ELSEVIER

Lossy Compression of Hyperspectral Data Using Vector Quantization

Michael J. Ryan and John F. Arnold**

Efficient compression techniques are required for the coding of hyperspectral data. Lossless compression is required in the transmission and storage of data within the distribution system. Lossy techniques have a role in the initial analysis of hyperspectral data where large quantities of data are evaluated to select smaller areas for more detailed evaluation. Central to lossy compression is the development of a suitable distortion measure, and this work discusses the applicability of extant measures in video coding to the compression of hyperspectral imagery. Criteria for a remote sensing distortion measure are developed and suitable distortion measures are discussed. One measure [the percentage maximum absolute distortion (PMAD) measure] is considered to be a suitable candidate for application to remotely sensed images. The effect of lossy compression is then investigated on the maximum likelihood classification of hyperspectral images, both directly on the original reconstructed data and on features extracted by the decision boundary feature extraction (DBFE) technique. The effect of the PMAD measure is determined on the classification of an image reconstructed with varying degrees of distortion. Despite some anomalies caused by challenging discrimination tasks, the classification accuracy of both the total image and its constituent classes remains predictable as the level of distortion increases. Although total classification accuracy is reduced from 96.8% for the original image to 82.8% for the image compressed with 4% PMAD, the loss in accuracy is not significant (less that 8%) for most classes other than those that present a challenging classification problem. Yet the compressed image is 1/17 the size of the original. ©*Elsevier Science Inc., 1997*

*School of Electrical Engineering, University College, The University of New South Wales, Australian Defence Force Academy, Canberra, Australia

Address correspondence to M. J. Ryan, School of Electrical Engineering, Univ. Coll., Univ. of New South Wales, Austr. Defence Force Acad., Canberra, ACT 2600, Australia.

Received 9 October 1996; revised 12 March 1997.

INTRODUCTION

The next decade will see a dramatic increase in the quantity of available remotely sensed data due to the launching and commissioning of sensors that, in addition to high spatial resolution, will have high spectral resolution covering up to several hundred bands. Since these hyperspectral data must be transmitted to the ground and distributed to users, the economics of transmission and storage dictate that image compression will become an essential feature of handling such large volumes of data. Efficient compression techniques are therefore required for the coding of remotely sensed images.

There are two broad types of image compression: lossless and lossy compression. The essential difference between the two is whether the original image can be precisely regenerated from the compressed data. The type of information that can be discarded, however, will depend on which data is unnecessary for a given application. There is therefore a role for lossless compression in the transmission and storage of data within the distribution system.

Lossy techniques have a role in the initial analysis of hyperspectral data where large quantities of hyperspectral data are evaluated to select smaller areas for detailed evaluation. To reduce the cost of purchase of these data and the storage required to retain and process them, a suitable distortion measure could be used to produce lossy data at an acceptable distortion level and at much reduced image sizes. The purchase price, transmission times, and storage requirements for these images can therefore be reduced dramatically. Central to lossy compression is the development of a suitable distortion measure so that the performance of an algorithm on lossy data produces results, from which its performance on the original data can be inferred.

This article begins with a short introduction to vector quantization as a compression technique. Distortion measures for image and video compression are briefly

REMOTE SENS. ENVIRON. 61:419–436 (1997)

655 Avenue of the Americas, New York, NY 10010

0034-4257/97/$17.00
PII S0034-4257(97)00054-0

discussed and criterion developed for a suitable measure for coding of remotely sensed images. The percentage maximum absolute distortion measure (PMAD) is proposed as the most suitable measure and is then applied to the compression of a small hyperspectral image. The effect of compression on classification (before and after feature extraction) is then examined for various levels of allowed distortion.

Original contributions reported in the article are: development of criteria specifying a suitable distortion measure for hyperspectral data, and development of the PMAD measure, which is demonstrated to assist in the compression of a hyperspectral image to 1/17 of its original size with less than 10% loss of classification accuracy for well-behaved classes.

VECTOR QUANTIZATION

The structure of hyperspectral images reveals spectral responses that can be considered as brightness vectors which would seem ideal for compression by vector quantization. The first step in vector quantization is the decomposition of an image into a set of vectors from which a training set of vectors is selected. A codebook of representative vectors is then generated from the training set using an iterative clustering algorithm. Finally, each of the image vectors is quantized to the closest representative vector in the codebook based on the minimization of an appropriate distortion measure. Compression is achieved by replacing each vector in the image by a scalar label. Reconstruction of the image is performed by a look-up procedure using the label as an entry to a table containing the representative vectors of the codebook. A simple vector quantization scheme is illustrated in Figure 1.

Lossless compression requires the transmission of both the codebook addresses and the exact difference between the reconstructed and original vectors, so that the receiver can precisely reconstruct the original image. In lossy compression, the distortion measure is used as a guide to reduce the entropy of the difference image to achieve a desired encoded entropy for an acceptable level of distortion (or, conversely, for an acceptable level of quality) in the reconstructed image.

The distortion measure used must allow the encoder, on a pixel-by-pixel basis, to discard information from the image on the basis that it is not required to maintain the desired level of quality. This implementation on a pixel-by-pixel basis is necessary to allow real-time compression as an image is recorded by the sensor and also to allow efficient off-line compression of an image for archival purposes. Therefore, the key to lossy coding of hyperspectral images is the development of a suitable distortion measure.

Due to the wide dynamic range of hyperspectral images, some normalization of the input vectors is required. The preferred technique for this work is mean-normalized vector quantization (M-NVQ) (Ryan and Arnold, 1997), in which input vectors are preprocessed by normalizing them in accordance with the average internal reflectance (or the mean) of each vector. M-NVQ has been demonstrated to be the preferred vector quantization technique for application to hyperspectral data, as it provides the best compromise between the reduction in difference image entropy and the increase in entropy of the normalization parameters (such as vector means) that must be included in the compressed image. These results compare favorably with an extensive array of DPCM techniques for the lossless compression of AVIRIS data.

DISTORTION MEASURES FOR IMAGE AND VIDEO COMPRESSION

The objective of lossy image compression is to code an image with as few symbols as possible while maintaining the quality of the reconstructed image at some acceptable standard. That is, the aim is to obtain a compression strategy that minimizes the average distortion in the image.

Models of the Human Visual System

In the case of video images the acceptable standard for reconstruction is set by the receiver of the image—the human observer. At an early stage in video image coding investigations, therefore, it was recognized that research into image compression would be enhanced by an understanding of the human visual system so that an image quality measure could be found that would correspond to the way in which the human visual system assesses image fidelity. Many attempts have been made to derive suitable distortion measures based on the various models of the human visual system, producing a large array of candidate measures of image quality. However, despite these investigations, an accurate visual model, and thence an accurate distortion measure, has not been found since visual perception is still not well understood.

Additionally, the most appropriate test of video and image coding is whether the quality of the reconstructed image is adequate, a decision best made by the viewer. Most image coding systems are therefore assessed through a subjective rating methodology (CCIR, 1974). However, to aid the coding process, most video compression relies on distortion measures such as the mean square error (MSE) as a means of comparing coding techniques.

Although there is no theoretical justification for its use, the MSE distortion measure has been used extensively because of its mathematical tractability and its computability. It also has some intuitive appeal since large errors are given more importance than small ones. It is, however, a poor criterion for video images as it is not subjectively meaningful and is often found to correlate poorly with subjective ratings (Pratt, 1979) and to lead to quite a large spread in reported subjective quality for images having similar measured distortions (Limb, 1979).

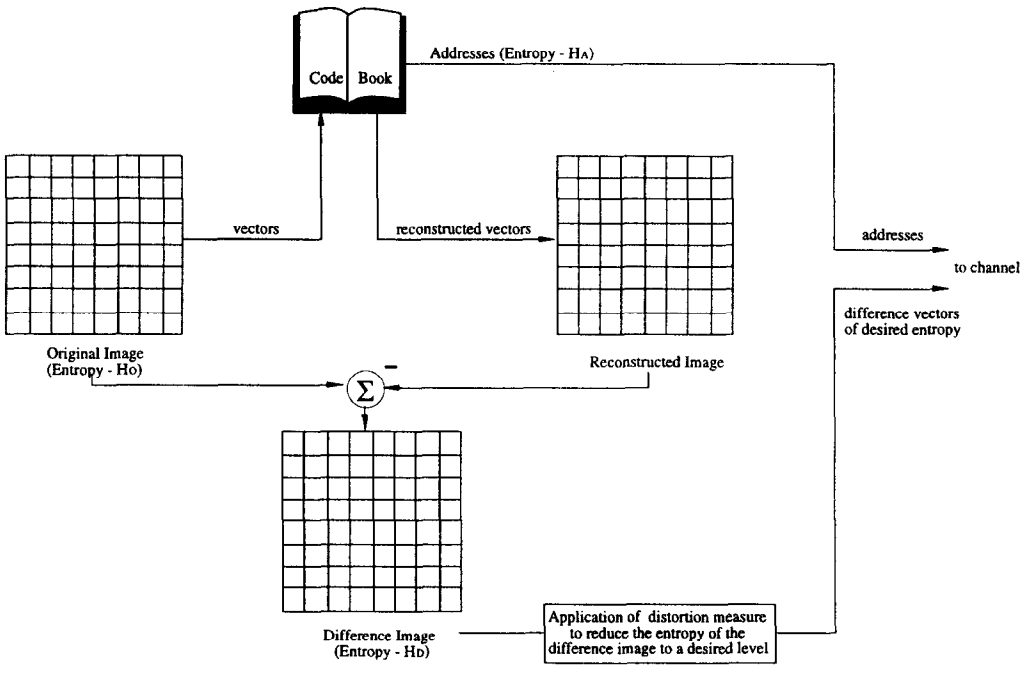


Figure 1. A schematic illustration of lossy coding by vector quantization.

