LARS Technical Report 062683

FEATURE SELECTION METHODOLOGIES USING SIMULATED THEMATIC MAPPER DATA

M.E. DEAN, R.M. HOFFER

Purdue University/Laboratory for
Applications of Remote Sensing
West Lafayette, Indiana

ABSTRACT

This study* investigated computer classification performances for forest and other cover types using Thematic Mapper Simulator (TMS) data collected by NASA's NS001 scanner. Specifically, results based on the use of a common feature selection measure -- transformed divergence (TD) -- were compared to those based on a principal component transformation for the purpose of evaluating the capabilities of each technique to define: (1) the optimum dimensionality for data sets of this type, and (2) the relative significance of the various wavelength bands with respect to their ability to discriminate among the various cover classes. Expected classification performances as indicated by a minimum Transformed Divergence (TDmin) criteria were compared to actual test classification results. The eigenvectors (i.e. principal components) and eigenvalues for both the overall and the individual class statistics used to classify the TMS data were also used to select waveband subsets to compare to the results from the subsets defined by TD(min).

The results indicated that the use of four wavelength bands will produce considerably better classification than the use of only two or three wavelength bands. However, when more than four wavelength bands were used, overall and individual class performances increased only slightly, thereby indicating that an appropriate set of four wavelength bands probably provide the 'optimum' dimensionality. Classifications using various four wavelength band combinations showed the individual cover class preferences for certain wavebands. These preferences of both individual cover classes and of all classes combined were better indicated by a principal component analysis of the data than by a Transformed Divergence criteria. Further, the results support the use of eigenvectors for identifying the optimal or 'intrinsic' dimensionality of data sets of this type.

I. INTRODUCTION

Much of the previous work with Landsat MSS data has involved the use of all four wavelength bands of the early satellites for distinguishing a wide variety of cover types. Many analysis procedures, including methods for developing training statistics and the development and use of optimum classification algorithms, have been well established through work with Landsat MSS data from a variety of geographic locations. With the advent of the Thematic Mapper (TM) scanner on Landsat IV, questions involving effective and efficient techniques for handling the increased spatial resolution and number of spectral bands once again confront the remote sensing community.

Since remote sensor data often has high interband correlations (6), there is a redundancy of information which is source dependent such that the 'intrinsic dimensionality' or the dimensionality required to characterize a specific data set is often less than the number of available bands (4). The value of data compression is evident when one considers the cost of storage and classification of data sets having many wavebands such as those obtained from the Thematic Mapper.

There are two common approaches to reduce the dimensionality or feature space of the data. One approach that has been used frequently involves manual selection

* This work was supported by NASA under Contract No. NAS9-15889.

of an optimum subset of the original bands based upon either *a priori* knowledge and/or upon one of a number of statistical separability measures. The second approach involves a linear transformation of the original bands to a set of uncorrelated new orthogonal transformed components in which a maximum amount of spectral variation is accounted for in descending order along the transformed components; i.e. the maximum variation of the data is accounted for in both direction and magnitude by the first component, the second greatest amount of variation by the second component, and so on. One such linear transformation is the Karhunen-Loêve or principal component transformation. In this procedure, the eigenvectors or latent roots, ($\vec{x}$), of an NxN matrix, A, satisfying the equation,

$$A\vec{x} = \lambda\vec{x} \qquad (1)$$

are found by solving the characteristic equation

$$f(\lambda) = |A - \lambda I| = 0. \qquad (2)$$

The roots of this polynomial (values of λ_i which make the polynomial 0) are the eigenvalues of matrix A. These are ordered such that:

$$\lambda_1 \geq \lambda_2 \geq \cdot \cdot \cdot \geq \lambda_N$$

The eigenvectors ($\vec{x}$) of matrix A are found by substituting the values of λ_i in equation (1). In essence, the eigenvectors define a new set of orthogonal coordinates whose direction cosines are the normalized characteristic vectors corresponding to the ordered characteristic roots, λ_i, of matrix A. In the case where A is an NxN covariance matrix, the eigenvectors define a coordinate system projected through the directions of maximum variance of the data in N-dimensions. The first component has direction through the maximum variance with length proportional to the square root of the first eigenvalue. The remaining characteristic roots, λ_i, and vectors, ($\vec{x}_i$), of A determine the lengths and orientations of the second and higher component axes, each in the direction of the maximum variance remaining in the data (3,5). By compressing the data variance or information content onto a fewer number of coordinate axes, a principal component transformation of multispectral scanner data can provide an efficient method of dimensionality reduction. Generally, a subset of the three or four higher ordered eigenvectors will account for almost all of the information contained in the entire set of the original wavelength bands. These components, therefore, can be used to classify the data with a minimum number of features and will result in approximately the same classification performance as if all of the original wavelength bands had been used.

