to make discoveries in hyperspectral imagery, dimensionality reduction has not been successful. The best example we have seen is perhaps by [4], who reduced a 224-channel AVIRIS image to 35 channels using the same neural network that was subsequently used for the classification of the eight surface units that they were interested in. The point we make here is that the feature extraction method was a powerful non-linear (neural network) method, claimed to produce the best possible image band selection for the particular case, yet 35 bands were necessary for the 8-class categorization. By costumary measures (such as, for example, PCA) the intrinsic spectral dimensionality of hyperspectral images is low, $5-10$ at most. Yet, efforts to reduce dimension to the predicted low numbers without significant loss of class distinctions have remained unsuccessful so far. A more detailed list of dimension reduction studies of hyperspectral data (including PCA, wavelets and Self-Organizing Maps (SOMs)) and results is given in [5, 6]. It appears that the intrinsic dimensionality (ID) of hyperspectral images is not well understood. The spectral bands, many of which are highly correlated, may lie on a low-dimensional but non-linear manifold, which is a scenario that eludes many conventional classification and clustering techniques.

Dimension reduction, in any case, carries the risk of missing a discovery. Later in this paper we will see examples of discovering small, statistically insignificant spectral units, which are very interesting, and which may not be seen from less than full spectral resolution data. It is imperative therefore to have methods that are capable of handling the full dimensionality, to enable discoveries, and to establish benchmarks for dimension reduction and classification performance studies.

What are then, the challenges associated with classification and clustering data of hyperspectral resolution? First, time tested favorites, covariance based classifiers such as Maximum Likelihood, Mahalonobis Distance, require at least as many training samples per class as the number of bands plus one, in order to have a non-singular covariance matrix. This creates a severe problem in remote sensing applications where often only a few precious training samples are available for some rare class. It may also happen that the entire class of interest contains less than the required number of training pixels! In such situations the above classifiers break down for mathematical reasons. Dimensionality reduction (or feature extraction) is frequently accepted in order to accomodate the data for traditional methods but this may result in an undesirable loss of information. In the presence of sufficient number of training pixels, classifiers based on second order statistics (covariance based classifiers) will still miss interesting variations in the data. Table 1 illustrates the highly non-Gaussian nature of a hyperspectral image cube: the third and fourth order statistics, skew and kurtosis, which would both be 0 for Gaussian data distribution, show strong variations along different slices across the image cube. Some of the higher values are typed in bold face for emphasis. Indeed, as we will see below, on real hyperspectral imagery the Maximum Likelihood and Mahalonobis Distance classifiers perform poorer than others that we tested [7].

band #	Skewness	Kurtosis	band #	Skewness	Kurtosis
1	0.0387754	-0.325047	101	0.30173	0.951647
11	-0.0071132	-0.17189	111	0.183944	0.147923
21	-0.204628	0.325475	121	0.133933	-0.141926
31	-0.554751	0.588788	131	0.240551	0.176545
41	-0.970494	1.11788	141	-0.427421	1.01846
51	**-1.2609**	1.69146	151	0.0547962	0.73411
61	**-1.26399**	**2.02341**	161	-0.190321	1.06755
71	-0.440014	**1.84514**	171	-0.500422	**2.09482**
81	0.0691932	**8.51355**	181	-0.843997	**2.45075**
91	0.292169	0.833326	191	-0.768847	1.7723
All	**-1.81785**	**3.09868**			

Table 1:
The non-Gaussianity of a hyperspectral image cube. Statistics of 20 selected bands. Data: AVIRIS 224-band image of the Lunar Crater Volcanic Field, Nevada, 1994.

Another important analysis tool is clustering (unsupervised classification), for discovery or in support of subsequent supervised classification. With high spectral dimensionality, many of the methods used for low-dimensional data spaces become impractical or fail for other reasons. For example, pairwise scattergrams or ratios, costumarily used for images with a few bands, are untractable for a human beyond a dozen or so bands. The performance of familiar statistical clustering techniques tend to decline at high dimensions, and beyond 3-dimensional spaces the difficulty of visualizing high-dimensional cluster structure arises. Yet, sensitive automated cluster detection is critical for high-dimensional data spaces because of the possibility of discovering knew knowledge through the detection of surprising relationships.