A common fault with MSE distortion measures is that a small average distortion can result in a strong visual artefact if all the error is concentrated in a small, but important region. Additionally, a perfect reconstruction with a 1-pixel shift will lead to a high MSE, which may be interpreted incorrectly as a poor reproduction.

A DISTORTION MEASURE FOR HYPERSPECTRAL IMAGES

In the compression of hyperspectral data, it is particularly important for a suitable distortion measure to be determined in the context of the application of the data since, in general, the receiver of the encoded images will not be a human viewer but will be some form of automated image processing system. To model the human visual system is therefore not relevant, and most of the rationale behind video coding distortion measures does not apply. Additionally, while video and image coding generally have the human viewer as their sole user, hyperspectral data is manipulated by a range of remote sensing applications based on automated algorithms.

The major difficulty is that images that have been compressed in a lossy way using a particular distortion measure can be used accurately only with those applications for which that distortion measure has been developed. In other words, the information that can be discarded for one particular application may be necessary for use in another. Since the distortion measure is used to decide which information can be discarded, a general distortion measure is required that can be applied, even though the application for the coded hyperspectral image is not known in advance.

Therefore, for hyperspectral images, it is also more appropriate to minimize a distortion measure that reflects the loss in scientific value of the original data. Unfortunately, such quantitative measures do not exist, either for specific applications or for general application to hyperspectral data.

Before discussing the difficulties associated with the development of suitable quantitative measures, it is instructive to develop appropriate criteria for a suitable distortion measure for the lossy compression of remote sensing images. Table 1 lists appropriate criteria developed in this work and compares them to those relevant to a distortion measure for video and speech coding, which are well known (Gray et al., 1980).

The first and second criteria are the same for the coding of both video and remote sensing data. The third and fourth criteria represent a significant difference and are discussed further in the following sections.

Quantitatively Meaningful

The requirement for a distortion measure to be *quantitatively meaningful* for remote sensing images is significantly different from the *subjectively meaningful* requirement for video coding, principally because of the "use" of the data. Invariably, in video coding the end user is the viewer whose eye is a very forgiving device due to its substantial ability to integrate spatial information. Large errors are therefore less important, provided that they are small in number and are surrounded by pixels with

Table 1. Criteria for a Distortion Measure for Video and Speech Coding (Gray et al., 1980) and for Remote Sensing Coding

Video Coding	*Remote Sensing Coding*
Mathematically tractable to permit analysis	Mathematically tractable to permit analysis
Computable in real time	Computable in real time (although not as essential as in video coding)
Subjectively meaningful so that large or small errors correspond to bad or good image quality	Quantitatively meaningful in that differences in the coded image correspond to the same relative magnitude of differences in the original image
	Independent of the application to be applied to the coded data

errors of small magnitude. In remote sensing applications, the user is invariably an algorithm, which does not necessarily have the same ability to ignore isolated large errors. An example is the matching of library spectra, where the differences in the coded data must correspond to the same relative magnitude of differences in the original image; otherwise the input vectors cannot be correctly matched. Another way to state the requirement to be quantitatively meaningful is that the measure must, in a quantitative way, reflect the loss in scientific value of the original data.

Application Independence

Again, the requirement for the measure to be independent of the application is significantly different for remotely sensed data. The end result of video coding is generally used by one application—the human eye. The end result of remote sensing coding is subjected to one or more of a wide range of remote sensing algorithms. Examples include classification, feature extraction, linear transformation, library spectra matching, and so on. Each of these applications may require a different degree of accuracy. If results are to be listed in a thematic map, accuracy of classification is of great importance. If image spectra are to be matched with library spectra, accuracy of reconstruction is very important. The development of a general distortion measure must accommodate these diverse requirements.

Development of a Distortion Measure with a Theoretical Base

Ideally, as in image coding, it would be possible to develop a distortion measure using a unified model that adequately describes each application. For example, what is required for maximum likelihood classification is a general distortion measure that can accurately predict the amount of distortion allowable in each decision that still allows the discriminant functions to remain valid and be able to separate classes in the same manner as it would on the original data.

While it may be possible to develop such a distortion measure, it is unlikely to be suited to any other remote sensing algorithm since it would have been derived specifically for maximum likelihood classification. That is, the distortion measure would reflect the scientific value of data relevant to maximum likelihood classification, and not necessarily any other remote sensing application. The problem of resolving a general distortion measure with application to all remote sensing algorithms is a complex one which is likely to prove as difficult to achieve as a general model of human vision has been in video coding.

Even if such a general distortion measure was developed, it is not likely to be able to be implemented on a pixel-by-pixel basis as the image is compressed. This is a very real issue as the encoder must be able to decide for each pixel (or each vector in the case of vector quantization) the allowable level of distortion to maintain a desired level of accuracy in all remote sensing applications. It is therefore desirable to examine the available simple distortion measures to determine if any appear suited to application to hyperspectral data.

A Suitable Distortion Measure

A particularly relevant requirement of a suitable distortion measure is implicit in the criteria developed earlier. The distortion measure must produce predictable results in the performance of each remote sensing application for a given level of distortion. The application scientist must be able to apply a particular algorithm to the compressed data and, by determining the algorithm's performance in the presence of the given level of distortion, must be able to infer the performance of the algorithm on the original (undistorted) data.

This requirement is very important. The compressed data is of little use to the application scientist unless the performance of an algorithm can be observed on the distorted data and an accurate assessment made of the performance of the algorithm, had it been applied to the original data. This requirement also implies that the algorithm's performance is consistently predictable for a range of levels of distortion. Figure 2a illustrates the concept in an ideal way for a classification application, which evaluates classification accuracy for various numbers of bands/features. Figure 2b provides another view of the

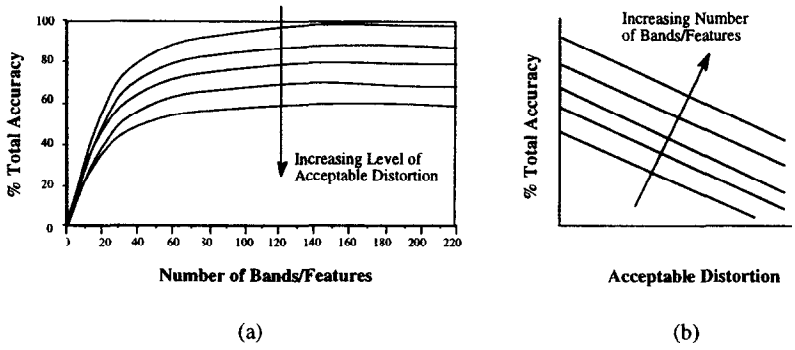


Figure 2. Ideal performance of a classifier in the presence of increasing levels of acceptable distortion as a function of a) the number of bands/features and b) the level of acceptable distortion.

requirement and illustrates the ideal variation in accuracy for a given number of bands/features in the presence of increasing levels of acceptable distortion.

These responses of a classifier in the presence of differing levels of distortion are ideal because they provide a degree of predictability to the application scientist who, having achieved a certain accuracy at one distortion level, could use a graph similar to Figure 2b to infer something about the accuracy to be expected from classification on the original image (or at some other distortion level).

Image providers could determine the performance of their data sets after lossy compression using an appropriate distortion measure and develop such graphs. In addition to a catalogue of available data sets, the service providers could also present a graph similar to that in Figure 2, which advises the application scientist of the classification performance of that data set in the presence of varying levels of distortion (or, in other words, compressed to varying degrees). The application scientist could then purchase compressed data for initial investigations, confident that the performance of a particular algorithm on the lossy data would produce a reasonably predictable performance on the original image. Having conducted initial investigations on the cheaper data with reduced transmission and storage requirements, small lossless images of the particular area of interest could then be procured to conduct detailed investigations.

Alternatively, having procured a full data set, the application scientist could use an accompanying accuracy/distortion graph to compress the image to reduce storage requirements. The full data set could then be used when investigation areas and parameters have been refined.

The Suitability of Mean Distortion Measures

The difficulty with *mean distortion* measures is that they cannot be obtained directly since a decision on a mean distortion level cannot be made on a pixel-by-pixel basis. To accommodate this in video coding, it is normal to encode using scalar quantization on a pixel-by-pixel basis, which results in a desired mean distortion across the image. The other difficulty is that the specification of a mean level of distortion tells the user something about the image in a general sense but does not really assist in the application of remote sensing algorithms, since it does not give any indication of the specific magnitude of the errors in any particular spatial or spectral portion of the image.

Unfortunately, in almost all remote sensing applications, algorithms make local decisions whether it is in spectral matching or the application of discriminant functions. Information as to the specific location and magnitudes of errors can only really be provided by an absolute distortion measure that indicates something quantitative about both the local and global distortion in an image.

The Suitability of Absolute Distortion Measures

Absolute distortion measures not only give precise information about local distortion, but they have the added advantage of being able to be applied on a pixel-by-pixel basis as the image is being encoded. Even so, the absolute distortion measure must be quoted as a maximum error rather than a mean error, for the reasons discussed earlier. Additionally, since hyperspectral images have a wide dynamic range, a suitable distortion measure must be relative to the original pixel intensity. The best way of achieving this is to use a *percentage maximum absolute distortion (PMAD)* measure.

The PMAD measure guarantees that every pixel $\hat{B}(x,y,\lambda)$ in the reconstructed image is within a maximum distance of $p \cdot 100\%$ to its original value $B(x,y,\lambda)$. That is,

$$(1-p)\cdot B(x,y,\lambda) < \hat{B}(x,y,\lambda) < (1+p)\cdot B(x,y,\lambda). \quad (1)$$

This PMAD distortion measure appears to be the most suitable for application to the lossy compression of hyperspectral images. It gives a direct indication of the loss of scientific value for a particular level of distortion and is able to be implemented on a pixel-by-pixel basis. It also offers the best potential to produce predictable results in a range of remote sensing applications.