One concern, however, of the use of principal components is the potential loss of descriptive information about the relative importance of the various wavelength bands to the individual cover classes. However, the coefficients or loadings, as they are sometimes referred to, of the eigenvectors can often provide a qualitative indication of the relative importance or contribution of the original features to each of the eigenvectors (5). This type of qualitative analysis has been done using multispectral scanner data from earth surface features to define the optimum spectral bands or wavelength regions which best characterize those surface features (7,10). In this way the loadings can be used to identify those bands which best characterize a particular data set and can therefore be used as an alternative feature selection method; i.e. to use the loadings rather than statistical separability measures, such as transformed divergence, for selecting an optimum waveband subset. In addition, the eigenvalues provide an indication of the intrinsic dimensionality of a data set. In summary, therefore, the optimum dimensionality of the data set can be determined from the eigenvalues and, in addition, the specific wavelength bands having the greatest information content can be defined using the eigenvector coefficients.

II. OBJECTIVES

The objectives of this study were:

1) to determine the intrinsic dimensionality of this simulated Thematic Mapper data set, and
2) to evaluate the effectiveness and sensitivity of 'standard' statistical separability measures (i.e. transformed divergence) in comparison to eigenvectors for identifying the optimum subset of the original TMS bands for classifying the various cover types.

III. MATERIALS AND METHODS

Thematic Mapper Simulator (TMS) data were collected on May 2, 1979 by NASA's NS001 aircraft multispectral scanner over a bottomland forested area in South Carolina near the city of Camden. Wavelength bands on this scanner include three bands in the visible portion of the spectrum (CH1:0.45 - 0.52µm; CH2:0.52 - 0.60µm; CH3:0.63 - 0.69µm), two bands in the near IR (CH4:0.76 - 0.90µm; CH5:1.00 - 1.30µm), one band in the middle IR (CH6:1.55 - 1.75µm) and one band in the thermal IR region (CH7:10.40 - 12.50µm). The test site is located in an area between the Piedmont plateau and the coastal plain. This area is characterized by large tracts of bottomland hardwoods, but includes smaller tracts of pine plantations and of agriculatural fields at varying stages of growth. Table 1 lists the designated cover classes found in the Camden test site.

Table 1. Descriptions of the various cover classes in the Camden test site.

Cover Class	Description
PINE	Pine forest areas, primarily plantations of slash and loblolly of varying age.
HDWD	Bottomland hardwoods such as sweetgum, willow, and bottomland oaks; mostly in dense old age stands.
TUPE	Water tupelo, primarily associated with narrow oxbow lakes and other areas of inundated soils.
CCUT	Areas subjected to clearcut forestry practices; clearcuts are in various stages of regrowth and may include windrowed slash.
PAST	Pastures and old fields.
CROP	Agricultural crops at various stages of development.
SOIL	Primarily areas of recently tilled agricultural fields, but may include some minespoil and recent clearcut areas.
WATER	Water areas include the Wateree River, small lakes and ponds, and turbid minespoil ponds.

In order to achieve the stated objectives, the following set of analysis procedures were used:

(1) Supervised training statistics were generated for the classes listed in Table 1 using a transformed divergence (TD) measure to evaluate the spectral separability of the cover class statistics (8). Likewise, the selection of optimum waveband subsets of two and greater were based upon a minimum TD criterion (8). The training areas were carefully selected so that they would comprise an exhaustive and representative set of all spectral classes within the scene.

(2) A statistical sample of test areas was selected using a procedure described previously (1).

(3) A set of eigenvectors and their associated eigenvalues were calculated both for individual cover class training statistics and for a combined or merged training data set generated from all spectral class training statistics.

(4) A Gaussian Maximum Likelihood algorithm (8) was used to classify the area using the set of defined supervised training statistics as input for each of the various wavelength band subsets.

