Classification of hyperspectral images using wavelet transforms and neural networks

Tom Moon

Montana Tech of the University of Montana
Department of Geophysical Engineering
1300 West Park Street, Butte, Montana 59701-8997

and

Erzsébet Merényi

Lunar and Planetary Laboratory, PIRL,
University of Arizona, Tucson, AZ 85721

June 23, 1995

To appear in the Proceedings of the Annual SPIE Conference,
July 9-14, 1995, San Diego, CA;
SPIE Volume 2569, "Wavelet Appllications and Signal Processing III"
Eds. A.F.Laine, M.A. Unser, and M.V. Wickerhauser
paper #2569-66

also to be submitted to
IEEE Transactions on Geoscience and Remote Sensing

T. Moon:
ph: (406) 496-4350, fax: (406) 496-4133
internet: tom@jove.mtech.edu

E. Merényi:
ph: (520) 621-4073, fax: (520) 621-4933
internet: erzsebet@lpl.arizona.edu

keywords: AVIRIS, hyperspectral, neural networks, wavelets, data compression

Classification of hyperspectral images using wavelet transforms and neural networks

Tom Moon

Montana Tech of the University of Montana, Department of Geophysical Engineering
1300 West Park Street, Butte, Montana 59701–8997

Erzsébet Merényi

University of Arizona, Planetary Image Research Laboratory, Lunar and Planetary Laboratory
Tucson, Arizona 85721

ABSTRACT

The Airborne Visible InfraRed Imaging Spectrometer (AVIRIS), presently being flown by the Jet Propulsion Laboratory, acquires images of the earth in the visible and reflected infrared. The wavelengths of the measured radiation range from about 400 nm to 2400 nm and are divided into 224 contiguous channels having a nominal spectral bandwidth of 10 nm. This means a high resolution radiance spectrum is acquired for each 20 m x 20 m ground cell in the AVIRIS scene. Geologic mapping from such data is possible by classifying each pixel based on the distinctive spectral signatures recorded in the channels. Artificial neural networks (ANN) have used these spectra successfully to classify an AVIRIS subscene of the Lunar Craters Volcanic Field (LCVF) in Nye County, Nevada. The size and number of spectra in an AVIRIS scene makes classifying these images a computationally intensive task. By classifying the data in a compressed format, savings in computer time may be realized. The wavelet bases have the desirable property of rendering signals similar to the AVIRIS spectra sparse in the wavelet domain. In this investigation, the discrete wavelet transform was applied to the spectra. This produced a set of wavelet coefficients for the spectra that could be made sparse with seemingly little loss of accuracy. Small subsets of the wavelet coefficients were used to classify the LCVF scene by ANN. The degree to which information was lost in the wavelet transform and the elimination of wavelet coefficients from the classification was assessed by making comparisons between the different ANN classifications. The ANN was chosen over more conventional classifiers because of its proven sensitivity in distinguishing subtle but geologically relevant features in these spectra.

KEYWORDS: AVIRIS, hyperspectral, neural networks, wavelets, data compression

1. BACKGROUND

The Jet Propulsion Laboratory's Airborne Visible InfraRed Imaging Spectrometer (AVIRIS) is an experimental earth imaging instrument with high spectral resolution. AVIRIS acquires a 224 channel spectrum of the upwelling radiance as a function of wavelength for each of the 20 m x 20 m ground cells in its scene. The 224 channels range from about 400 nm to 2500 nm and have a nominal spectral bandwidth of 10 nm. Figure 1 illustrates the two spatial dimensions and third spectral dimension of an AVIRIS hyperspectral data set or hypercube. This figure was constructed from an AVIRIS subscene of the Lunar Craters Volcanic Field (LCVF) in Nye County, Nevada. Thirteen representative spectra from this hypercube are plotted in figure 2.

Geologic mapping from AVIRIS data is possible by classifying the shape of the spectra recorded at each pixel. The sometimes subtle spectral differences between geological units along with the high dimensionality of the spectra presents a challenging pattern recognition task. Artificial neural networks (ANN) have been found to be powerful tools in separating geological units in the LCVF subscene. Merényi et al. (1) was successful in separating 13 classes from this subscene by using an ANN (see figure 3). Their ANN, similar to the one described below in section 3, used 158 channels of the spectra as its input.

A classification of an AVIRIS scene can be computationally intensive because of the size of the hyperspectral data set. If the number of inputs to the ANN can be reduced without significant loss of information, then it should be possible to successfully classify a scene based on a reduced or compressed data set. A smaller data set would mean a smaller ANN and a savings in time and programming complexity for the classification. The wavelet transform was chosen in this investigation for the compression because of the property of the wavelet bases to render certain signals sparse in the wavelet domain (2). The distinctive peaks and valleys of the hypercube spectra suggested that the N=2 orthonormal wavelet basis of Daubechies (DAUB4) was a reasonable choice for the task (3). We present in this study several classifications of the LCVF subscene based on wavelet transforms of the spectra. As described below in section 4, discrete wavelet transforms of the spectra were taken and subsets of the wavelet coefficients were used to train an ANN and classify the LCVF subscene. These classifications give a good indication of the information loss in the wavelet transform and compression. The following three sections briefly discuss the geology of the LCVF, the ANN, and the wavelet transform. Section 5 presents the results of our investigation.