In the following, we present successful Artificial Neural Network (ANN) clustering and classification of real hyperspectral imagery with as many as 200 data dimensions, and discrimination of many (20–30) classes. The strength of our ANN approaches lies in the application of Self-Organizing Maps (SOMs). After introducing the test data, techniques and theoretical considerations are described along with presentation of sample results.

2 Sample Hyperspctral Test Image: The Lunar Crater Volcanic Field

The Lunar Crater Volcanic Field (LCVF) has been one of several standard sites imaged yearly by AVIRIS. It was the focus of a number of independent, detailed studies such as the large-scale NASA-sponsored Geologic Remote Sensing Field Experiment (GRSFE) [8], individual PhD works and other projects through which substantial amount of field knowledge, ground truth and other supporting data were accumulated [9, 10, 11, 12]. Fig. 2 shows a natural color composite of the LCVF, with locations representative of various cover types marked by their respective class labels. The LCVF contains, among other materials, volcanic cinder cones (class A, reddest peaks) and weathered derivatives thereof such as ferric oxide rich soils (L, M, W), basalt flows of various ages (F, G, I), a dry lake divided into two halves of sandy (D) and clayey composition (E); a small rhyolitic outcrop (B); and some vegetation at the lower left corner (J), and along washes (C). Alluvial material (H), dry (N,O,P,U) and wet (Q,R,S,T) playa outwash with sediments of various clay contents as well as other sediments (V) in depressions of the mountain slopes, and basalt cobble stones strewn around the playa (K) form a challenging series of spectral signatures for pattern recognition as seen from Fig. 2, right panel. A long, NW-SE trending scarp, straddled by the label G, borders the vegetated area. On this color composite, containing information from only three image bands, many of the cover type variations do not manifest. They become visible in the cluster and class maps. After atmospheric correction and removal of excessively noisy bands (saturated water bands and overlapping detector channels), 194 image bands remained. These 194-dimensional spectra are the input patterns in the following analyses. Details of this data set are in [7].

3 Self-Organizing Maps for Clustering and Classification of Hyperspectral Images

The use of ANNs for large-scale and complex classification tasks such as for hyperspectral image interpretation is motivated a) by their massively parallel architecture, which provides speed and convenient hardware implementation for integrated, on-board, (near) real-time processing; b) by their power in pattern recognition. The