Certain limitations of using separability measures such as TD have been addressed elsewhere in the literature (8,9). If it is possible to assume that the training statistics actually represent an exhaustive set of all spectral variability within the scene, then these limitations, to a large extent, result from the fact that most separability measures, including transformed divergence, only have an indirect relationship to the probability of error. In addition, when calculating such separability measures, class *a priori* probabilities are often unknown and are therefore assumed to be equal; this can cause the estimation of P_e to deviate considerably from the actual P_e. Following this, one additional advantage of a principal component analysis is that the eigenvectors inherently incorporate *a priori* probabilities in their calculation as long as the sample covariance matrix has been generated from a representative, i.e. proportional, set of all cover classes. (This can be achieved if a statistical sample of points are taken from which the sample covariance matrix is calculated.) One of the main purposes of this study, therefore, is to evaluate the effectiveness of such feature selection procedures in defining optimum wavelength band subsets, i.e. subsets which minimize the probability of error, $P_e = 1 - P_c$.

IV. RESULTS AND DISCUSSION

A. Intrinsic Dimensionality

Data sets which have multivariate normal distribution in an N-dimensional feature space often times exhibit a non-spherical distribution in that feature space. That is, the variance is often not equal, but differs widely between bands, so that the data in N-dimensions resembles more of a multidimensional ellipsoid. The eigenvectors (principal components or latent roots) and eigenvalues of a source covariance matrix define a set of orthogonal axes which result from a rigid rotation of the original coordinate axes (variables) to an orientation determined by the direction of maximum data variance of this multidimensional ellipsoid. The first component is positioned through the maximum data spread, the second through the next greatest amount of data spread and so on (5). This linear transformation of the original bands eliminates any interband correlation and concentrates a maximum amount of the data variance onto a fewer number of features. If the potential for characterizing a remote sensing data set lies in the ability to define the distribution, i.e. variance or spread, of the data in the feature space, then such transformed axes theoretically allow the data to be characterized with a minimum number of variables or coordinate axes. The intrinsic dimensionality of a data set can therefore be determined by observing when most of the total data source variance has been accounted for by a subset of the ordered eigenvectors.

Table 2a. The ordered eigenvectors, their associated eigenvalues and the percent of total data variance they each account for of an overall category covariance matrix.

Wavelength Band	Eigenvector (Component) 1	2	3	4	5	6	7	
1	-0.12	0.22	-0.33	-0.17	0.21	-0.67	0.56	
2	-0.17	0.37	-0.52	-0.29	0.39	0.15	-0.55	
3	-0.15	0.46	-0.36	0.16	-0.54	0.44	0.35	
4	0.76	0.01	-0.24	-0.35	-0.42	-0.20	-0.17	
5	0.58	0.19	0.03	0.14	0.57	0.39	0.37	
6	0.16	0.49	0.13	0.68	-0.03	-0.38	-0.33	
7	-0.06	0.57	0.65	-0.51	-0.06	0.01	0.02	
	2128.2	1779.1	591.0	95.6	12.7	7.3	2.9	Eigenvalue
	46.1%	38.5	12.8	2.1	0.3	0.1	0.1	Percent Variation
	46.1%	84.6	97.4	99.5	99.8	99.9	100.0	Cumulative Percent Variation

Table 2b. The ordered eigenvectors, their associated eigenvalues and the percent of total data variance they each account for of a Tupelo category covariance matrix.

Wavelength Band	Eigenvector (Component) 1	2	3	4	5	6	7	
1	-0.02	0.30	0.10	0.18	0.80	0.39	-0.29	
2	-0.08	0.71	0.23	-0.40	-0.26	-0.19	-0.42	
3	-0.07	0.46	0.16	-0.09	0.04	0.20	0.84	
4	0.79	0.04	-0.06	-0.24	0.32	-0.44	0.12	
5	0.59	0.06	0.12	0.15	-0.41	0.65	-0.12	
6	0.07	0.23	0.31	0.82	-0.10	-0.41	0.01	
7	-0.02	-0.38	0.90	-0.22	0.09	0.00	0.01	
	305.2	23.3	17.7	2.0	1.2	1.0	0.5	Eigenvalue
	87.0%	6.6	5.0	0.6	0.3	0.3	0.2	Percent Variation
	87.0%	93.6	98.6	99.2	99.5	99.8	100.0	Cumulative Percent Variation

Table 2c. The ordered eigenvectors, their associated eigenvalues and the percent of total data variance they each account for of a Crop category covariance matrix.