2. THE GEOLOGY OF THE LUNAR CRATERS VOLCANIC FIELD

In the summer of 1989 NASA sponsored the Geologic Remote Sensing Field Experiment in which the AVIRIS instrument was used to image the LCVF in northern Nye County, Nevada. The 256 x 256 pixel subscene of the LCVF used in this study was collected on September 29, 1989 at 11:44 PDT (see figure 1). Vegetation in LCVF is predominant only within washes and near springs (4). The geological units are principally Quaternary (< 1 million years) and Tertiary (1 to 70 million years) in age. Quaternary basaltic pyroclastic and flow deposits lie atop ignimbrites and silicic lava flows of Tertiary age and in turn are overlain by Quaternary alluvial and playa deposits. The LCVF subscene in this study contains oxidized basaltic cinder deposits, the southern half of the Lunar Lake Playa, and outcrops of the Rhyolite of Big Sand Spring Valley (5). The legend to figure 3 lists the geologic types found in the LCVF subscene. This is based on the classification by Merényi et al. (1) which is reproduced in figure 3.

3. THE ARTIFICIAL NEURAL NETWORK

An ANN is a densely interconnected set of a large number of simple processing units, organized into different functional groups. The units work parallel in time, receiving and passing processed data among themselves as their connection paths define the flow of information. There is an input layer where data are introduced, one or more hidden layers to process the data, and one output layer where class predictions are represented. The system is a learning machine that derives its knowledge from examples shown to it many times. In our case the representative spectra for each surface cover type, together with their class designations, make up the training material. The learning takes place by repetitive adjustment of connection weights (the analog of the synaptic strength in biological nervous systems) among the processing units (the nerve cells). The network is trained until it correctly classifies all training samples. Then it can make class predictions for unknown spectra based on the class properties that it derived from the training samples. A particularly attractive ability of ANNs is the modeling of complex shapes (such as spectra) without analytical description. This is due, in great part, to their non–linear processing capabilities and high connectivity. They have proved to produce equal or higher quality classification results than conventional classifiers (6), while being more tolerant of noise.

There are numerous ANN paradigms developed for various kinds of problems. The particular ANN paradigm we applied here is a hybrid architecture. It contains a 2–D Kohonen Self–Organizing Map (7) in its hidden layer, connected to an output layer with a Widrow–Hoff learning rule (8), which is similar to Back–Propagation. The particular implementation is by NeuralWare Inc, (1991) (9). This network scales up well to handle large number of channels, training convergence is fast and easy relative to Back–Propagation networks, as demonstrated by previous works for up to 300 channels and well over a dozen classes (1, 10).

Training difficulties often discourage people from using ANNs. Dimensionality reduction (and with that the loss of subtle discrimination) often is accepted in order to use a Back–Propagation network. The ANN described above gave more accurate results than Back–Propagation for underrepresented and unevenly represented classes in the 158–channel AVIRIS LCVF spectral image (1) and in developing asteroid taxonomy (10). By first forming a cluster map of the data in an unsupervised learning phase that takes place exclusively in the 2–D Kohonen layer, it is less likely to learn inconsistent class labels in the subsequent supervised phase than Back–Propagation. This results in better generalization from even a small number of samples and in higher classification accuracy (11). This capability is especially valuable in remote sensing problems in view of costly field sampling for training data and for the verification of results.

4. THE WAVELET TRANSFORM

The AVIRIS hypercube of the LCVF subscene contains 256 x 256 radiance spectra, one for each pixel in the scene (see figure 1). The original 224 channels (or bands) in the scene were first reduced to 158 by removing spectrometer overlap, excessively noisy channels and channels which fell within atmospheric absorption. Figure 2 shows the 158 channel spectra from 13 representative pixels in the scene. In Merényi et al. (1), all 158 of these bands were used in their ANN classification. By changing the basis system of the spectral data, it may be possible to reduce this number of inputs to the classifying ANN. Applying the DAUB4 wavelet transform to one of the LCVF spectra renders what appears to be a sparse set of wavelet coefficients (see figure 4a). Much of this is due to zero padding the 158 bands out to 256 for the transform; however, it's clear that the spectra are dominated by a relatively few large coefficients. Our objective was to see whether or not a subset numbering 158 or less of these wavelet coefficients could be used successfully in a classification by an ANN. The simplest method for choosing the subset was to use only the largest wavelet coefficients. To determine the amount of information loss in eliminating wavelet coefficients, comparisons were made between the new classifications and the classification of Merényi et al. (1) By running classifications on the inverse transformed wavelet coefficients, the amount of loss in the wavelet transform itself was assessed.

To see the immediate effects of eliminating (zeroing) the smallest coefficients, three inverse wavelet transforms are shown in figure 4b. Line A is the original 158 channel spectrum from a pixel in class A of Merényi et al. (1) All but the largest 158, 118 or 79 wavelet coefficients were zeroed before the inverse wavelet transform to generate the three spectra A@100%, A@75%, and A@50% respectively. The percentages correspond to the number of non–zero wavelet coefficients used in the inverse transform relative to 158. As more coefficients were eliminated, the smallest peaks and valleys in the spectra were removed, but the spectra retain major features. Treating the spectra as vectors, the Euclidean distance between A and A@50% is 68 (the length of data vector A is 3046). The angular distance between A and A@50% is 1.28^{o}. The Euclidean distances between spectra of the 13 classes of Merényi et al. (1) typically range between 300 to 8000, and their angular distances range between 1.2^{o} and 14.2^{o}. Although this suggests that little information might be lost by using these subsets of the wavelet coefficients, the importance of these smaller features are perhaps best assessed by comparing ANN classifications. This is discussed in the following section.