The evaluation of lossy compression techniques requires an understanding of the applications that will be applied to the reconstructed data. However, the analysis of hyperspectral data is still not completely understood with new applications continually being developed. Therefore, in this work it is proposed to narrow the scope of the

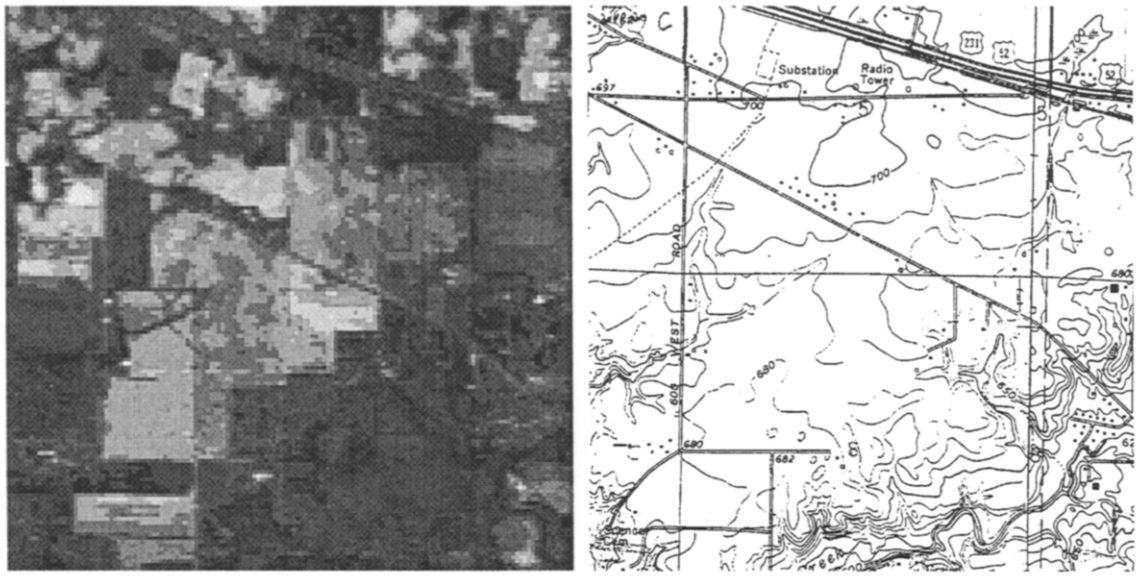


Figure 3. The test area illustrated by a) a grey-scale image of the data set and b) a USGS Quadrangle map (after Landgrebe, 1994).

analysis to the problem of classification using *maximum likelihood classification (MLC)* (Richards, 1993) on original pixels and after feature extraction by the *decision boundary feature extraction (DBFE)* technique (Lee and Landgrebe, 1993).

THE DATA SET

This work makes use of a data set provided with the Multispec© program provided by the Purdue Research Foundation (Landgrebe and Biehl, 1995). This has two advantages: Ground truth data is available for the data set, and benchmark results have already been presented in work published to support Multispec© (Landgrebe, 1994). The selected image is from a June 1992 AVIRIS data set of a mixed agriculture/forestry area in the Indian Pine Test Site, a 100 square mile area in northwestern Indiana. The data set is designated 92AV3C. It is 9.04 Mbytes in size and contains 145 lines with 145 pixels in each of 220 bands. Figure 3a shows an image of the area.

The features and the terrain relief of the area are shown more clearly on the USGS quadrangle map of the area (Fig. 3b). Ground-truth data for the data set has been provided (Landgrebe, 1994) to produce training samples for the ground cover of the area. The types of ground cover are illustrated in the generalized reconnaissance map shown in Figure 4.

This map is highly generalized since many of the small variations within fields that are obvious in the image data are ignored in the reconnaissance map. Therefore, not every pixel inside a designated area on the reconnaissance map should be expected to belong to the class associated with that area, and a classifier should not be expected to classify accurately all pixels as belonging to the correct class since the area designated as belonging to a particular class may in fact contain small numbers of pixels from other classes.

Additionally, the map gives land-use classes, rather than land-cover classes. An area marked *Corn-notill* on the map may really have a land cover of bare soil and residue from previous vegetation, with only a small percentage of corn vegetation as land cover. This is an important issue because it means that the data set contains some particularly challenging classes (Landgrebe, 1994), since the data were collected in the early part of the

Figure 4. A generalized reconnaissance map of the area of 92AV3C (after Landgrebe, 1994).

growing season when the canopies for corn and soybean presented only about 5% ground cover. The spectral response was therefore dominated, less by the type of crop, and more by the type of soil in the field and the amount of residue from last year's crop. To account for this fact, the corn and soybean classes were divided further into three subclasses depending on the amount of tillage employed in each field: *no-till*, *minimum-till*, and *clean*. The no-till fields have a substantial amount of crop residue from last year, the minimum-till fields have a moderate amount, and the clean fields have no residue and would therefore show just the crop canopy against the soil background.

Six classes were therefore formed from the two crops in three conditions. The separation of these six classes represents a very challenging classification problem (Landgrebe, 1994) due to the very poor signal-to-noise ratio, where the spectral response from the crop leaves represents the signal and the noise is represented by the soil and residue content of the background.

MAXIMUM LIKELIHOOD CLASSIFICATION (MLC)

This section investigates the maximum likelihood classification (MLC) technique using the Multispec© application. There are a number of issues associated with maximum likelihood classification. namely, the number of training pixels required for each class, the method for assessment of total accuracy, the generalization of results, and the selection of bands for inclusion in classification.

Number of Training Pixels for each Class

Reasonable estimates of the mean and covariance matrix for each class require that sufficient training pixels are available for each class. At least $N+1$ training pixels are required for an N-dimensional decision space to avoid producing a covariance matrix that is singular. A singular covariance matrix cannot be inverted, preventing the MLC technique from determining discriminant functions. (It should be noted that. even though having $N+1$ training pixels guarantees nonsingularity of the covariance matrix, it does not guarantee that that number of pixels is necessarily sufficiently representative of the classes.) Therefore, any class with less than $N+1$ training pixels is removed from the training set prior to application of the MLC technique. Since the smallest class is Oats with 20 members, the smallest number of bands that can be used for all classes is 19.

Table 2 illustrates the classification accuracy for the entire 220 original bands. In this case, any class with less than 221 pixels must be ignored (viz., Alfalfa, Grass/Pasture-mowed, Oats, Wheat, and Stone-steeltowers).

Assessment of Total Accuracy

Since the classifier is tasked to classify the training set, 100% accuracy is to be expected if all 220 bands are used. The actual accuracy is 96%. This occurs for two reasons. First, as noted earlier, the reconnaissance map for the area is highly generalized and not every pixel inside a designated area can be expected to belong to the class associated with that area. Some pixels will therefore be incorrectly classified according to the generalized reconnaissance map and the classifier will appear not to be working to its full potential. The second reason for the difference is related to the method used to calculate accuracy.

Table 2 illustrates how the total accuracy of the MLC classifier is assessed by the Multispec© application. As shown in Table 2, when training sizes for classes are too small, that class is ignored. The classifier is recorded to have 0% accuracy for that class. This is a little unfair as the classifier did not attempt to classify that class. The accuracy of the classifier on those classes it classified is therefore better than the total accuracy listed.

Table 2 lists the accuracy of the MLC classifier of 220 bands as classifying 9952 of the 10,366 pixels correctly giving an accuracy of 96%. If the bands with less than 221 pixels are not included in the assessment, the accuracy is 9952 correctly classified of the total of 9959 pixels submitted to the classifier. This performance of 99.9% is closer to that expected if the full 220 bands are included. Despite this, the inclusion of the bands with less than $N+1$ pixels is appropriate in the assessment of accuracy as it more correctly reflects the ability of the classifier to classify the image. It is a little unfair to include the classes that the classifier cannot even attempt to classify, but it is the most appropriate way to compare the classifier's performance using different numbers of bands.

Generalization of Results

The data set chosen for this work is too small to be of any practical significance in terms of classification of the data, but was chosen because a larger image was not available with ground-truth data. A particular limitation caused by the small image size is the need to have large training sets for the classification of 220 bands so that all pixels for which ground-truth data is available had to be included in the training sets. Since such a large proportion of the image had to be included in the training sets, classification could only be conducted on the training data, and it is not possible to generalize the performance of the classifiers beyond their training sets. This is limiting in a classification sense, but does not detract from this work which concentrates on the *relative* performance of a classifier in the presence of distortion introduced by degrees of lossy compression.