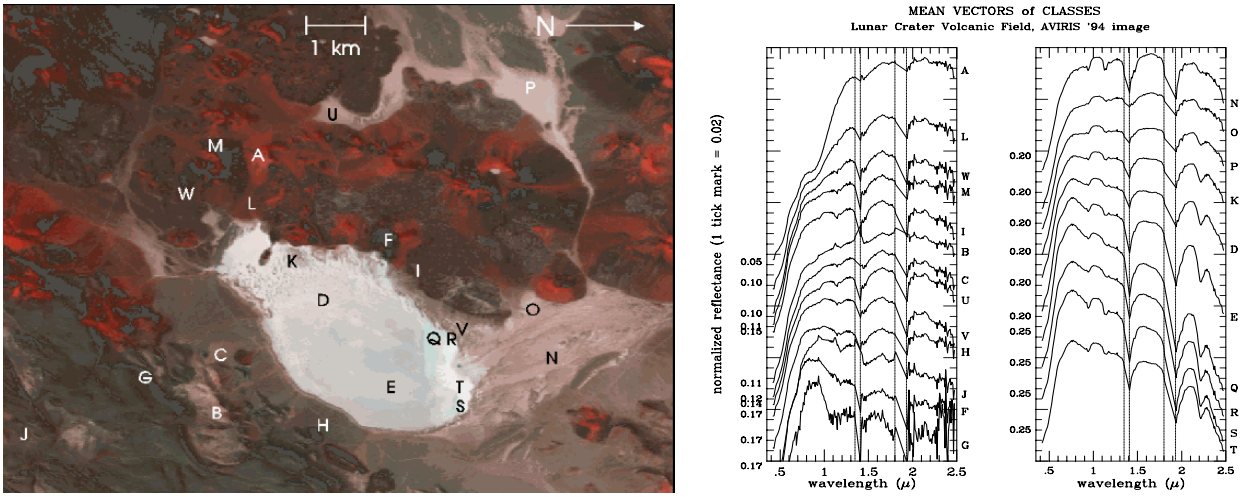

Figure 2: a) Left: Color composite of the Lunar Crater Volcanic Field, Nevada, USA. The original image, taken by AVIRIS in 1994, comprises 140 Mbytes of data, 512 x 614 pixels, covering a 10 x 12 square km area. The 224 spectral bands span the 0.38 to 2.5 micron wavelength range. One red, one green, and one blue band was selected from the 224 bands, and combined into an RGB representation for this display. Labels mark 23 main geologic features as well as locations of training pixels for the respective class categories. The letters correspond to the class labels used throughout the paper. b) Right: Typical spectral signatures of the 23 geologic units indicated by the labels on the color composite. The spectra are offset vertically for viewing convenience, with reference reflectance values for each curve at left, and the labels of the respective cover types at right. Original image is in color. Download from http://www.ece.rice.edu/∼erzsebet/emann.html, ICICIS'02 paper.

focus of this paper is to provide a contribution to this second point. For a comprehensive review on ANNs, see *e.g.,* [13]. Earlier works documented capabilities of ANNs (mostly Back Propagation ANNs) for remote sensing spectra for a relatively small number of classes (5–12) and for low-to-moderate number of channels (6 – 60). A review is provided in [6]. The case of highest spectral resolution, 60 channels, in a previous study [14] uses synthetic data which, by virtue of Gaussian data construction, favors the Maximum Likelihood classifier. While all earlier works provide valuable insights to various aspects of classification, full exploitation of hyperspectral data requires further efforts.

A small number of works attempted to apply ANNs to hyperspectral imagery [4, 15, 6]. [4] reduced the original 224 AVIRIS bands to 35 prior to classification and found that the 35 bands they obtained with feature extraction by a Back Propagation ANN, contained enough information for the discrimination of the 8 classes that they were interested in. [7] present a rigorous and systematic performance evaluation of four conventional and an SOM-based ANN classifier on the LCVF image introduced above, using the *full* spectral resolution, discriminating among 23 classes. The ANN classifier, with 90% accuracy, outperformed the others. In this case, reduction to 30 bands led to the loss of a number of the 23 classes, demonstrating that by using all (or nearly all of the 224) AVIRIS bands, detection of more, geologically meaningful spectral variations, and increased classification accuracy is to be gained through a sophisticated enough method. The key to success is the Self-Organizing Map.

We use SOMs in two ways in the interpretation of hyperspectral imagery, or high-dimensional data spaces in general. One is clustering, with the objective of discovering similarity groups, unknown structure in the high-dimensional space. The other is to aid in supervised classification, by first performing the clustering, and then connect the converged SOM to a categorization layer. This latter approach helps prevent teaching inconsistent information, and produces better generalization from a small number of training samples, and thus it results in higher classification accuracy, as described below. In both cases, faithful SOM representation of the high-dimensional clusters is crucial, therefore we discuss some aspects of that next.

3.1 Clustering and discovery in hyperspectral images with SOMs

Self-Organizing Maps (SOMs) [16] implement a mapping procedure which, if correctly applied, should end in a topology preserving projection of the high-dimensional data onto a low-dimensional lattice. In most cases a two-dimensional SOM grid is the common choice of lattice structure because of its easy visualization [17]. In general, this choice is not guaranteed to produce a topology preserving mapping which, informally speaking, means that similar data vectors are mapped onto the same or neighbored locations in the lattice and vica versa

[18]. If neighborhood preservation fails the map becomes "twisted" leading to misinterpretation of the cluster structure of the high-dimensional input space [19]. A further aspect that should be addressed for faithful SOM mapping is the problem of the detection of rarely occurring surface material classes in the images.