5. RESULTS

Eight ANN classifications are presented in this study. They are listed along with their descriptions and abbreviations in table 1. They are hereafter referred by their abbreviations. All classifications are of the LCVF 256x256 pixel subscene. The classifications fall into two categories: classifications based on spectra in the 158 channel spectral domain, and classifications based on the wavelet coefficients of the spectra after transformation to the wavelet domain. Classifications REF, S118, S79 fall in the first category. W158, W118, W79, Wrl4567, and Wrl6 are based on wavelet coefficients and therefore fall into the second category. Although qualitative comparisons between classifications would require color images, grayscale renditions of the classifications are presented in figures 5 through 8 (color reprints are available upon request). Table 2 presents a more quantitative summary of the results.

Table 1. Descriptions of the eight ANN classifications presented in this paper. The classifications are referred to in the text by their abbreviations. Figures of the classifications are listed at right.

abbreviation	description	figure
REF	This is a classification of Merényi et al. (1) It is based on the original 158 channel spectra of the LCVF subscene.	3 , 5a
W158	In this classification only the largest 158 wavelet coefficients of the DAUB4 wavelet transform of the LCVF spectra were used.	5b
S118	The spectra are wavelet transformed and all but the largest 118 wavelet coefficients are zeroed. Then, the wavelet coefficients are inverse transformed back to the spectral domain of 158 channels. This classification is based on this 158 channel spectra.	6a
W118	Classification using only the largest 118 wavelet coefficients of the wavelet transform of the LCVF spectra.	6b
S79	As in classification S118, this classification is based on the back transformation of the largest 79 wavelet coefficients (all other coefficients were zeroed).	7a
W79	Classification using only the largest 79 wavelet coefficients of the wavelet transform of the LCVF spectra.	7b
Wrl4567	The wavelet transform of a 158 channel spectrum generates a set of 256 wavelet coefficients. Within this set of coefficients are wavelet resolution levels –1, 0, 1, ... , 6, and 7 (12). This classification used only the wavelet coefficients of resolution levels 4, 5, 6 and 7.	8a
Wrl6	Classification using only the wavelet coefficients of resolution level 6.	8b

In table 2, the Merényi et al. (1) (REF) classification acts as our "ground truth." Geologic maps and field experiments do in fact confirm that the thirteen classes found in REF correspond to geologic units. Each "hit" column in this table gives the percentage of pixels which were classified in agreement with REF. The "miss" columns give the percentage of pixels erroneously placed in another class. The difference between the hit and miss percentages is the percentage of pixels unclassified. The bottom row lists the average hit and miss percentages for the classifications. The ANN in each case was trained on exactly the same pixels as in the REF case. Training accuracy was 100% for most classes, with a minimum of 90%, in all cases except for the last two, where minimum training accuracy was 60% across classes. More detailed inspection of the confusion matrices for the training and test cases allow tracking which classes suffer loss of distinction among themselves as the spectral information decreases.

Table 2. Summary of classification results. The "class" and "pixels" columns list the 13 classes and the number of pixels in each class from the classification of Merényi et al. (1) (REF). Using the REF classification as "truth", the "hit" column for each classification gives the percentage of pixels correctly classified. The "miss" column gives the percentage of pixels incorrectly classified into one of the other 13 classes. The difference between the hit and miss percentages gives the number of pixels unclassified.

	REF	W158		S118		W118		S79		W79		Wrl4567		Wrl6	
class	pixels	hit	miss	hit	miss	hit	miss	hit	miss	hit	miss	hit	miss	hit	miss
A	2076	74	2	78	5	91	3	86	4	92	8	81	12	65	19
B	491	79	20	64	36	89	11	95	3	77	20	81	16	42	33
C	2783	92	8	94	6	86	14	72	17	83	15	72	15	28	38
D	3722	55	0	38	0	31	69	47	0	66	0	94	0	83	2
E	8257	56	0	63	0	50	43	85	0	69	28	16	0	45	14
F	1170	59	41	49	51	49	51	41	59	68	32	31	70	16	72
G	1637	75	17	37	63	83	9	53	19	91	9	19	59	10	44
H	7907	98	2	80	16	86	14	94	1	92	8	79	16	52	36
I	13400	77	23	61	39	79	21	73	27	80	20	61	29	54	23
J	1064	68	32	77	23	72	29	71	29	59	41	59	41	18	72
K	1694	87	13	78	22	60	40	87	12	82	6	75	16	25	41
L	8772	60	39	61	36	57	27	96	4	64	32	55	41	22	54
M	467	97	0	77	16	84	12	73	27	77	23	96	4	75	11
avg		75	15	66	24	71	26	75	15	77	19	63	24	41	35

A number of questions can be raised by these classifications. The two broader issues of using data in a compressed format for an ANN classification and the effect the wavelet transform itself has on the classification are the primary focus of this study. Inspection of table 2 and figures 5 through 8 suggest that the ANN classification can be fairly sensitive to lossy compression of the spectra. Comparing classifications W158, W118 and W79, however, does not indicate any simple trends with respect to the amount of compression. For example, hit percentages for classes A and I increase from W158 to W79, but decrease for classes C and M. Other classes show no obvious trends. There is some indication that information in the spectra is spread throughout the resolution levels. The process of removing the smaller wavelet coefficients for the W158, W118 and W79 classifications tends to eliminate more coefficients from the higher resolution levels than the lower. While classes C, H and M were barely affected by this type of lossy compression, classes D, E, and F were significantly affected. Classifications Wrl4567 and Wrl6 further indicate this information spread throughout the resolution levels.