Table 2. Multispec© Output-Classification Accuracy of MLC of All 220 Bands

Class Name	*Class No.*	*Percent Correct*	*Number Samples*	*2*	*3*	*4*	*5*	*6*	*8*	*10*	*11*	*12*	*14*	*15*
Alfalfa	1	0.0	54	1	0	0	0	1	47	0	5	0	0	0
Corn-notill	2	99.8	1434	1431	0	0	0	2	0	0	1	0	0	0
Corn-min	3	100.0	834	0	834	0	0	0	0	0	0	0	0	0
Corn	4	100.0	234	0	0	234	0	0	0	0	0	0	0	0
Grass/Pasture	5	100.0	497	0	0	0	497	0	0	0	0	0	0	0
Grass/Trees	6	100.0	747	0	0	0	0	747	0	0	0	0	0	0
Grass/Pasture-mowed	7	0.0	26	0	0	0	0	0	17	0	9	0	0	0
Hay-windrowed	8	100.0	489	0	0	0	0	0	489	0	0	0	0	0
Oats	9	0.0	20	15	0	0	0	0	0	0	5	0	0	0
Soybeans-notill	10	100.0	968	0	0	0	0	0	0	968	0	0	0	0
Soybeans-min	11	99.8	2468	0	0	0	0	3	0	1	2464	0	0	0
Soybeans-clean	12	100.0	614	0	0	0	0	0	0	0	0	614	0	0
Wheat	13	0.0	212	0	0	0	8	85	0	0	2	0	117	0
Woods	14	100.0	1294	0	0	0	0	0	0	0	0	0	1294	0
Bldg-Grass-Tree	15	100.0	380	0	0	0	0	0	0	0	0	0	0	380
Stone-steeltowers	16	0.0	95	28	0	0	0	0	0	0	8	58	0	1
Total			10,366	1475	834	234	505	838	553	969	2494	672	1411	381
Overall performance (9952/10,366)=96.0%														

Table 3. Multispec© Output-Classification Accuracy of MLC of 50 DBFE Features

Class Name	*Class No.*	*% Correct*	*Number Samples*	*1*	*2*	*3*	*4*	*5*	*6*	*8*	*10*	*11*	*12*	*13*	*14*	*15*	*16*
Alfalfa	1	100.0	54	54	0	0	0	0	0	0	0	0	0	0	0	0	0
Corn-notill	2	95.0	1434	0	1363	10	1	2	2	0	16	37	2	0	0	1	0
Corn-min	3	97.6	834	0	6	814	0	0	0	0	1	12	1	0	0	0	0
Corn	4	100.0	234	0	0	0	234	0	0	0	0	0	0	0	0	0	0
Grass/Pasture	5	99.0	497	0	0	0	0	492	0	0	0	4	1	0	0	0	0
Grass/Trees	6	99.9	747	0	0	0	0	0	746	0	0	0	0	0	0	1	0
Grass/Pasture-mowed	7	0.0	26	0	0	0	0	1	1	23	0	1	0	0	0	0	0
Hay-windrowed	8	100.0	489	0	0	0	0	0	0	489	0	0	0	0	0	0	0
Oats	9	0.0	20	0	5	0	0	9	3	0	0	1	0	0	0	2	0
Soybeans-notill	10	97.2	968	0	3	1	0	2	2	0	941	19	0	0	0	0	0
Soybeans-min	11	89.1	2468	0	100	36	0	12	6	0	80	2198	34	0	0	2	0
Soybeans-clean	12	98.5	614	0	0	3	0	0	0	0	0	5	605	0	0	0	1
Wheat	13	99.5	212	0	0	0	0	0	0	0	0	0	0	211	1	0	0
Woods	14	98.5	1294	0	0	0	0	1	2	0	0	0	0	0	1274	17	0
Bldg-Grass-Tree	15	96.3	380	0	0	0	0	0	0	0	0	0	0	0	14	366	0
Stone-steeltowers	16	100.0	95	0	0	0	0	0	0	0	0	0	0	0	0	0	95
Total			10,366	54	1477	864	235	519	762	512	1038	2277	643	211	1289	389	96
Overall performance (9882/10,366)=95.3%																	

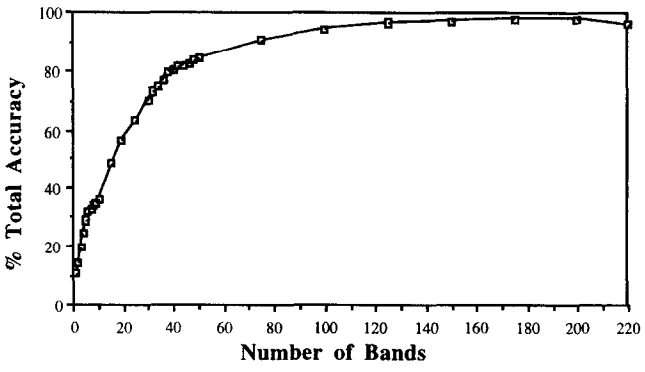


Figure 5. Percentage total accuracy of classification for maximum likelihood classification of various numbers of original bands of 92AV3C.

Selection of Bands for Inclusion in MLC Classification

When classification is conducted using a subset of the original bands as the decision space, it is important to identify which bands have been selected. For example, the performance of the classifier using a subset of four bands will depend on which four bands are selected. In this work, bands are selected sequentially. Figure 5 illustrates the total classification accuracy as the number of bands included in the decision space is increased. The number of bands is selected by including additional bands in recorded order. Thus the figure does not represent the best performance available for the selection of a particular number of bands, but represents performance for a particular sequence of including bands (that is, in the order in which they were recorded), which may not be the optimum but is adequate for the purposes of comparison.

As expected, the total accuracy of the classifier in Figure 5 increases as more bands are included. This will only occur when the classifier is tasked to classify the data on which it was trained, that is on the training set. It should be noted that for classification of large test areas, where data other than the training set is to be classified, the increase in accuracy will improve up to a certain point as additional features are added, and then deteriorate as more features are added. This is the Hughes effect (Hughes, 1968), which arises because of the finite number of training samples available and the limited precision to which the class distributions can be estimated.

The effect of including ignored classes in the assessment of total accuracy can be seen in the dip at the end of the curve in Figure 5. The class Wheat (212 pixels) is included in the evaluation until the number of bands reaches 212. When 212 or more bands are included, Wheat must be ignored and the sudden "misclassification" of 212 pixels is not offset by the increase in the performance of the classifier on other classes. Consequently, the classifier suddenly appears to perform less well.

DECISION BOUNDARY FEATURE EXTRACTION (DBFE)

The difficulty in selecting bands and the large number of bands available mean that most classification is conducted after some form of feature extraction technique has been applied to the data. This has the advantage of reducing the decision space to the desired size while including the most discriminating features, in order of importance. This section examines the performance of an MLC classifier after the *decision boundary feature extraction (DBFE)* (Lee and Landgrebe, 1993) technique is applied to the original 92AV3C image.

Number of Training Pixels for each Class

To again avoid singularity problems with matrix inversion, the use of the DBFE feature extractor again requires the selection of a subset of classes which excludes any class that has less than $N+1$ (221) training samples. The classes Alfalfa, Grass/Pasture-mowed, Oats, Wheat, and Stone-steeltowers are therefore omitted from the classes used for feature extraction, which then only utilizes 11 of the available 16 classes.

It is possible for the feature extractor to produce up to N (220) features from the N bands of original data. However, once again, the number of features used for the MLC classification will dictate the size of classes that can be presented to the classifier without causing singularity problems in matrix inversion. As before, 19 features is the largest number that can be used and still include all classes in the classification, since Oats has only 20 pixels.

Table 3 illustrates the Multispec© output for the MLC classification of 50 DBFE features. Note that, for reasons discussed earlier, the Grass/Pasture-mowed class and the Oats class are ignored in the classification since they have less than 50 pixels in the class.

A maximum of 50 features was used in this investigation representing a trade-off between demonstrating a sufficiently large range of DBFE performance and the time taken to perform feature extraction and classification. Approximately 20 h was required to generate DBFE features and another 12 h to generate enough points to plot a single curve such as those shown in Figures 5 and 6. Additional features will provide a gradual increase in performance, but since this investigation is in comparative performance, 50 features were deemed sufficient.

Figure 6 presents results in a similar format to that of Figure 5. Percentage total accuracy is plotted as a function of the number of DBFE features. Again, as the number of DBFE features is increased, the classifier performance improves. As the number of features is increased above 50, the performance can be expected to increase fairly linearly to provide a total accuracy close to 100% if all 220 features were included.

The curves described in Figures 5 and 6 provide a

benchmark for the performance of MLC and MLC after DBFE. In the following sections the performance of MLC classification is investigated (on the original bands and on DBFE features) in the presence of distortion added due to the lossy compression of 92AV3C by mean-normalized vector quantization.

THE EFFECT OF LOSSY COMPRESSION ON THE CLASSIFICATION OF HYPERSPECTRAL IMAGES

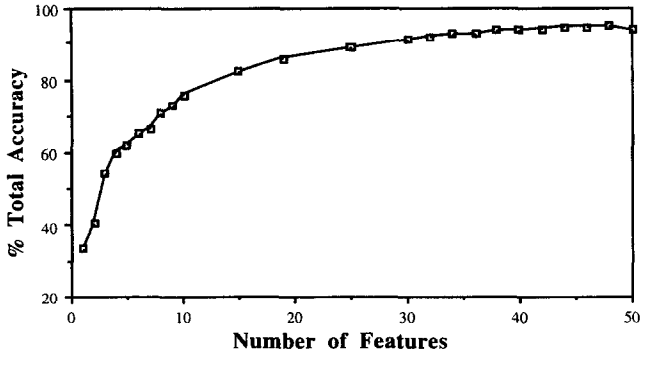
Mean-Normalized Vector Quantization (M-NVQ)

The original image 92AV3C was vector quantized using the M-NVQ technique, and then reconstructed to varying degrees of accuracy using the PMAD distortion measure. Vector quantization was achieved by clustering 1000 reference vectors to form a codebook of 32 codevectors. The length of vector was chosen to be 22, giving 10 blocks of vectors in the 220-band image. This was not necessarily the optimum length, but was deemed adequate as it was computationally much more straightforward to have the same number of vectors in each band block to be coded.

Compressed images were obtained by performing M-NVQ and then determining, on a pixel-by-pixel basis, how close the reconstructed pixel value was to the original value. If greater than the allowed distortion level, an error (residual) vector was included to ensure that the reconstructed pixel was within the allowed distortion. In this way, every pixel in the compressed image could be guaranteed to be within the allowed degree of PMAD distortion when reconstructed.

The full residual vectors are stored with the appropriate means and codebook addressed in the compressed image. On reconstruction, the address is used to select the appropriate codevector from the codebook, the vector mean is used to reverse the normalization process, and the reconstructed vector is examined to determine whether each pixel is within the allowed distortion level.

Figure 6. Percentage total accuracy of classification for maximum likelihood classification of up to 50 DBFE features for 92AV3C.



If the reconstructed pixel is between $(1\pm p)x_i$, no residual is stored. If the pixel is reconstructed outside $(1\pm p)x_i$, a residual is added to take the reconstructed vector to the appropriate error boundary, and the residual is stored in the difference image. Each pixel of the reconstructed image is then guaranteed to be within $\pm p\%$ of its original value.