Wavelength Band	Eigenvector (Component) 1	2	3	4	5	6	7	
1	0.07	-0.05	0.26	0.27	-0.42	0.82	0.08	
2	-0.04	0.09	0.60	0.22	0.27	-0.18	0.68	
3	-0.11	0.17	0.55	0.23	0.30	0.00	-0.71	
4	0.87	0.16	0.26	-0.28	-0.20	-0.15	-0.07	
5	0.37	0.36	-0.42	0.38	0.57	0.28	0.07	
6	0.01	0.29	-0.13	0.67	-0.52	-0.42	-0.04	
7	-0.29	0.85	0.02	-0.39	-0.14	0.13	0.07	
	1036.2	69.9	19.1	9.3	2.7	1.8	0.8	Eigenvalue
	90.9%	6.1	1.7	0.8	0.3	0.1	0.1	Percent Variation
	90.9%	97.0	98.7	99.5	99.8	99.9	100.0	Cumulative Percent Variation

Table 2a lists the eigenvectors (i.e. principal compontents) defined for the covariance matrix of the combined supervised statistics, and Tables 2b and 2c show the covariance matrices defined for two of the individual cover class statistics (Tupelo and Crop, respectively). Figure 1 shows graphically the amount of information associated with each of the ordered eigenvectors or components for the combined supervised statistics; i.e. it graphically depicts the eigenvalues listed in Table 2a. In examining the cumulative percent variation indicated in Tables 2a,b, and c, it is evident that the intrinsic dimensionality of this data set, as described by the eigenvalues of the ordered eigenvectors, appears to be approximately four; in other words, the majority of the data variance has been accounted for by the first four eigenvectors of each of these sample covariance matrices.

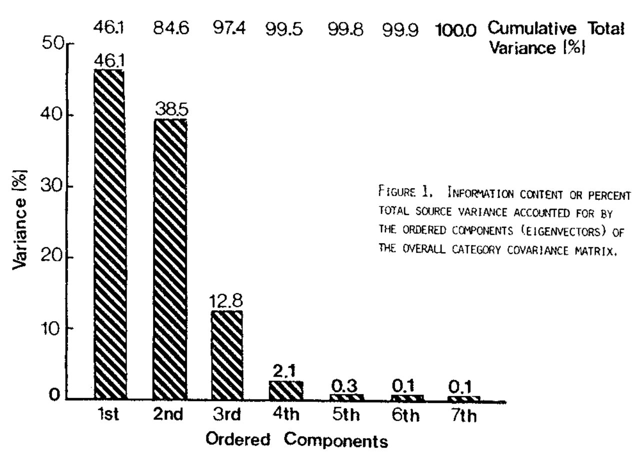

FIGURE 1. INFORMATION CONTENT OR PERCENT TOTAL SOURCE VARIANCE ACCOUNTED FOR BY THE ORDERED COMPONENTS (EIGENVECTORS) OF THE OVERALL CATEGORY COVARIANCE MATRIX.

Table 3 lists the optimum waveband subsets as selected by a minimum transformed divergence criteria, TD(min). Figures 2 and 3 show the overall and individual class performances, respectively, for each of these waveband subsets. Actual classification performance values for these are given in Table 4.

Table 3. Optimum waveband subsets by combination level as determined by a TD(min) criterion.

Combination Level	Waveband Subset
2	(2,5)
3	(1,3,6)
4	(2,4,5,7)
5	(2,3,4,6,7)
6	(1,2,4,5,6,7)
7	(1,2,3,4,5,6,7)

These response surfaces further corroborate that the intrinsic dimensionality is approximately four; i.e. any dimensionality greater than four does not result in a significant increase in either the individual class or overall classification performances.

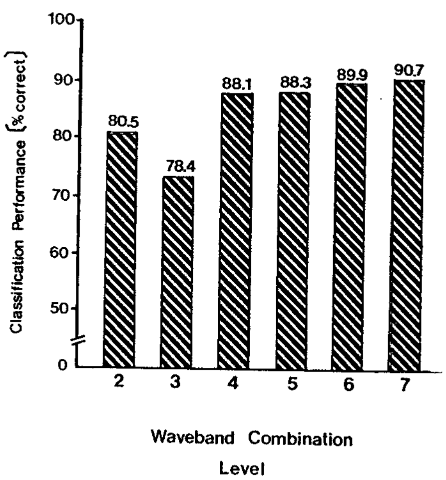

FIGURE 2. OVERALL CLASSIFICATION PERFORMANCE OF THE "BEST" 2 THROUGH ALL 7 WAVEBAND SUBSETS AS SELECTED BY A TD(MIN) CRITERION.