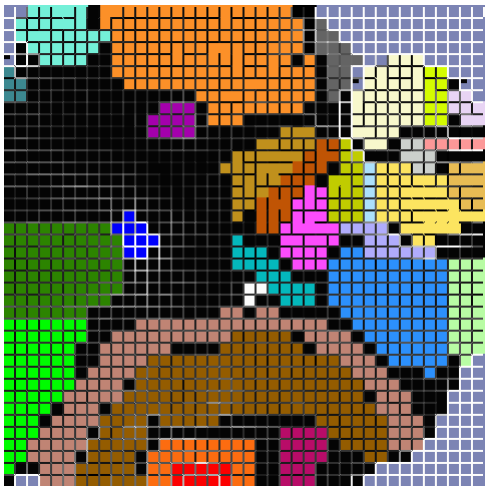

Figure 3: Clusters identified in a 40 x 40 SOM. Each grid location represents one neuron. The SOM was trained on the entire 194-band LCVF image, using the DeSieno (1988) algorithm. Original image is in color. Download from http://www.ece.rice.edu/∼erzsebet/emann.html, ICICIS'02 paper.

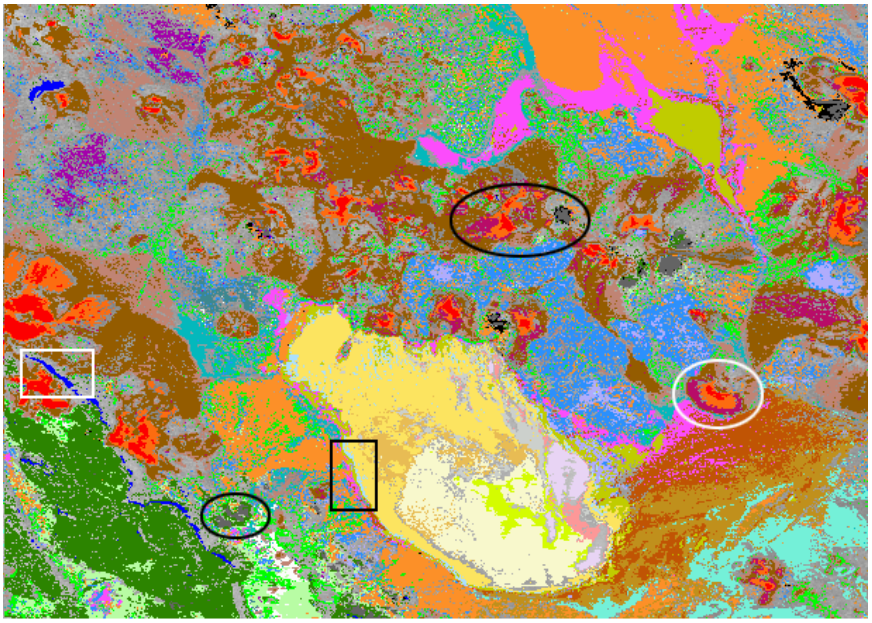

Figure 4: The clusters from Fig. 3 remapped to the original spatial image, to show where the different spectral types originated from. Ovals and rectangles highlight examples discussed in the text. Original image is in color. Download from http://www.ece.rice.edu/∼erzsebet/emann.html, ICICIS'02 paper.

Figures 3 and 4 show a successful mapping of the 194-dimensional data space of the LCVF hyperspectral image onto a 2-dimensional SOM lattice of 40 by 40 neurons. It was generated by a modification of the original SOM [16], the conscience algorithm of DeSieno [20]. 32 groups of neurons that were found to be sensitized to groups of similar spectra in the 194-dimensional LCVF input data, are indicated by various colors. The boundaries of these clusters were determined by a somewhat modified version of the Ultsch and Simeon method [21]. Areas where no data points (spectra) were mapped are the grey corners with uniformly high fences, and are relatively small. The black background in the SOM lattice shows areas that have not been evaluated for cluster detection. The spatial locations of the image pixels mapped onto the groups of neurons in Fig. 3, are shown in the same colors in Fig. 4. Color coding for clusters here is the same as that of corresponding classes in the supervised classification in [6] to show similarities. (Those class maps are not shown here for lack of space

but [6] is accessible for download, see link in reference list.) Colors for additional groups were added.