More efficient methods are possible that have the potential to lead to lower difference image entropies. For example, the residual could be scalar quantized so that, while the residual can be guaranteed to be within the allowable distortion, lower difference image entropies will result. However, in this work the full residual was stored in the difference image to provide a benchmark for later investigations using scalar quantization of the residuals.

Properties of the Compressed Images

The compressed image contains residual vectors, vector means, and codebook addresses so that the entropy of the compressed image is a result of the combination of the residual vector entropy (H_R), the entropy of the vector means (H_μ), and the entropy of the codebook addresses (H_A). Table 4 gives these entropies and the total normalized compressed entropies (H_C) for the 92AV3C image compressed to varying degrees of PMAD distortion. The normalized compressed image entropy (measured in bits per original pixel) is employed so that the compressed image can be compared directly to the original image storage requirement (H_O), which has the same units. That is,

$$H_C = \frac{\text{no. of residual bits} + \text{no. of address bits} + \text{no. of parameter bits}}{\text{no. of pixels in total image}}. \quad (2)$$

Note, therefore, that $H_C \neq H_R + H_A + H_\mu$.

Another way to view these results is from the point of view of the application scientist who is interested in the relative classification accuracy of images compressed to varying degrees. The relationship between PMAD and the total compressed entropy is given in Table 4. Using these values and an original image storage requirement of 16 bits per pixel the compression ratio can be calculated. The compression ratio C_R is the ratio of the original image storage requirement H_O (16 bits/pixel) to the compressed entropy H_C:

$$C_R = \frac{H_O}{H_C}. \quad (3)$$

The values are plotted in Figure 7 to illustrate the variation of each of the entropies with PMAD. Figure 7 shows the complete picture of the entropies of the compressed images. Since the address and the mean entropies are the same regardless of the level of PMAD (0.19 bits/pixel and 0.35 bits/pixel, respectively), the significant

Table 4. Residual, Address, Mean, Total Compressed Entropies, and Compression Ratios for M-NVQ Compression of 92AV3C to Varying Levels of PMAD

	PMAD (%)													
	0	*0.25*	*0.5*	*0.75*	*0.825*	*1*	*1.5*	*2*	*2.5*	*3*	*4*	*5*	*7.5*	*10*
H_R	5.98	4.82	3.56	2.49	2.26	1.77	0.92	0.50	0.28	0.17	0.07	0.03	0.01	0.00
H_A	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19
H_μ	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35	0.35
H_C	6.65	5.46	4.22	3.19	2.96	2.49	1.70	1.31	1.11	1.01	0.93	0.90	0.88	0.88
C_R	**2.41**	**2.93**	**3.79**	**5.02**	**5.41**	**6.43**	**9.41**	**12.21**	**14.41**	**15.84**	**17.02**	**17.78**	**18.18**	**18.18**

factor is the variation of residual entropy (H_R) with PMAD. That is, as the level of PMAD is allowed to increase, there is a corresponding decrease in the additional information required to be added to the compressed image to bring it within the required distortion level.

It should be noted that, once a level of approximately 4–5% PMAD is reached, little additional information is required to be added. In other words, the reconstruction of the image from codevectors (based on the transmitted means and addresses) is already within the allowable level of distortion. This illustrates the effectiveness of mean-normalized vector quantization, which, for the 92AV3C image, produces a reconstructed image from the codebook, which has each pixel within 4–5% of its original value without having to transmit any significant additional information. M-NVQ therefore provides a powerful technique for the lossy compression of hyperspectral images.

EFFECT OF COMPRESSION ON MLC OF ORIGINAL PIXEL VALUES

PMAD Performance during MLC Classification

The original image 92AV3C was compressed to varying degrees of PMAD distortion, and each reconstructed image was classified with an MLC classifier using a varying number of bands in the classification process. Figure 8 illustrates the result of the classification.

As the level of distortion is allowed to increase, the classification accuracy decreases until a distortion level of approximately 4% is reached when the accuracy begins to increase again. This is difficult to see in Figure 8, but is more easily seen in Figure 9, which plots the percentage total accuracy versus PMAD for a number of bands. This is in effect a series of vertical slices of Figure 8 taken at 80, 100, 130, 150, and 220 bands.

The response to PMAD distortion illustrated in Figures 8 and 9 is not in accordance with the ideal response of Figures 2a and 2b. While the accuracy decreases with increased distortion until approximately 4%, it begins to increase again after that level.

The reason for this increase in accuracy is due to the Corn and Soybean subclasses, which, as noted earlier, are particularly difficult to discriminate. The spectral response was therefore dominated, less by the type of crop, and more by the type of soil in the field and the amount of residue from last year's crop. To account for this fact, the Corn and Soybean classes were divided further into three subclasses depending on the amount of tillage employed in each field: *no-till*, *minimum-till*, and *clean*. The no-till fields have a substantial amount of crop residue from last year, the minimum-till field has a moderate amount, and the clean field has no residue and

Figure 7. Total, residual, address and mean entropies required to ensure that M-NVQ of 92AV3C is within varying levels of PMAD.

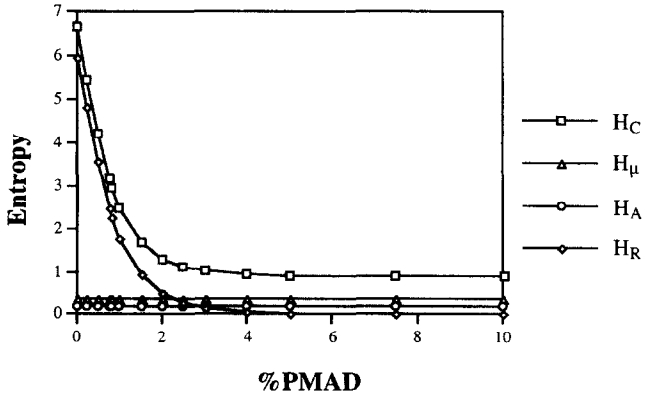
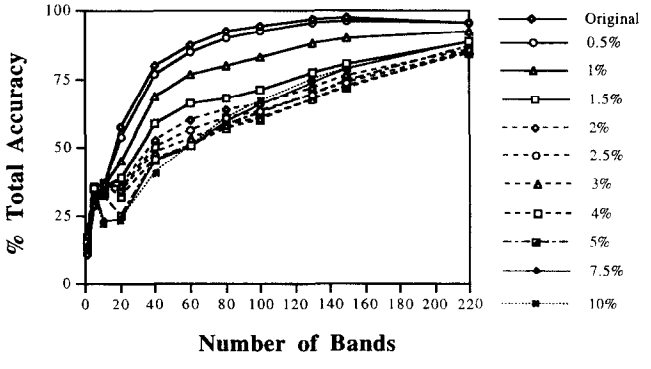


Figure 8. Percentage total accuracy of classification for maximum likelihood classification of various numbers of original bands of 92AV3C in the presence of varying levels of PMAD distortion.



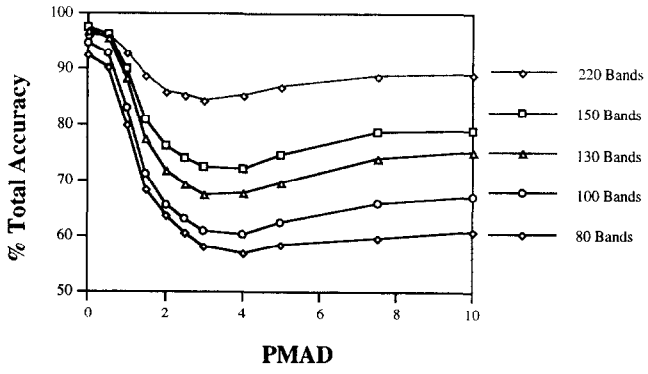


Figure 9. Percentage total accuracy of MLC classification versus PMAD for a selection of bands.

would therefore show just the crop canopy against the soil background.

Six classes were therefore formed from the two crops in three conditions. The separation of these six classes represents a very challenging classification problem due to the very poor signal-to-noise ratio, where the spectral response from the crop leaves represent the signal and the noise is represented by the soil and residue content of the background.

Table 5 illustrates this further with the example of classification accuracy for MLC of 220 bands with 4% distortion. The major errors in classification in the classes of Corn and Soybeans, which are confused with each other and with other types of the same class group. In particular, the Soybean-min class is confused with other Soybean classes and Corn classes. The small amount of foliage (5%) in each of these classes is the major cause of the confusion so that as the level of distortion varies, the background is classified rather than the foliage. This is most likely due to improvements in class normality, which means that classes begin to look more like each other in the presence of certain levels of distortion.

Since this "misclassification" also occurs with the MLC of DBFE features, further detailed investigation and a pictorial representation of the effect are given in the next section.

EFFECT OF COMPRESSION ON MLC OF DBFE FEATURES

MLC classification is rarely employed on original bands. Normally, some form of feature extraction is performed first. The feature extraction technique chosen for this work is the DBFE method. This section conducts a similar investigation to that of the previous section and considers the effect of the PMAD distortion measure on the accuracy of MLC classification of DBFE features. The image used is again the 92AV3C image.