Some individual class performances actually decrease, slightly, due to their specific preferences for certain wavebands. It should be noted, however, that the use of transformed divergence measures do not provide as effective an indication of the intrinsic dimensionality of the data as is the case with a principal component analysis.

B. Waveband Analysis

Since four TM bands appear to be 'optimum' for both individual and overall classification for this data set, a waveband analysis with various four band subsets was performed in order to evaluate the impact of certain wavelength bands on individual cover classes. The results of both the individual cover class and overall classification performances are shown in Figures 4 and 5. Actual classification performance values for these are given in Table 5.

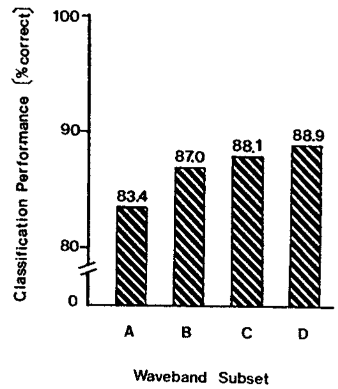

FIGURE 4. OVERALL CLASSIFICATION PERFORMANCE OF VARIOUS FOUR WAVEBAND SUBSETS.*

*SUBSET	WAVEBANDS	
A	(3,5,6,7)	
B	(2,4,6,7)	
C	(2,4,5,7)	("BEST" AS SELECTED BY TD)
D	(2,3,4,5)	

Although a TD(min) criterion selected the four band subset (2,4,5,7) as the best (see Table 3), overall classification performance increased slightly and many individual class performances increased significantly with the use of bands (2,3,4,5) as shown in Figures 4 and 5. Transformed divergence measures between all possible combinations of spectral class pairs for each of the four band subsets contained only one or two class pairs which had TD values less than 1800, while all other class pairs in all of the four band subsets were greater than 1900*. These results therefore suggest that TD is a relatively insensitive measure for estimating the probability of error and, subsequently, the probability of correct classification, for a given data set.

Further analysis involved the use of the loadings of the eigenvectors of both individual cover class and overall or combined class covariance matrices for identifying important (significant) wavelength

* Transformed divergence (TD) values range from zero (identical spectral classes) to 2000 (completely separable spectral classes).