The classification Wrl4567 used only the wavelet coefficients in resolution levels 4, 5, 6, and 7 (12). Resolution levels 4 through 7 are the wavelet coefficients for wavelet functions which render the peaks and valleys in the spectra ranging in width from about 16 channels to 2 channels respectively. Classification Wrl6 used only the wavelet coefficients in resolution level 6. This resolution level contains the wavelets that render features in the spectra that are approximately 4 channels in width. Comparison between Wrl4567 and W158 indicates that the distinguishing features of class D are the narrower, lower amplitude peaks and valleys in the spectra, while the broader, low frequency features set class H spectra apart. All this suggests that a more complex method of lossy compression is required for a generally successful classification based on subsets of wavelet coefficients.

Classifications W118 and S118 used data which are essentially wavelet transform pairs, as are the data used in classifications W79 and S79. Assuming no loss in the wavelet transform, it appears from table 2 and figures 6a and 6b that the ANN perceives W118 and S118 differently. This also appears to be true for W79 and S79 (see figures 7a and 7b). Taking the hit and miss percentages at face value shows that W118 and W79 performed slightly better on the average than their counterparts S118 and S79 respectively (although perhaps not significantly better). Closer inspection shows that hit percentages for some classes were higher in the wavelet domain than in the spectral domain, while for others the opposite was true. The reason for this is not entirely clear at this time, but probably tells us as much about the ANN as the wavelet transform.

6. ACKNOWLEDGMENTS

This research was supported in part by the NASA/JOVE project. A special thanks to Karl Glass for suggesting the problem and the Jet Propulsion Lab for the use of their AVIRIS data. The computing facilities of the Lunar and Planetary Laboratory and the Planetary Image Research Laboratory, of the University of Arizona, software support by James Winburn and contributions by NASA Space Grant Interns Trevor Laing and Michael Shipman are also cheerfully acknowledged.

7. REFERENCES

1. E. Merényi, R. B. Singer, R.B. and W. H. Farrand, "Classification of the LCVF AVIRIS test site with a Kohonen artificial neural network," *Proceedings of the Fourth Airborne Geoscience Workshop, Washington, D.C., Oct. 25–29, JPL Publication 93–26 1993*, Vol. 1, pp. 117–120, 1993.

2. W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, *Numerical Recipes in C, 2ed.*, Cambridge University Press, pp. 591–606, 1992.

3. I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Communications on Pure and Applied Mathematics*, Vol. XLI, pp. 909–996, 1988.

4. D. H. Scott and N.J. Trask, "Geology of the Lunar Crater Volcanic Field, Nye County, Nevada," *USGS Prof. Paper 599–I*, 1971.

5. E. B. Ekren, E.N. Hinrichs and G.L. Dixon, "Geologic map of the wall quadrangle, Nye County, Nevada," *USGS Misc. Geol. Inv. Map I–719, scale 1:48,000*, 1973.

6. W. Huang and R. Lippman, "Comparisions Between Neural Net and Conventional Classifiers," *IEEE First International Conference on Neural Networks*, Vol. IV, pp. 485–494, 1987.

7. T. Kohonen, *Self–Organization and Associative Memory*, Springer–Verlag, New York, 1988.

8. B. Widro and F. W. Smith, "Pattern–recognizing control systems," *Computer and Information Science Symposium Proceedings, Spartan Books, Washington, D.C.*, 1963.

9. *NeuralWare, Inc.*, Neural Computing, NeuralWorks Professional II, 1991.

10. E. S. Howell, E.S., E. Merényi, and L. A. Lebofsky, "Classification of asteroid spectra using a neural network," *JGR Planets*, Vol. 99, No. E5, pp. 10,847–10,865, 1994.

11. J.A. Benediktsson, P.H. Swain, O. K. Ersoy and D. Hong, "Classification of very high dimensional data using neural networks," *IGARSS'90 10th Annual International Geoscience and Remote Sensing Symp.*, Vol. 2 , p. 1269, 1990.

12. D. E. Newland, *An Introduction to Random Vibrations, Spectral and Wavelet Analysis, 3ed.*, Longman Scientific and Technical, John Wiley and Sons, Inc., New York, 1993.