This cluster map is remarkably accurate in outlining geological variations in the LCVF. The SOM detected all classes that were known to us as meaningful geological units and that were classified in [6]. The "discovery" of classes B (rhyolitic outcrop, white), F (young basalt flows, dark grey and black, some shown in the black ovals), G (a different basalt, exposed along the scarp, dark blue, one segment outlined in the white rectangle), K (basalt cobbles, light blue, one segment shown in the black rectangle), and other spatially small classes such as the series of playa deposits (N, O, P, Q, R, S, T) is significant. This is the capability we need for sifting through high-volume, high-information-content data to be alerted to interesting, novel, or hard-to-find units. In addition, the SOM detected more, spatially coherent, clusters than the 23 classes that were defined originally for supervised classification. The SOM's view of the data is more refined and more precise than that of the geologist's! [6] contains a direct comparison of this cluster map with known surface features. For example, class A (red in Fig. 3 in [6]) is split here into a red (peak of cinder cones) and a dark orange (flanks of cinder cones) cluster, that make geologic sense. The maroon cluster to the right of the red and dark orange clusters at the bottom of the SOM fills in some areas that remained unclassified (bg) in the ANN class map, in Fig. 3 in [6]. An example is the arcuate feature at the base of the cinder cone in the white oval, that apparently contains a material different enough to merit a separate spectral cluster. This material fills other areas too, consistently at the foot of cinder cones (another example is seen in the large black oval). Observation of further refinements are left to the reader.

Correctness of this mapping was verified semimanually, by observing that all the known surface units were found, and by validating the geologic meaning of the additional discoveries (the 9 clusters that were identified by the SOM in addition to the 23 originally specified classes.) However, for large-scale production we need a more automated approach to ensure correct mapping. A growing SOM (GSOM) method was developed to generate a guaranteed topology preserving mapping in a simple hypercube structure of the lattice [22]. By defining an exact measure, and monitoring topology violation during the learning of the SOM, [23] can determine the best strategy for alleviating the topology violation: either increase the number of neurons in the present dimension of the SOM, or increase the dimension of the lattice. Preliminary experiments with the GSOM resulted in a 3-dimensional SOM for this LCVF image, producing a clustering that is comparable to that in Fig. 4 [24]. We are in the process of thorough evaluation of the GSOM for hyperspectral imagery.

A mapping that matches the probability density (*pdf*) of the input data is preferable, because it represents regions of the input space with resolutions appropriate for their relative densities. If the SOM is used as a classifier system the distribution of the weight vectors across the SOM lattice is determined by a power function of the probability density of the data vectors [25], with the so-called magnification factor as the power. The map magnification factor, $m(x)$, which is the number of neurons representing a subsection dx of the input space, is required to be proportional to the *pdf*. For magnification factors less than 1, which is the case for the standard Kohonen SOM, it may be difficult or impossible to separate spectral patterns of seldom occurring surface materials. The original SOM algorithm does not meet the above requirement and tends to overrepresent regions of low density, and underrepresent regions of high density [13]. This could prevent sensitive separation of spectral clusters with subtle signature differences, such as the sandy and clayey part of the playa (classes D and E, dark and light yellow), or the splitting of class A into the red and dark orange clusters. The *conscience algorithm* of DeSieno [20] adjusts the winning frequencies of the SOM neurons so as to ensure proper spreading (compaction) of information that is dense (sparse) in the input feature space. See details in [6]. Van Hulle points out that adding a conscience algorithm to the SOM does not equate to equiprobabilistic mapping, in general. However, for *very high dimensions*, a minimum distortion quantizer (such as the conscience algorithm) approaches an equiprobable quantizer [26] (page 93). The above clustering seems to support this statement, as seen from an empirical evaluation in Fig. 6 of [6].

As shown in [27] it is possible to control the magnification of the SOM by the introduction of a local learning rate that is dependent on the data density. The corresponding learning scheme can easily be implemented into the standard learning rule. Since this is a rigorous control for *pdf* matching that could be very valuable in a fully automated environment (such as on-board computational intelligence in space missions or Earth monitoring) we are interested to see if magnification control can replace the more heuristic DeSieno algorithm, with equal success for high-dimensional data.