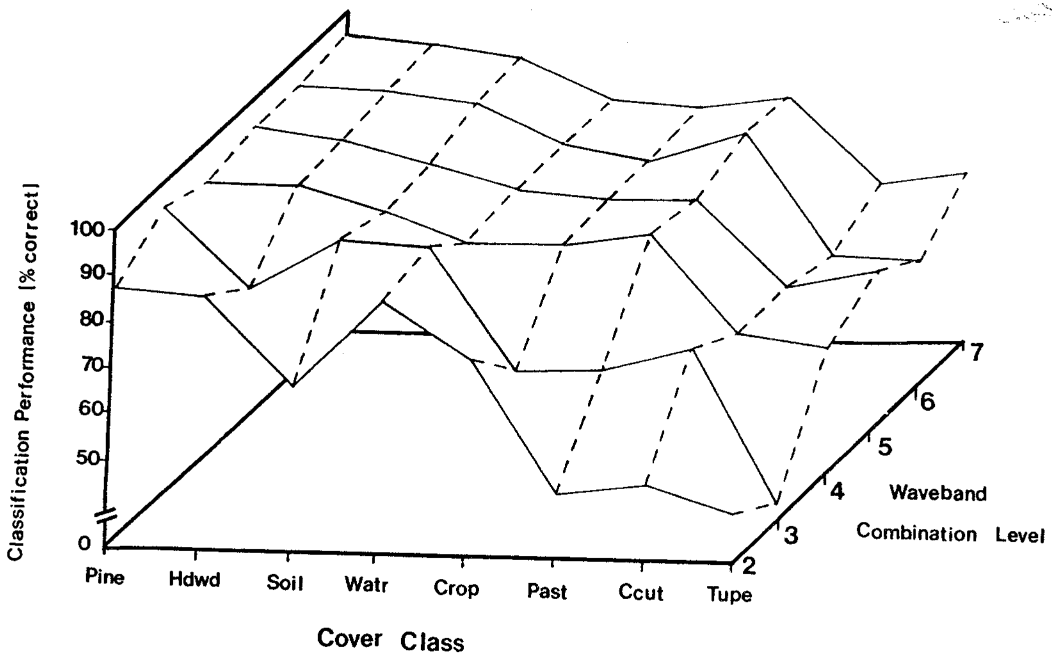

FIGURE 3. INDIVIDUAL COVER CLASS PERFORMANCES OF THE "BEST" 2 THROUGH ALL 7 WAVEBAND SUBSETS AS SELECTED BY A TD(MIN) CRITERION.

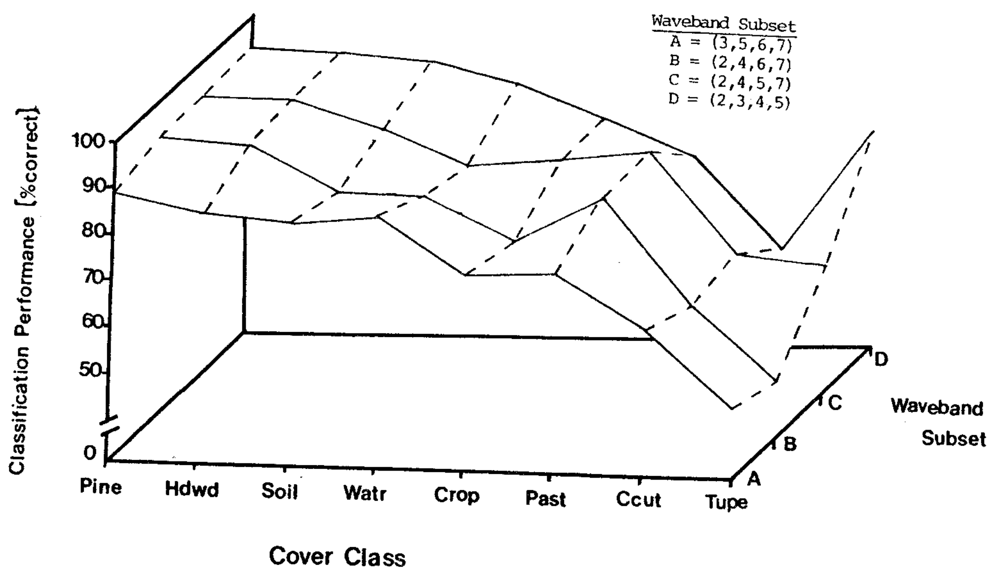

FIGURE 5. INDIVIDUAL COVER CLASS PERFORMANCES OF VARIOUS FOUR WAVEBAND SUBSETS.

bands. The loadings of the first four eigenvectors for the covariance matrices of the combined category, the Tupelo category and Crop category, respectively, are shown in Figures 6-8. These graphically portray the values listed in Tables 2a-2c. As shown in Figure 6, the first eigenvector of the combined category weighted bands 4 and 5 as the highest, and the second eigenvector weighted bands 2,3,6 and 7 relatively high. Comparing these values with the overall classification performances, it is apparent that although the subset of bands (2,3,4,5) performed slightly better than bands (2,4,5,7), both subsets suggest the importance of bands 4 and 5 which were both heavily weighted in the first eigenvector. The other two bands included in each of these subsets were all fairly significant in the second eigenvector.

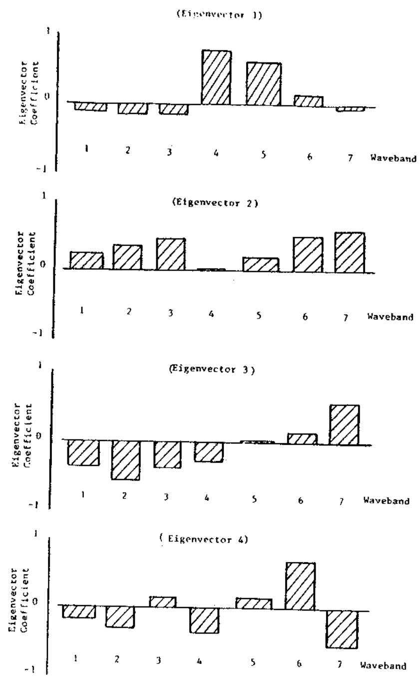

Figure 6. The eigenvector coefficients or loadings of the first four ordered eigenvectors (principal components) from the overall category covariance matrix.