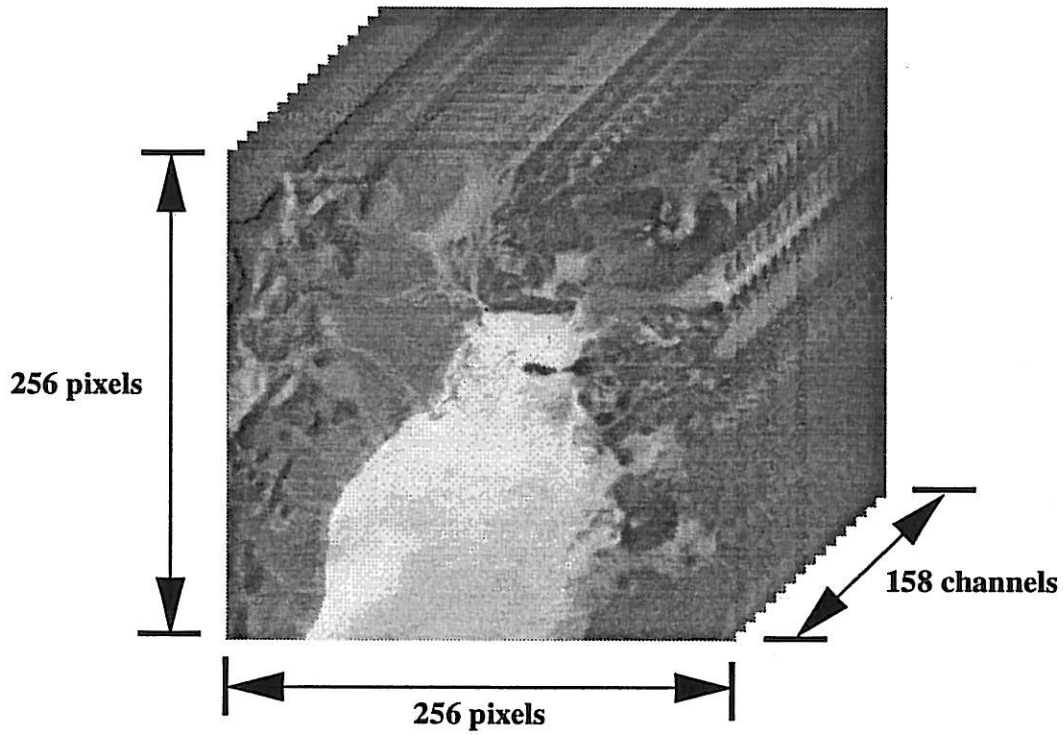


Figure 1. The LCVF hypercube is a deck of 158 perfectly registered images, one for each channel. It is also a set of 256 x 256 spectra, one for each pixel. Each spectrum contains 158 spectral channels in the LCVF subscene.

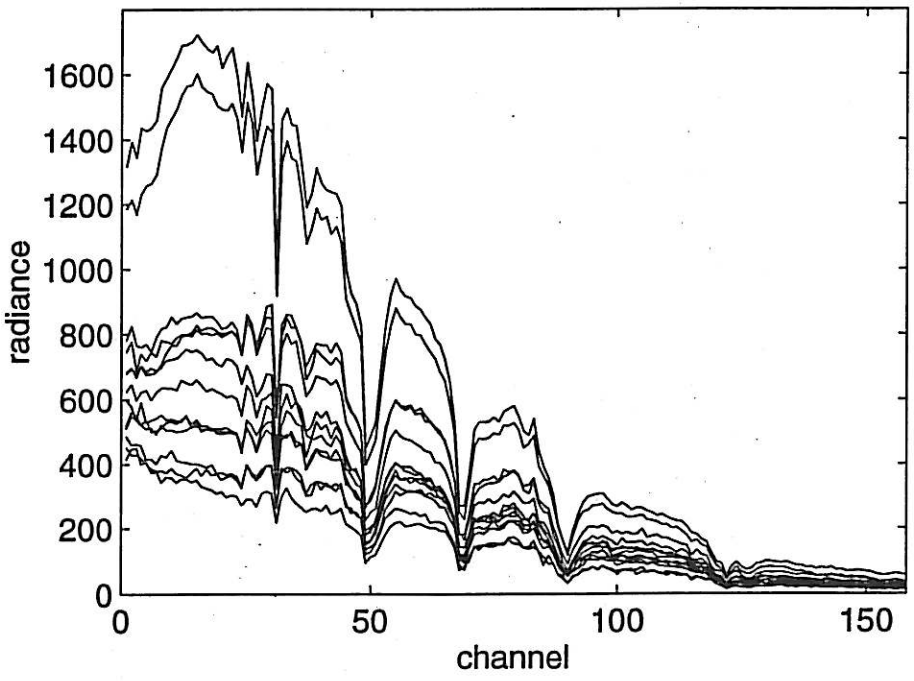
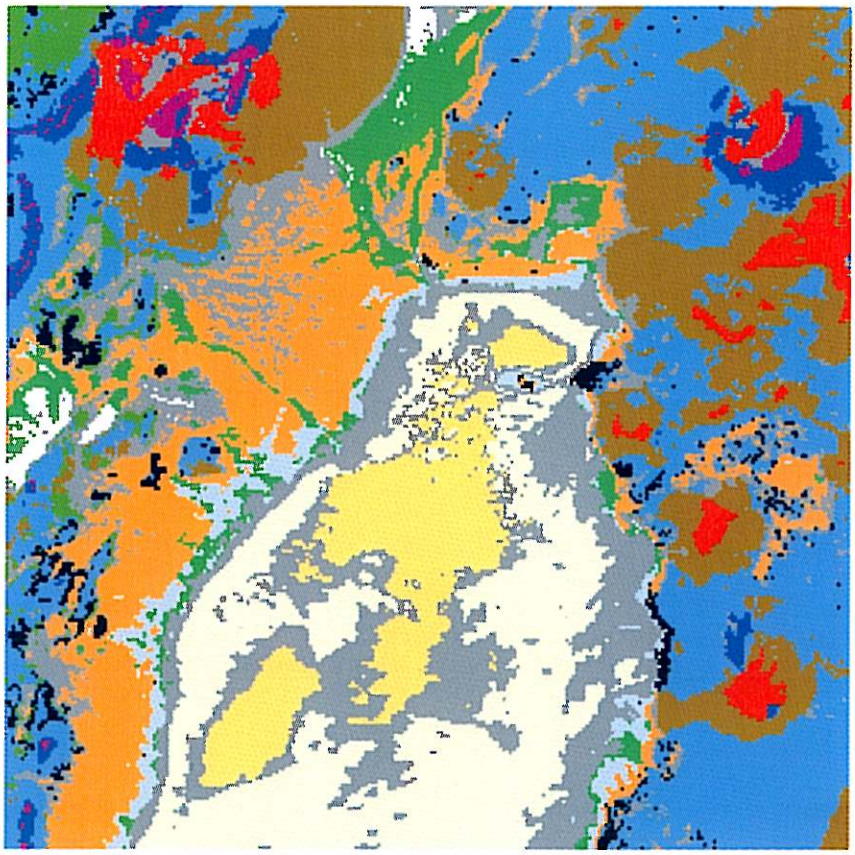


Figure 2. Thirteen representative spectra from the LCVF subscene. Each spectrum belongs to one of the A through M classes ofMerényi et al. (1) (see also figure 3).