3.2 Supervised classification of hyperspectral images with SOM hybrid ANN

Back Propagation neural networks, which are the most popular and best known ANN paradigms, can be very difficult to train as their complexity increases non-linearly with the number of spectral bands and the

possibility for the gradient descent learning rule to get stuck in local minima increases dramatically with the high dimensionality of the weight space. Dimensionality reduction is frequently applied to high spectral resolution data to achieve tolerable training time, or training convergence at all (*e.g.,* [4, 28]). We use an hybrid ANN architecture with a 2-dimensional Self-Organizing Map connected to an output layer with a Widrow-Hoff learning rule [29, 30]. This network scales up well, training convergence is relatively fast and easy as demonstrated by previous works with up to 200 channels and over a dozen classes. This ANN construction gave us more accurate results than Back Propagation for under- and unevenly represented classes in improving asteroid taxonomy [31, 32], and discovering Martian surface units [15]. By first forming a cluster map of the data in an unsupervised learning phase, it is less likely to learn inconsistent class labels in the subsequent supervised phase than Back Propagation (will not merely memorize the training information). This results in better generalization from even a small set of unevenly distributed samples, and consequently in higher classification accuracy. [33] recently offered proof of the advantage of combining unsupervised traning with supervised, in case of small training population. For details on the SOM-hybrid ANN used in our work we refer the reader to [32] and [31]. For classification of the LCVF image we used a network configuration of 194 input (30 and 13 for additional subsampled cases), 23 output nodes, and a 40-by-40 2-dimensional SOM in the hidden layer.

Our SOM-hybrid ANN was built and tested in NeuralWare's NeuralWorks Professional II/PLus [30] graphical interface environment, then deployed using NeuralWare's Designer Pack, and embedded in our own software environment that has specifically been developed for the exploitation of high dimensional data such as large hyperspectral images. Knowledge extraction, process monitoring and visualization capabilities were also developed in-house.

Table 2. summarizes the performance of the five classifiers compared in [7], for three different band sets. ANN refers to our SOM-hybrid neural net described above. The other four classification methods are common techniques well documented in remote sensing and statistics texts. The Minimum Euclidean Distance (MED), Maximum Likelihood (MLH), Mahalonobis Distance (MHD) and Spectral Angle Mapper (SAM) classifications were performed with the ENVI image processing software [34]. Rigorous statistical estimation of classification accuracy was done as prescribed by accepted literature (*e.g.,* [35, 36, 37, 38]). The evaluation was based on a ground truth image file, painstakingly constructed for this study containing up to 200 test pixels for each class, as allowed by the size of the class. The details are provided in [7].

For training of the covariance based classifiers (MLH, MHD), a minimum of 4485 samples would be required for 194 bands and 23 classes ($(194 + 1) * 23 = 4485$), in addition to the test samples. Our 942 original training pixels sampled the classes unevenly, with some classes represented by only a handful of spectra, and thus only sufficed to classify a 13-band subset of the AVIRIS image with MLH and MHD. To have a better comparison with the MLH and MHD capabilities, a 30-band data set was classified using an augmented training set. The creation of an augmented training data set (raising the number of training samples to at least 31 in each class) presented a challenge but it was doable, with careful work by a domain expert. There was, however, no way to identify more than 31, let alone 195, reliable training pixels for every class since several of the small classes (such as B, the rhyolitic outcrop, or the small classes Q, R, S, T), contain too few pixels. Therefore the MLH and MHD classifications could only be run for the 30-band and 13-band data sets. This is a typical scenario in remote sensing, illustrating one of the pressing reasons for the need for new computational intelligence techniques.

	194-band %	classification κ	30-band %	classification κ	13-band %	classification κ
ANN	89.80	0.8925	75.02	0.7386	77.08	0.7598
MED	82.04	0.8126	72.85	0.7166	73.29	0.7187
SAM	80.21	0.7914	67.10	0.6538	67.54	0.6583
MLH			50.87	0.4822	63.23	0.6172
MHD			72.53	0.7140	73.87	0.7276

Table 2:
Summary statistics of the performance of five classifiers. Data: AVIRIS 1994 hyperspectral image of the Lunar Crater Volcanic Field, Nevada, USA.

For lack of space, to view the classification maps and for detailed discussion of the classifiers' performances we refer the reader to [7]. Table 2, taken from [7] indicates that the SOM-hybrid ANN did best with nearly 90% accuracy on the full spectral resolution, and it outperformed the other four classifiers on the lower resolutions as well. In addition to the classification accuracy Table 2. lists the so-called *kappa*-statistics, a frequently used index of classification performance in remote sensing.