The original image 92AV3C was again compressed to varying degrees of PMAD distortion. DBFE features were then extracted from each reconstructed image. The

Table 5. Multispec© Output-Classification Accuracy of MLC of First 220 Bands of 92AV3C with a PMAD Distortion Level of 4%

Class Name	*Class No.*	*Percent Correct*	*Number Samples*	*2*	*3*	*4*	*5*	*6*	*8*	*10*	*11*	*12*	*14*	*15*
Alfalfa	1	0.0	54	0	0	0	2	2	34	3	0	1	0	12
Corn-notill	2	92.1	1434	**1320**	**14**	**0**	0	2	0	**82**	**10**	**6**	0	0
Corn-min	3	97.7	834	**6**	**815**	**0**	0	0	0	**3**	**4**	**6**	0	0
Corn	4	100.0	234	**0**	**0**	**234**	0	0	0	**0**	**0**	**0**	0	0
Grass/Pasture	5	96.8	497	0	0	0	481	0	0	2	0	2	12	0
Grass/Trees	6	99.7	747	0	0	0	0	745	0	1	0	0	1	0
Grass/Pasture-mowed	7	0.0	26	0	0	0	4	0	21	0	1	0	0	0
Hay-windrowed	8	100.0	489	0	0	0	0	0	489	0	0	0	0	0
Oats	9	0.0	20	0	0	0	0	18	0	0	0	0	0	2
Soybeans-notill	10	98.1	968	**10**	**0**	**0**	0	1	0	**950**	**6**	**1**	0	0
Soybeans-min	11	67.1	2468	**204**	**87**	**0**	2	15	1	**470**	**1655**	**27**	0	7
Soybeans-clean	12	95.1	614	**8**	**7**	**0**	0	0	0	**11**	**4**	**584**	0	0
Wheat	13	0.0	212	0	0	0	0	2	0	0	0	0	0	210
Woods	14	98.3	1294	0	0	0	0	3	0	0	0	0	1272	19
Bldg-Grass-Tree	15	75.8	380	0	0	0	1	5	0	1	0	0	85	288
Stone-steeltowers	16	0.0	95	0	3	0	0	2	0	1	0	1	0	88
Total			10,366	1548	926	234	490	795	545	1524	1680	628	1370	626

Overall performance (8833/10,366) = 85.2%

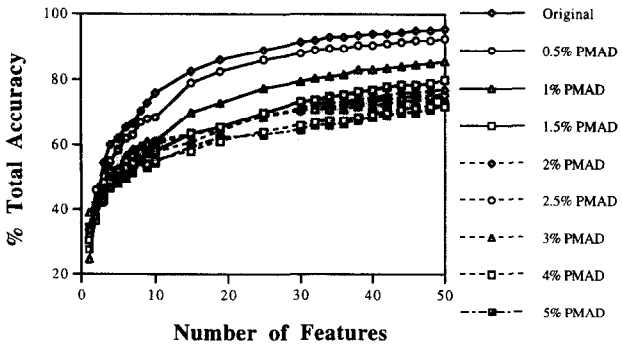


Figure 10. Percentage total accuracy of classification for maximum likelihood classification of up to 50 DBFE features for 92AV3C in the presence of varying levels of PMAD distortion.

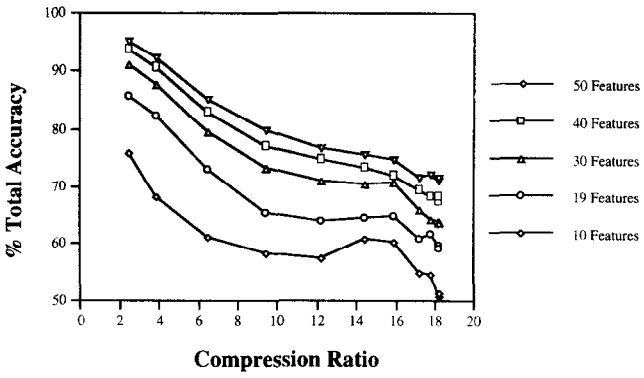


Figure 12. Percentage total accuracy of classification for maximum likelihood classification of a range of DBFE features for 92AV3C as a function of varying degrees of compression.

transformed image was then classified with an MLC classifier using a varying number of features in the classification process. Figure 10 illustrates the results of the classification investigation.

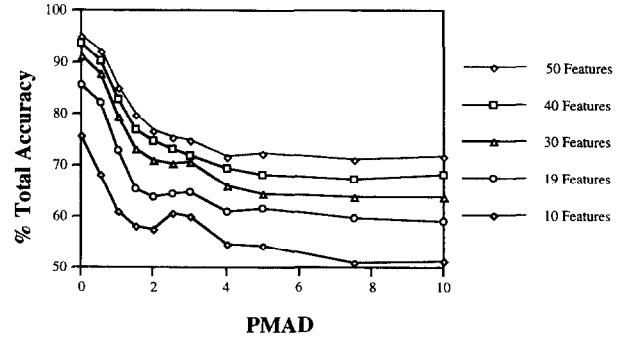
The performance of MLC classification after DBFE feature extraction is closer to the ideal performance illustrated by Figure 2. As the level of distortion is increased, the classifier performance decreases. As desired, it decreases in what appears to be a regular manner commensurate with the increase in acceptable distortion.

The performance for a range of numbers of features is better shown in Figure 11, which plots percentage total accuracy versus PMAD for a selection of DBFE features. Again, the performance is close to the predictability desired for classification of a distorted image. The family of curves has a regular shape, and, once this shape is known, it appears that the performance of the classifier can be predicted for a given level of distortion and a given number of features.

Figure 12 shows the variation of total classification accuracy as a function of the compression ratio (which is still, in effect, a surrogate measure for PMAD due to the relationship shown in Table 4).

As required by the ideal performance criterion of Figure 2b, the relationship is regular, allowing the application scientist to be confident of the accuracy that could be expected to be achieved using a required degree of compression. In particular, should a certain accuracy be gained for a particular degree of compression (that is, a particular level of PMAD distortion), the performance of the classifier on the original image (with zero distortion) can be inferred from the information provided in Figure 12.

Figure 11. Percentage total accuracy of classification for maximum likelihood classification of a range of DBFE features for 92AV3C in the presence of varying levels of PMAD distortion.



These results are further visually illustrated by examining the thematic maps for MLC classification of 50 DBFE features for the 92AV3C image reconstructed after compression by M-NVQ to varying degrees of PMAD. These maps are presented in Figure 13.

The effect of the distortion is perhaps best highlighted in the comparison between the thematic maps for the original image and the image compressed to 4% PMAD. The two thematic maps have a number of differences; but the map from the image compressed to 4% PMAD is still a very good representation of the original image, and yet is obtained from an image that is only 1/17 the size of the original. In particular, Figure 13 illustrates the difficulty in distinguishing Corn classes from Soybean classes, in both the original and distorted images.

It should be noted that there is considerable amount of "salt and pepper" error in the images of Figure 13. This error could be removed by a number of techniques which would result in images of improved appearance as well as improving the classification accuracy. This was not done in this work as the aim was to compare performance of the basic MLC algorithm.

The next section looks at whether the classification accuracy can be improved for the distorted images, particularly if that challenging classification problem is reduced.

FURTHER CONSIDERATION OF CLASSIFICATION ACCURACY

The investigations in the preceding sections noted, and accounted for, the difficult classification problem presented by the difficulty in separating the classes and sub-

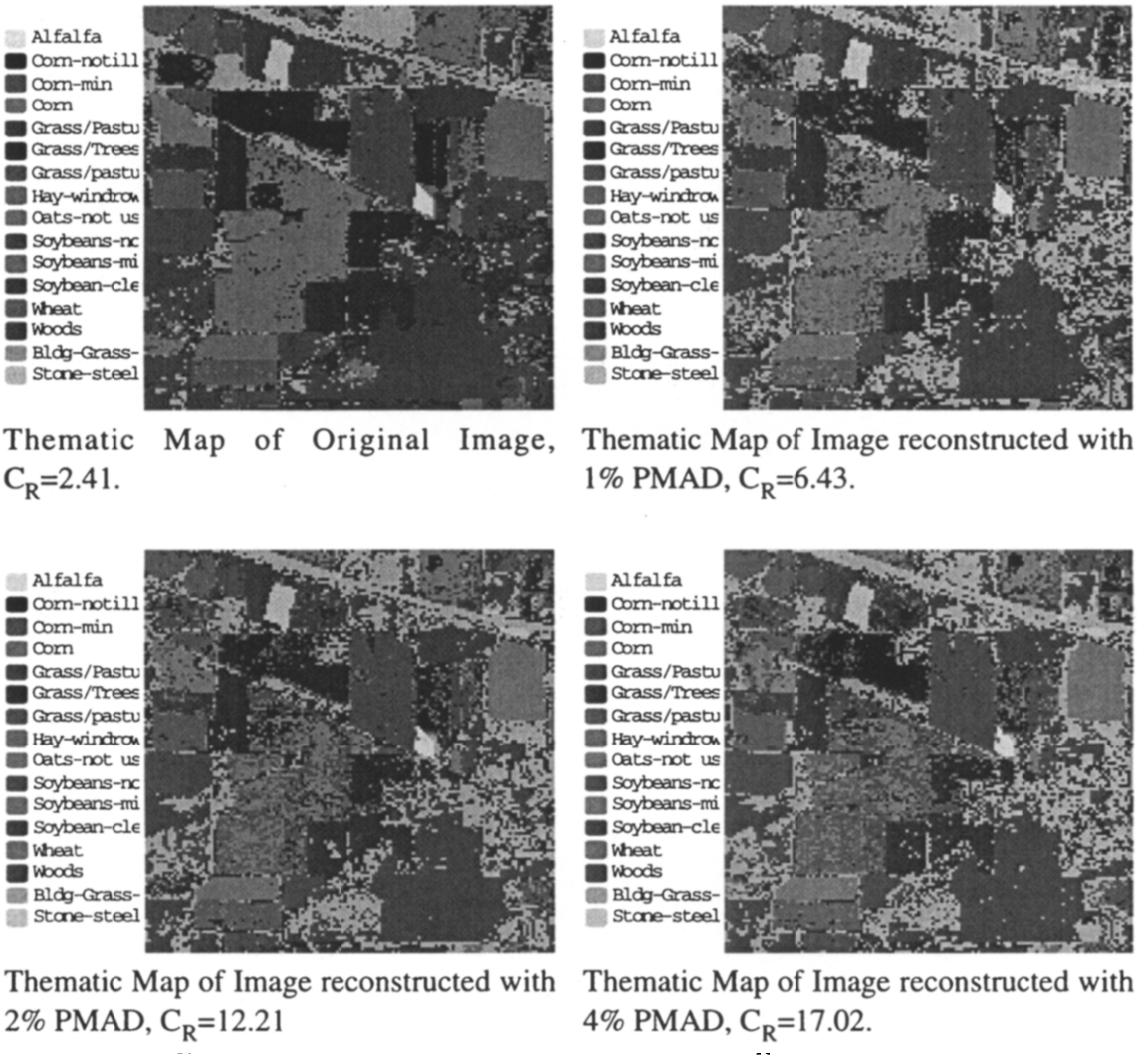


Figure 13. Thematic maps produced for MLC classification of 50 DBFE features for the 92AV3C image reconstructed after compression by M-NVQ to varying degrees of PMAD.

classes of Corn and Soybeans. In both the MLC classification of the original bands and DBFE features these classes were the most difficult to classify accurately as the distortion level increased. Tables 6 and 7 give two extremes of the problem presenting the classification accuracies for 50 DBFE features for the original (undistorted) image and for the image with a 4% PMAD level of allowed distortion. These tables show that the Corn and Soybean subclasses are the cause of most of the error in classification as the distortion level is allowed to increase to 4% PMAD.