A - highly oxidized cinders
B - rhyolite of Big Sand Springs Valley
C - vegetation type 1, probably grass
D - southern playa
E - northern playa (clay-rich)
F - young basalt flows
G - Shingle Pass Tuff
H - Quaternary alluvium derived from silicic volcanics
I - old basalt flow
J - vegetation type 2, probably scrub brush
K - basalt cobbles on playa surface
L - ferric oxide rich soil
M - poorly oxidized dark cinders
U - unclassified



A B C D E F G H I J K L M U

Figure 3. The Merényi et al. classification. Geologic maps and field experiments confirm that these 13 classes correspond to known geologic units.

Figure 4a

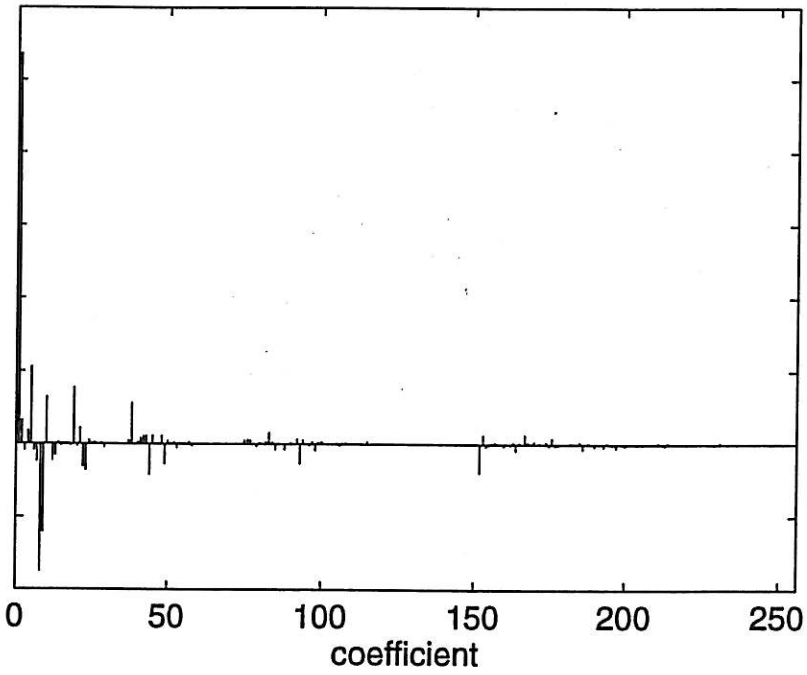


Figure 4b

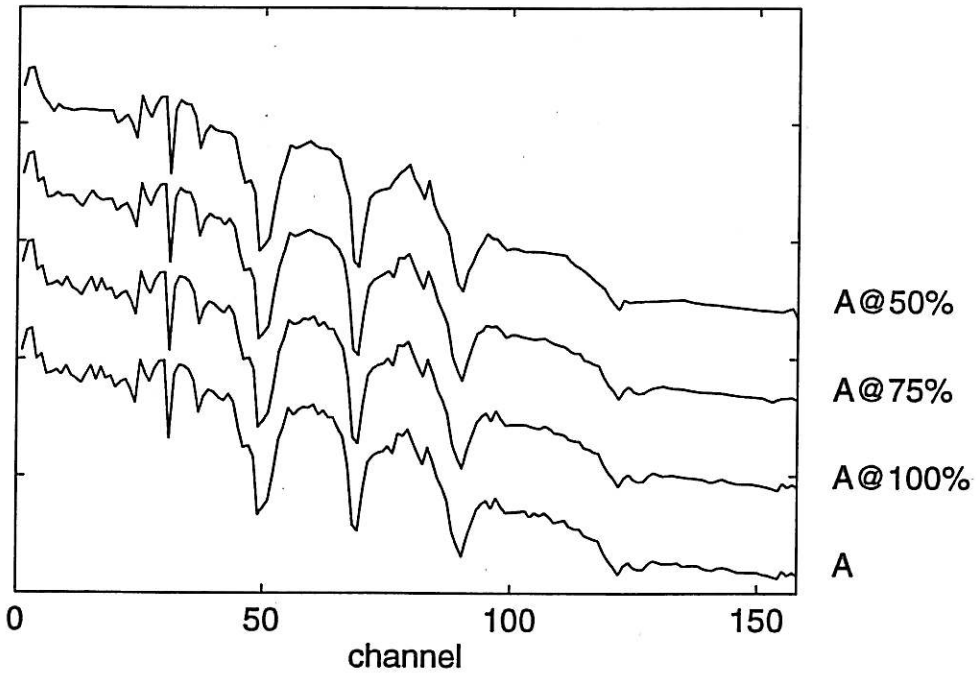


Figure 4. **(a)** The DAUB4 wavelet transform of the spectrum located at row 24 and column 32 in the LCVF hypercube of figure 1. **(b)** Line "A" is a spectrum from a pixel in class A of Merényi et al. (1). The DAUB4 wavelet transfrom of this spectrum is shown in (a). Line "A@100%" is the spectrum obtained by inverse transforming the wavelet coefficients of figure 4a after all but the largest 158 have been zeroed. Likewise, line "A@75%" is the spectrum obtained by inverse transforming the largest 118 wavelet coefficients and line "A@50%" is the spectrum obtained by inverse transforming the largest 79 coefficients. The "%" refers to the percentage of non–zeroed wavelet coefficients relative to 158.