Another successful recent hyperspectral image analysis with the same method, for ecosystem studies in the Grand Canyon, Arizona, USA, is reported in [39].

4 CONCLUSIONS and FUTURE DIRECTIONS

Principled applications of Self-Organizing Maps and hybrid ANNs containing SOMs are extremely promising for clustering and classification of very high-dimensional data spaces. Correct and rigorous approaches such as described in this summary allow reliable "precision mining" of highly complex data such as hyperspectral imagery, at full spectral resolution. This enables fine discrimination of a large number of surface materials in classifications, and it provides the potential of discoveries to the extent of detail that the advanced spectral imagers capture. Current and near future work is focused on thorough testing of the theories outlined and integration thereof into one powerful tool for hyperspectral image analyses.

Acknowledgments

This research was partially supported by the Applied Information Systems Research Program of NASA, Office of Space Science, NAG9–10432.

References

[1] J. Campbell, *Introduction to Remote Sensing*, The Guilford Press, U.S.A., 1996.

[2] R. O. Green, "Summaries of the 6th Annual JPL Airborne Geoscience Workshop, 1. AVIRIS Workshop," Pasadena, CA, March 4–6 1996.

[3] R. W. Basedow, D. C. Carmer, and M. L. Anderson, "HYDICE: An airborne system for hyperspectral imaging," in *Proc. SPIE*, Orlando, FL, 17-18 April 1995, vol. 2480, pp. 258–267.

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

[5] E. Merényi, "The challenges in spectral image analysis: An introduction and review of ANN approaches," in *Proc. Of European Symposium on Artificial Neural Networks (ESANN'99)*, Brussels, Belgium, 1999, pp. 93–98, D facto publications.

[6] 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.)*, Download from http://www.ece.rice.edu/ erzsebet/emann.html, 2000, Physica Verlag.

[7] E. Merényi, W. H. Farrand, and et al., "Efficient geologic mapping from hyperspectral images with artificial neural networks classification: a comparison to conventional tools," *IEEE TGARS*, p. (submitted), 2002.

[8] R. E. Arvidson and M. Dale-Bannister and. et al., "Archiving and distribution of Geologic Remote Sensing Field Experiment data.," *EOS, Transactions of the American Geophysical Union*, vol. 72, no. 17, pp. 176, 1991.

[9] W. H. Farrand, *VIS/NIR Reflectance Spectroscopy of Tuff Rings and Tuff Cones*, Ph.D. thesis, University of Arizona, 1991.

[10] W. H. Farrand and R. B. Singer, "Analysis of altered volcanic pyroclasts using AVIRIS data," in *Proceedings of the Third Airborne Visble/Infrared Imaging Spectrometer (AVIRIS) Workshop*, Pasadena, CA, March 20–21 1991.

[11] D. H. Scott and N. J. Trask, "Geology of the Lunar Crater Volcanic Field, Nye County, Nevada," Tech. Rep., USGS, 1973.

[12] M. K. Shepard, R. E. Arvidson, E. A. Guinness, and D. W. Deering, "Scattering behavior of Lunar Lake playa determined from PARABOLA bidirectional reflectance data," , vol. 18, pp. 2241–2244, 1991.

[13] S. Haykin, *Neural Networks. A Comprehensive Foundation*, McMillan, New Jersey, 1999.

[14] J. A. Benediktsson, P. H. Swain, and et al., "Classification of very high dimensional data using neural networks," in *IGARSS'90 10th Annual International Geoscience and Remote Sensing Symp.*, 1990, vol. 2, p. 1269.

[15] E. Merényi, R. B. Singer, and J. S. Miller, "Mapping of spectral variations on the surface of Mars from high spectral resolution telescopic images," *ICARUS*, vol. 124, pp. 280–295, 1996.

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

[17] Samuel Kaski, Janne Nikkilä, and Teuvo Kohonen, "Methods for interpreting a self-organized map in data analysis," in *Proc. Of European Symposium on Artificial Neural Networks (ESANN'98)*, Brussels, Belgium, 1998, pp. 185–190, D facto publications.