The regions with boldface elements in Table 7 show that a significant amount of the confusion arises due to the misclassification of soybeans and corn, either as the same crop with one of the other types of tillage (regions with boldface elements), or as the other type of crop (boxed regions). In particular, the soybeans group of classes is difficult as some 943 of its 4050 pixels are misclassified as a type of corn. A further 824 pixels are misclassified as the wrong type of tillage. Only 44% of soybean pixels are correctly classified. This is very evident visually in the thematic maps of Figure 13.

It is instructive to examine the effect of the PMAD distortion measure if the fine distinctions between these troublesome classes were ignored. It would be artificial to ignore them completely, but given the closeness of the classes, it was considered appropriate to combine each of the three subclasses into the one class group. That is, the three subclasses of Corn were grouped together as one group of Corn-all and the three subclasses of Soybeans were grouped into the one group of Soybeans-all. Tables 8 and 9 present the classification accuracies for the original image and the image with 4% PMAD using the other 10 classes and the two class groups of Corn-all and Soybeans-all.

The effect of the previous "misclassification" of the Corn and Soybean classes is now evident. If classification within these challenging classes is overlooked, the total

Table 6. Multispec© Output-Classification Accuracy for Original 92AV3C Image (MLC of 50 DBFE Features)

Class Name	*Class No.*	*% Correct*	*Number Samples*	*1*	*2*	*3*	*4*	*5*	*6*	*8*	*10*	*11*	*12*	*13*	*14*	*15*	*16*
Alfalfa	1	100.0	54	54	0	0	0	0	0	0	0	0	0	0	0	0	0
Corn-notill	2	95.0	1434	0	1363	10	1	2	2	0	16	37	2	0	0	1	0
Corn-min	3	97.6	834	0	6	814	0	0	0	0	1	12	1	0	0	0	0
Corn	4	100.0	234	0	0	0	234	0	0	0	0	0	0	0	0	0	0
Grass/Pasture	5	99.0	497	0	0	0	0	492	0	0	0	4	1	0	0	0	0
Grass/Trees	6	99.9	747	0	0	0	0	0	746	0	0	0	0	0	0	1	0
Grass/Pasture-mowed	7	0.0	26	0	0	0	0	1	1	23	0	1	0	0	0	0	0
Hay-windrowed	8	100.0	489	0	0	0	0	0	0	489	0	0	0	0	0	0	0
Oats	9	0.0	20	0	5	0	0	9	3	0	0	1	0	0	0	2	0
Soybeans-notill	10	97.2	968	0	3	1	0	2	2	0	941	19	0	0	0	0	0
Soybeans-min	11	89.1	2468	0	100	36	0	12	6	0	80	2198	34	0	0	2	0
Soybeans-clean	12	98.5	614	0	0	3	0	0	0	0	0	5	605	0	0	0	1
Wheat	13	99.5	212	0	0	0	0	0	0	0	0	0	0	211	1	0	0
Woods	14	98.5	1294	0	0	0	0	1	2	0	0	0	0	0	1274	17	0
Bldg-Grass-Tree	15	96.3	380	0	0	0	0	0	0	0	0	0	0	0	14	366	0
Stone-steeltowers	16	100.0	95	0	0	0	0	0	0	0	0	0	0	0	0	0	95
Total			10,366	54	1477	864	235	519	762	512	1038	2277	643	211	1289	389	96
Overall performance (9882/10,366)=95.3%																	

Table 7. Multispec© Output-Classification Accuracy for 92AV3C Compressed with 4% PMAD Distortion (MLC of 50 DBFE Features)

Class Name	*Class No.*	*Percent Correct*	*Number Samples*	*1*	*2*	*3*	*4*	*5*	*6*	*8*	*10*	*11*	*12*	*13*	*14*	*15*	*16*
Alfalfa	1	92.6	54	50	0	0	0	0	0	4	0	0	0	0	0	0	0
Corn-notill	2	60.6	1434	0	**869**	**176**	**17**	0	10	0	**240**	**90**	**23**	0	0	9	0
Corn-min	3	86.9	834	0	**32**	**725**	**8**	0	2	0	**14**	**11**	**37**	0	0	5	0
Corn	4	58.5	234	0	**30**	**59**	**137**	0	2	0	**2**	**1**	**2**	0	0	1	0
Grass/Pasture	5	91.3	497	0	0	0	7	454	9	0	1	1	3	0	12	10	0
Grass/Trees	6	95.7	747	0	0	0	2	4	715	0	1	0	0	0	0	25	0
Grass/Pasture-mowed	7	0.0	26	0	0	0	0	6	0	19	0	0	0	0	0	1	0
Hay-windrowed	8	99.2	489	0	0	0	0	0	0	485	0	0	0	0	0	4	0
Oats	9	0.0	20	0	0	0	0	0	5	0	0	0	0	0	0	15	0
Soybeans-notill	10	94.2	968	0	**12**	**18**	**5**	4	5	0	**912**	**5**	**3**	0	0	4	0
Soybeans-min	11	34.5	2468	0	**179**	**586**	**25**	5	18	2	**696**	**851**	**87**	0	0	19	0
Soybeans-clean	12	73.5	614	0	**31**	**73**	**14**	1	5	0	**24**	**9**	**451**	0	0	4	2
Wheat	13	98.1	212	0	0	0	0	0	2	0	0	0	0	208	0	2	0
Woods	14	94.7	1294	0	0	0	0	8	5	0	0	0	0	1	1226	54	0
Bldg-Grass-Tree	15	69.5	380	0	0	0	0	3	32	0	1	0	1	1	78	264	0
Stone-steeltowers	16	98.9	95	0	0	1	0	0	0	0	0	0	0	0	0	0	94
Total			10,366	50	1153	1638	215	485	810	510	1891	968	607	210	1316	417	96
Overall performance (7441/10,366)=71.8%																	

Table 8. Multispec© Output Showing Classification Accuracy for Original Image with the Six Classes of Corn and Soybeans Combined into Two Groups of Corn-all and Soybeans-all (MLC after DBFE Feature Extraction Using 50 Features)

Class Name	*Group No.*	*Percent Correct*	*Number Samples*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
Alfalfa	1	100.0	54	54	0	0	0	0	0	0	0	0	0	0	0
Grass/Pasture	2	99.0	497	0	492	0	0	0	0	0	0	0	0	0	5
Grass/Trees	3	99.9	747	0	0	746	0	0	0	0	0	0	0	0	0
Grass/Pasture-mowed	4	0.0	26	0	1	1	0	23	0	0	0	0	0	0	1
Hay-windrowed	5	100.0	489	0	0	0	0	489	0	0	0	0	0	0	0
Oats	6	0.0	20	0	9	3	0	0	0	0	0	2	0	5	1
Wheat	7	99.5	212	0	0	0	0	0	0	211	1	0	0	0	0
Woods	8	98.5	1294	0	1	2	0	0	0	0	1274	17	0	0	0
Bldg-Grass-Tree	9	96.3	380	0	0	0	0	0	0	0	14	366	0	0	0
Stone-steeltowers	10	100.0	95	0	0	0	0	0	0	0	0	0	95	0	0
Corn-notill	11	95.8	1434	0	2	2	0	0	0	0	0	1	0	1374	**55**
Corn-min	11	98.3	834	0	0	0	0	0	0	0	0	0	0	820	**14**
Corn	11	100.0	234	0	0	0	0	0	0	0	0	0	0	234	**0**
Soybeans-notill	12	99.2	968	0	2	2	0	0	0	0	0	0	0	**4**	960
Soybeans-min	12	93.7	2468	0	12	6	0	0	0	0	0	2	0	**136**	2312
Soybeans-clean	12	99.3	614	0	0	0	0	0	0	0	0	0	1	**3**	610
Total			10,366	54	519	762	0	512	0	211	1289	389	96	2576	3958
Overall performance (10,037/10,366)=96.8%															

Table 9. Multispec© Output Showing Classification Accuracy for 4% PMAD Image with the Six Classes of Corn and Soybeans Combined into Two Groups of Corn-all and Soybeans-all (MLC after DBFE Feature Extraction Using 50 Features)