Figure 5a. Classification REF

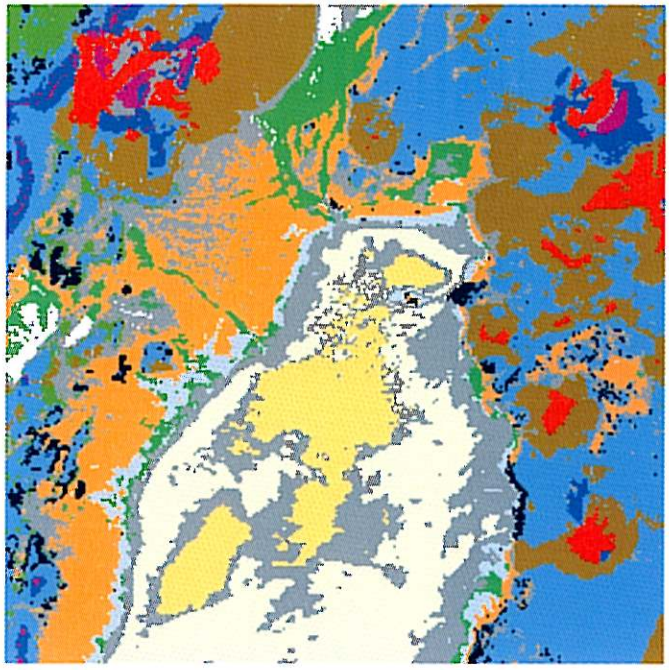


Figure 5b. Classification W158

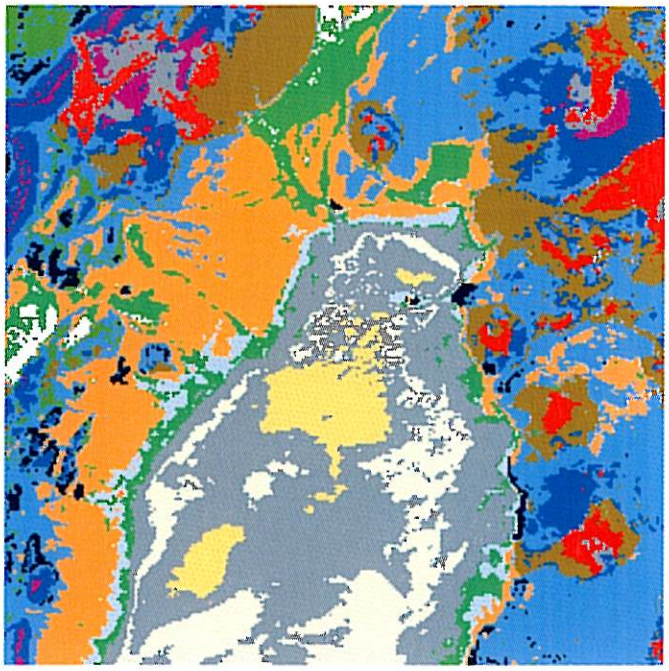


Figure 6a. Classification S118

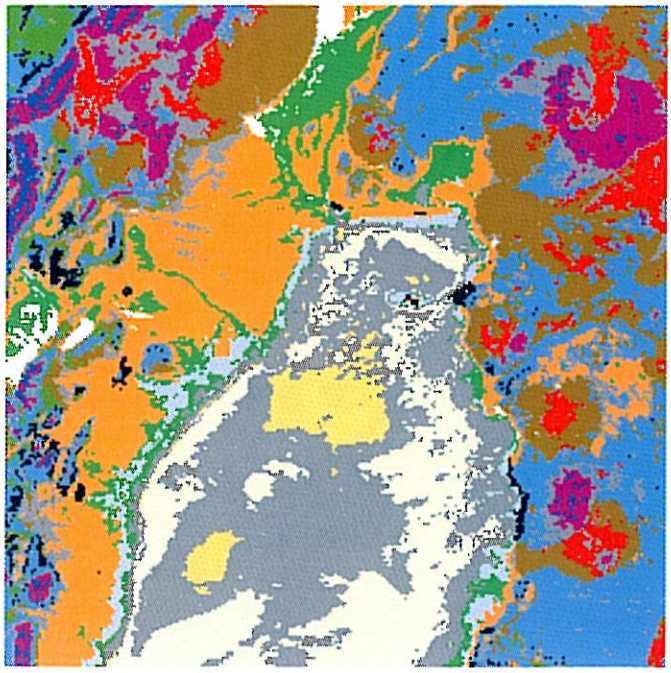


Figure 6b. Classification W118