[18] Thomas Villmann, "Topology preservation in self-organizing maps," in *Kohonen Maps*, E. Oja and S. Kaski, Eds., Amsterdam (Holland), June 1999, Helsinki, number ISBN 951-22-3589-7, pp. 279–292, Elsevier.

[19] Th. Villmann, "Benefits and limits of the self-organizing map and its variants in the area of satellite remote sensoring processing," in *Proc. Of European Symposium on Artificial Neural Networks (ESANN'99)*, Brussels, Belgium, 1999, pp. 111–116, D facto publications.

[20] D. DeSieno, "Adding a conscience to competitive learning," in *Proc. ICNN, July 1988* **I**, New York, 1988, pp. 117–124.

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

[22] H.-U. Bauer and Th. Villmann, "Growing a Hypercubical Output Space in a Self–Organizing Feature Map," *IEEE Transactions on Neural Networks*, vol. 8, no. 2, pp. 218–226, 1997.

[23] 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*, vol. 8, no. 2, pp. 256–266, 1997.

[24] 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.)*. 2001, pp. 121–145, Springer-Verlag, Download from http://www.ece.rice.edu/ erzsebet/emann.html.

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

[26] Marc M. Van Hulle, *Faithful Representations and Topographic Maps*, Wiley Series and Adaptive Learning Systems for Signal Processing, Communications, and Control. Wiley & Sons, New York, 2000.

[27] H.-U. Bauer, R. Der, and M. Herrmann, "Controlling the magnification factor of self–organizing feature maps," *Neural Computation*, vol. 8, no. 4, pp. 757–771, 1996.

[28] J. D. Paola and R. A. Schowengerdt, "Comparison of neural network to standard techniques for image classification and correlation," in *Proc. Int'l Geosci. and Remote Sensing Symposium*, Caltech, Pasadena, CA, August 8–12 1994, vol. III, pp. 1404–1405.

[29] B. Widrow and F. W. Smith, "Pattern-recognizing control systems," in *Computer and Information Science Symposium Proceedings*, Washington, D. C., 1963, Spartan Books.

[30] Inc. NeuralWare, *Neural Computing, NeuralWorks Professional II/PLUS*, 1993.

[31] E. S. Howell, E. Merényi, and L. A. Lebofsky, "Classification of asteroid spectra using a neural network," *Jour. Geophys. Res.*, vol. 99, no. E5, pp. 10,847–10,865, 1994.

[32] E. Merényi, E. S. Howell, and et al., "Prediction of water in asteroids from spectral data shortward of 3 microns," *ICARUS*, vol. 129, pp. 421–439, 1997.

[33] M. T. Fardanesh and O. K. Ersoy, "Classification accuracy improvement of neural network classifiers by using unlabeled data," *IEEE. Trans. Geosci. and Remote Sens.*, vol. 36, no. 3, pp. 1020–1025, 1998.

[34] Research Systems Inc., *ENVI v.3 User's Guide*, 1997.

[35] K. Fitzpatrick-Lins, "Comparison of sampling procedures and data analysis for a land-use and land-cover map," *Photogrammetric Engineering & Remote Sensing*, vol. 47, no. 3, pp. 343–351, 1981.

[36] P. J. Curran and H. D. Williamson, "Sample size for ground and remotely sensed data," *Photogrammetric Engineering & Remote Sensing*, vol. 20, pp. 31–41, 1986.

[37] R. G. Congalton, "A comparison of sampling schemes used in generating error matrices for assessing the accuracy of maps generated from remotely sensed data," *Photogrammetric Engineering & Remote Sensing*, vol. 54, no. 5, pp. 593–600, 1988.

[38] R. G. Congalton and C. Green, *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, Lewis Publishers, U.S.A., 1999.

[39] E. Merényi, W. H. Farrand, L.E. Stevens, T.S. Melis, and K. Chhibber, "Mapping Colorado River ecosystem resources in Glen Canyon: Analysis of hyperspectral low-altitude AVIRIS imagery," in *Proc. ERIM, 14th Int'l Conference and Workshops on Applied Geologic Remote Sensing, 4–6 November, 2000, Las Vegas, Nevada*, Download from http://www.ece.rice.edu/ erzsebet/emann.html, 2000.