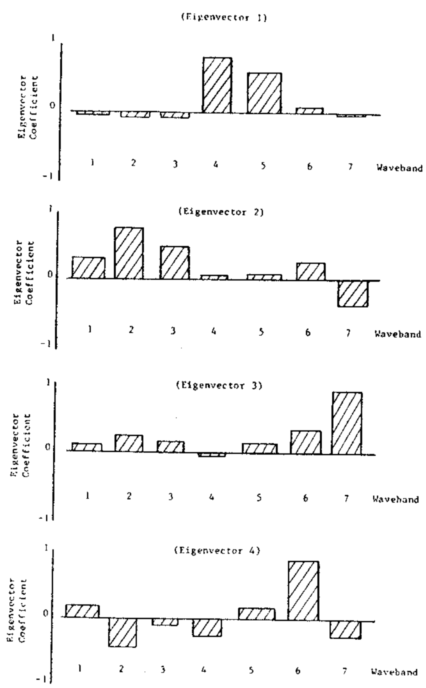

Figure 7. The eigenvector coefficients or loadings of the first four ordered eigenvectors (principal components) from the Tupelo category covariance matrix.

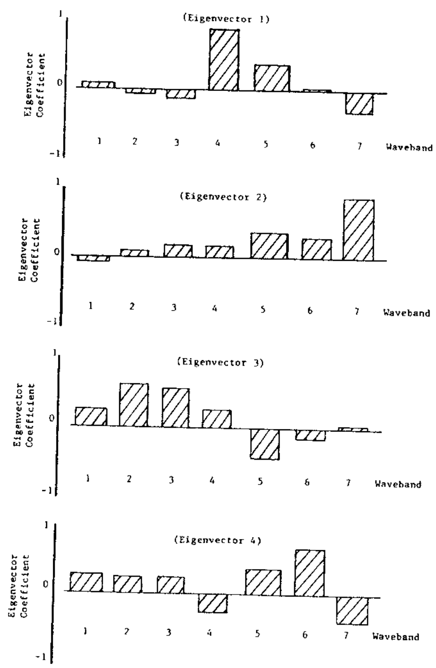

Figure 8. The eigenvector coefficients or loadings of the first four ordered eigenvectors (principal components) from the Crop category covariance matrix.

A more sensitive analysis can be seen with the individual cover class results. In the case of the Tupelo category, bands 4 and 5 in the first eigenvector and bands 2 and 3 in the second eigenvector have the highest loadings. All other bands in these first two eigenvectors had much smaller loadings. This preference for bands 2,3,4 and 5 is reflected in the significant increase in classification performance for Tupelo with bands (2,3,4,5) over any of the other wavelength bands. Further, the first two eigenvectors of the Tupelo covariance matrix account for 94% of the total variance of the Tupelo statistics, so that most of the information is contained in the first two eigenvectors. Therefore, even if a band had a high loading in the third or lower ordered eigenvector, it would not contribute significantly to the total variance of the Tupelo class.

Table 4. Overall and individual cover class performances for the waveband subsets selected by TD(min).

Cover Class	Waveband Combination Level 2	3	4	5	6	7
Pine	87.0%	94.7%	91.0%	93.8%	93.0%	95.0%
Hardwood	85.9	77.8	91.1	90.9	92.7	93.2
Tupelo	41.5	21.2	58.5	66.1	57.6	67.8
Clearcut	47.3	68.1	60.5	61.6	59.2	64.9
Pasture	44.6	62.3	82.6	80.6	85.7	83.4
Crop	73.7	61.5	79.7	79.9	78.9	81.0
Soil	66.1	89.8	85.6	86.2	90.4	90.6
Water	86.3	88.0	78.7	80.7	81.3	81.7
Overall	80.5%	78.4	88.1	88.3	89.9	90.7