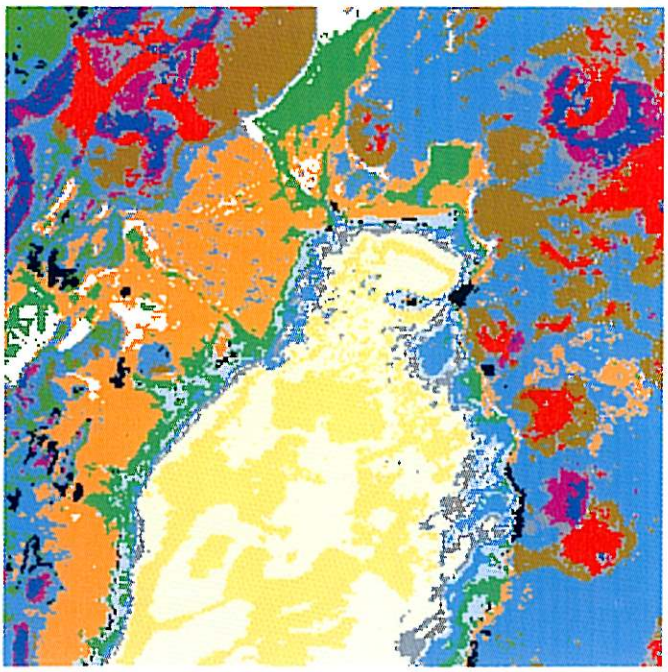


Figure 7a. Classification S79

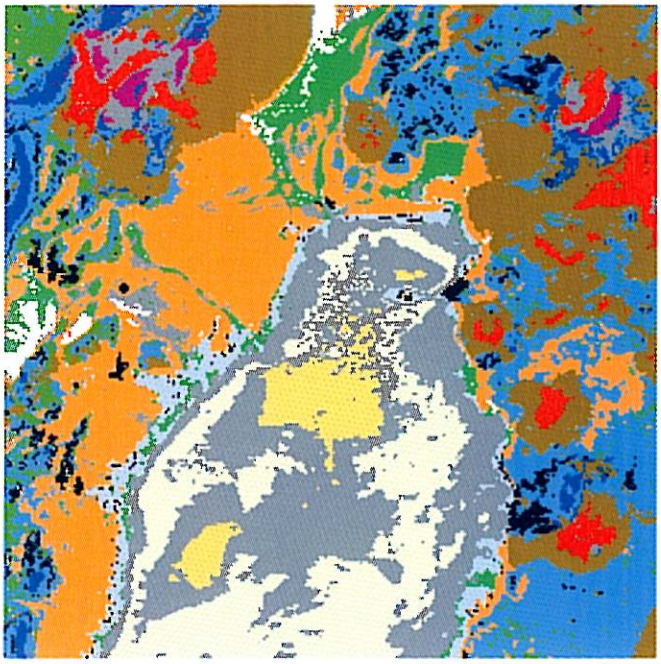


Figure 7b. Classification W79

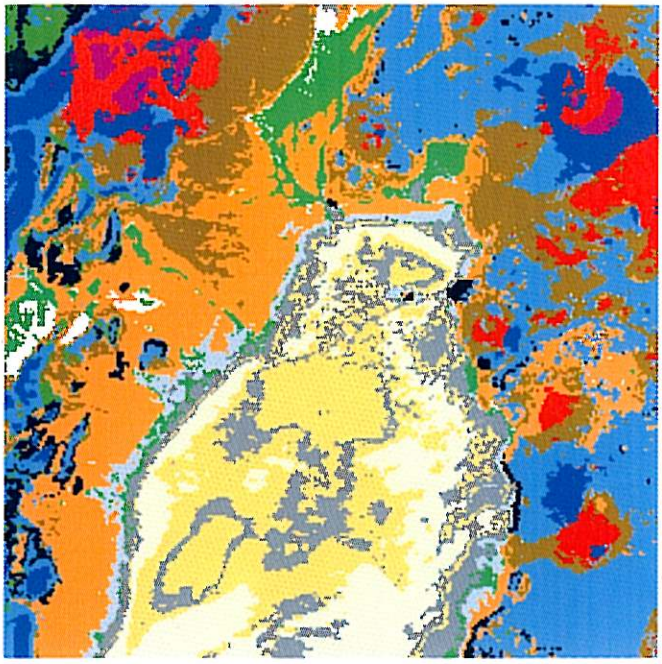


Figure 8a. Classification Wrl4567

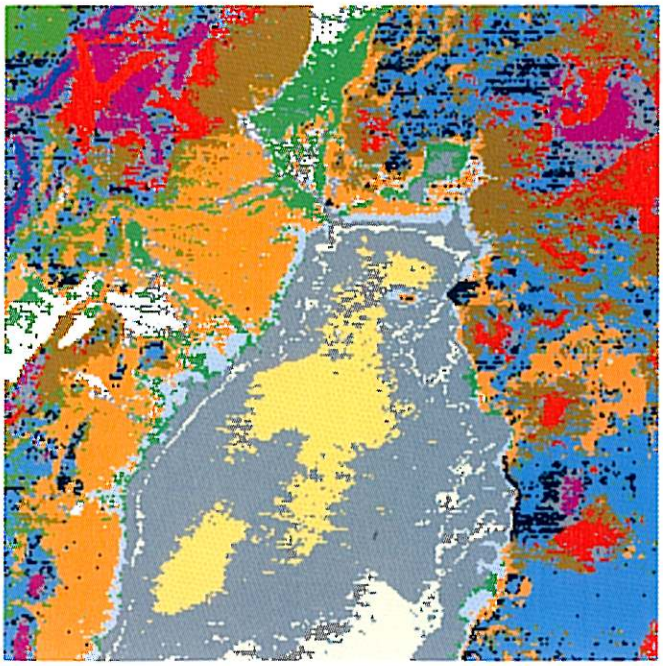


Figure 8b. Classification Wrl6