Class Name	*Group No.*	*Percent Correct*	*Number Samples*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*
Alfalfa	1	92.6	54	50	0	0	0	4	0	0	0	0	0	0	0
Grass/Pasture	2	91.3	497	0	454	9	0	0	0	0	12	10	0	7	5
Grass/Trees	3	95.7	747	0	4	715	0	0	0	0	0	25	0	2	1
Grass/Pasture-mowed	4	0.0	26	0	6	0	0	19	0	0	0	1	0	0	0
Hay-windrowed	5	99.2	489	0	0	0	0	485	0	0	0	4	0	0	0
Oats	6	0.0	20	0	0	5	0	0	0	0	0	15	0	0	0
Wheat	7	98.1	212	0	0	2	0	0	0	208	0	2	0	0	0
Woods	8	94.7	1294	0	8	5	0	0	0	1	1226	54	0	0	0
Bldg-Grass-Tree	9	69.5	380	0	3	32	0	0	0	1	78	264	0	0	2
Stone-steeltowers	10	98.9	95	0	0	0	0	0	0	0	0	0	94	1	0
Corn-notill	11	74.1	1434	0	0	10	0	0	0	0	0	9	0	1062	**353**
Corn-min	11	91.7	834	0	0	2	0	0	0	0	0	5	0	765	**62**
Corn	11	96.6	234	0	0	2	0	0	0	0	0	1	0	226	**5**
Soybeans-notill	12	95.0	968	0	4	5	0	0	0	0	0	4	0	**35**	920
Soybeans-min	12	66.2	2468	0	5	18	0	2	0	0	0	19	0	**790**	1634
Soybeans-clean	12	78.8	614	0	1	5	0	0	0	0	0	4	2	**118**	484
Total			10,366	50	485	810	0	510	0	210	1316	417	96	3006	3466
Overall performance (8587/10,366)=82.8%															

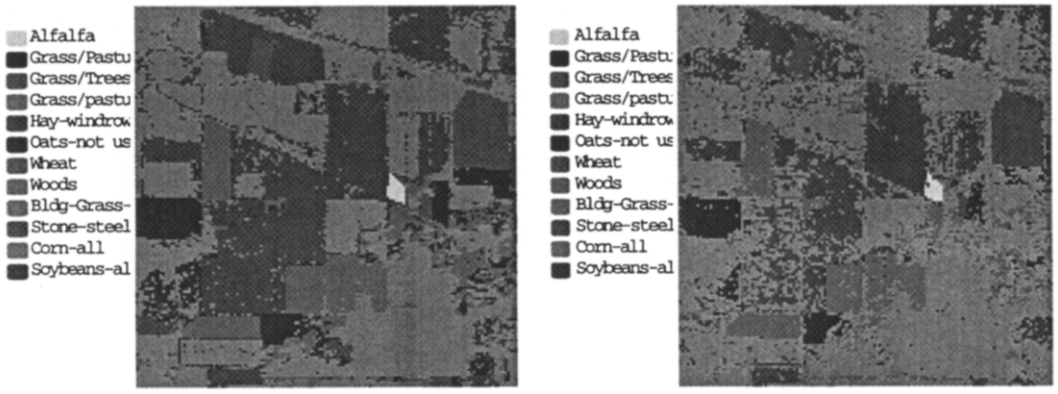


Figure 14. Thematic maps combining the six classes of corn and soybeans into two groups of corn and soybeans for a) the original image and b) the image with 4% PMAD. (MLC after DBFE feature extraction using 50 features.)

classification accuracy is reduced from 96.8% for the original image to 82.8% for the compressed image containing an allowable distortion level of 4% PMAD. This moderate reduction in total classification accuracy is achieved using a compressed image that is 1/17 the size of the original image. Further, it should be noted that, apart from the class of Bldg-Grass-Trees, the other seven classes are classified in the 4% PMAD compressed image with greater than 90% accuracy.

The effect of creating the Corn-all and Soybeans-all groups is further highlighted by Figure 14, which compares the two thematic maps for the original image and that compressed with 4% PMAD distortion. The smaller difference in total classification is evident between the two maps, although the difficulty in distinguishing corn from soybeans (illustrated in Table 9) is again evident.

When the challenging subclasses of Corn and Soybeans are grouped, it is evident from the above discussion and the summary shown in Table 10 that classification is not significantly affected by the additional distortion due to compression.

The total accuracy of classification is reduced from 96.8% for the original image to 82.8% for the image compressed with 4% PMAD. This reduction of 14% appears significant, although Table 10 shows that the majority of that loss is due to poor classification of the classes Bldg-

Table 10. Classification Accuracy for Original Image and 4% PMAD Image with the Six Classes of Corn and Soybeans Combined into Two Groups of Corn-all and Soybeans-all (MLC after DBFE Feature Extraction Using 50 Features)

Class Name	*Group No.*	*Number Samples*	*Percent Correct (0% PMAD)*	*Percent Correct (4% PMAD)*	*% Loss of Classification Accuracy*
Alfalfa	1	54	100.0	92.6	7.4
Grass/Pasture	2	497	99.0	91.3	7.7
Grass/Trees	3	747	99.9	95.7	4.2
Grass/Pasture-mowed	4	26	0.0	0.0	0.0
Hay-windrowed	5	489	100.0	99.2	0.8
Oats	6	20	0.0	0.0	0.0
Wheat	7	212	99.5	98.1	1.4
Woods	8	1294	98.5	94.7	3.8
Bldg-Grass-Tree	9	380	96.3	69.5	**26.8**
Stone-steeltowers	10	95	100.0	98.9	1.1
Corn-notill	11	1434	95.8	74.1	**21.7**
Corn-min	11	834	98.3	91.7	6.6
Corn	11	234	100.0	96.6	3.4
Soybeans-notill	12	968	99.2	95.0	4.2
Soybeans-min	12	2468	93.7	66.2	**27.5**
Soybeans-clean	12	614	99.3	78.8	**20.5**
Total		10,366	96.8	82.8	14.0

Grass-Tree, Corn, and Soybeans. The other seven classes suffer only small reductions in classification accuracy with Grass/Pasture suffering the largest loss of 7.7%.

As noted earlier, the class Bldg-Grass-Tree is not well formed and performs poorly since it is a rather odd assortment of subclasses and it basically represents an "all-other classes" type of grouping. Also, as discussed earlier, the Corn and Soybean classes have been selected because of the difficult classification problem they present. It is therefore to be expected that the introduction of distortion will exacerbate an already difficult decision, and it is unlikely that an application scientist who wanted to achieve such a task would entertain doing so with data that have been compressed in a lossy way.

It is therefore considered that those three class groups have an exaggerated effect on the total classification accuracy of the distorted image. The remainder of the classes do not suffer a significant loss in classification accuracy; yet the compressed image is 1/17 the size of the original. An application scientist could perform initial investigations on a much smaller image (with corresponding advantages in purchase price and storage requirements) and be confident of achieving results that are very close to those that would have been achieved on the original image. Detailed investigations could then be conducted on the original image in areas of particular interest.

This dramatic reduction in image size for a small reduction in classification accuracy reinforces the applicability of the PMAD distortion measure to the compression of hyperspectral imagery. The PMAD measure is particularly useful since the performance of a classifier at one level of distortion can be used to make a reasonable assessment of the performance of the same classifier at other levels of distortion.

SUMMARY

The performance of MLC classification after DBFE feature extraction is close to the ideal performance illustrated by Figure 2. As the level of distortion is increased, the classifier performance decreases. As is desired, it decreases in what appears to be a regular manner commensurate with the increase in distortion. The family of distortion/classification accuracy curves has a regular shape, and, once this shape has been determined for a single image or an image set, it appears that the performance of the classifier can be predicted for a given level of distortion and a given number of features.

Despite some anomalies caused by small classes and challenging discrimination tasks, the classification accuracy of both the total image and its constituent classes varies in a regular way as the level of PMAD distortion is increased. When the challenging subclasses of Corn and Soybeans are grouped, it is evident that classification accuracy is not significantly affected by the additional distortion due to compression. Although the total accuracy of classification is reduced from 96.8% for the original image to 82.8% for the image compressed with 4% PMAD, the loss in classification accuracy is not significant (less that 8%) for most classes other than those that present a challenging classification problem. Yet the compressed image is 1/17 the size of the original.

This dramatic reduction in image size for a small reduction in classification accuracy reinforces the applicability of the PMAD distortion measure for the compression of hyperspectral imagery. The PMAD measure appears particularly useful since the performance of a classifier at one level of compression can be used to make a reasonable assessment of the performance of the same classifier at other levels of compression, including lossless.

However, it must be noted that, in this work, the performance of the PMAD measure has only been investigated in the context of one image. Before more general comments can be made, more images must be obtained with appropriate ground-truth data so that our algorithm can be tested further.

The authors would like to thank the Purdue Research Foundation for freely supplying Multispec© and the 92AV3C data set.

REFERENCES

CCIR (Comité Consutatif International des Radiocommunications) (1974), Method for subjective assessment of the quality of television pictures. In *13th Plenary Assembly*, Recommendation 500. Vol. 11, pp. 65–68.

Gray, R., et al. (1980), Distortion measures for speech processing. *IEEE Trans. Acoust. Speech Signal Process.* ASSP-28(4): 367–376.

Hughes, G. (1968), On the mean accuracy of statistical pattern recognizers. *IEEE Trans. Inf. Theory* IT-14:55–63.

Landgrebe, D. (1994), *Multispectral Data Analysis: A Signal Theory Perspective*, School of Electrical Engineering, Purdue University, West Lafayette, IN.

Landgrebe, D., and Biehl, L. (1995), *An Introduction to Multispec, Version 2.15.95*, School of Electrical Engineering, Purdue University, West Lafayette, IN.

Lee, C., and Landgrebe, D. (1993), Feature extraction on decision boundaries. *IEEE Trans. Pattern Anal. Mach. Intell.* 15(4):388–400.

Limb, J. (1979), Distortion criteria of the human viewer. *IEEE Trans. Syst. Man Cybern.* SMC-9(12):778–793.

Pratt, W. (1979), *Image Transmission Techniques*, Academic, New York.

Richards, J. (1993), *Remote Sensing Digital Image Analysis: An Introduction*, 2nd ed., Springer-Verlag, New York.

Ryan, M., and Arnold, J. (1997), The lossless compression of AVIRIS images by vector quantization. *IEEE Trans. Geosci. Remote Sens.* 35(3):546–550.