The eigenvectors of the Crop class covariance matrix have the highest loadings from band 4 and, to a lesser extent, band 5 in the first eigenvector and from band 7 and, again to a lesser degree, band 5 in the second eigenvector. Again this preference of these bands is reflected in the performance of the Crop category: the subset of bands (2,4,5,7) had the highest individual class performance, followed by the subset (2,3,4,5). One reason that subset (2,4,6,7) didn't perform nearly as well for Crop as with the previous two subsets might be that although band 5 is not weighted as heavily in the first eigenvector as band 7 is on the second, band 5 may actually account for more spectral data variance since the first eigenvector accounts for almost 91% of the total data variance alone. Hence, bands 4 and 5 may be the most significant with band 7 providing some additional information, and thus a subset of bands including both bands 4 and 5 would provide optimum classification of the Crop category.

Finally, the coefficients of the eigenvectors from the combined cover class covariance matrix shows the highest loading from bands 4 and 5 in the first component (46.1% of the total data variance) and from bands 2,3,6, and 7 in the second component (38% of the total variance). Here again the overall performance was the best when both bands 4 and 5 were included, thereby indicating the importance of the near infrared portion of the spectrum. Although subsets (2,4,6,7) and (3,5,6,7) included all wavelength regions, (i.e. visible, near IR, middle IR and thermal IR), they did not perform as well overall as either of the subsets (2,4,5,7) or (2,3,4,5) which were the waveband combinations defined by one of the feature selection techniques being evaluated.

Table 5. Overall and individual cover class performances for selected four waveband subsets.

Cover Class	Waveband Subset A(3,5,6,7)	B(2,4,6,7)	C(2,4,5,7)	D(2,3,4,5)
Pine	89.5%	92.3%	91.0%	92.6%
Hardwood	85.7	90.7	91.1	91.8
Tupelo	46.6	42.4	58.5	78.0
Clearcut	63.0	58.6	60.5	51.4
Pasture	74.9	82.3	82.6	71.1
Crop	73.7	71.5	79.7	79.1
Soil	84.2	81.0	85.6	90.3
Water	86.3	81.0	78.7	86.3
Overall	83.4%	87.0%	88.1%	88.9%

In summary, the eigenvectors and eigenvalues of a covariance matrix from an MSS data set can be obtained without having to actually transform the data, and will provide descriptive information about the data including the relative importance of the wavebands and also the intrinsic dimensionality of the data set. Therefore, this type of analysis can provide an additional or alternative feature selection procedure which the analyst can use with the original data set. It may be that the user will want to actually transform the data set using a principal component transformation and then subsequently classify this transformed data set using a subset of the higher ordered components. However, since such data transformations usually require significant amounts of computer (CPU) time, this approach may not necessarily be desirable. In addition, it should be pointed out that the sensitivity of a principal component analysis is highly dependent upon the structure of the data set. As discussed by Jenson and Waltz (2), the effectiveness of orthogonal transformations such as canonical analysis and principal components lies, to a great extent, in the degree of the correlations among the bands for a given data set; thus, the greater the interband correlation, the more effective the transformation in dimensionality reduction. Further, it is possible to envision situations in which the maximum data spread might be defined by two or more relatively unimportant and/or infrequent spectral classes. In this case, the first eigenvector may be projected through this 'unimportant' data spread and actually cause other, more important spectral classes to lose some distinguishing spectral information as a result of the transformation. Therefore, a principal component analysis, as with other feature selection procedures, can only provide the analyst with a descriptive tool with which to analyze his data. Familiarity of the analyst with the characteristics of his particular data set, e.g. spectral variance and data structure, cannot be overemphasized.

V. CONCLUSIONS

The intrinsic dimensionality of data sets of this type appears to be about four; at any dimensionality greater than four, overall classification performance as well as performances for individual cover classes do not increase significantly. This dimensionality generally can not be inferred from separability measures such as transformed divergence, but can only be determined in such cases by performing a series of classifications using 'optimum' two and greater waveband combinations and comparing the resulting classification performances; i.e. finding where classification performance begins to level off. However, the sum of the eigenvalues of the ordered eigenvectors of a data set can provide insight into the amount of significant spectral 'information' within that data set, and hence give an indication as to the number of bands required to achieve a leveling off of classification performance; i.e. the intrinsic dimensionality of the data.

In addition, from this study there appears to be a correlation between the coefficients of a set of ordered eigenvectors or principal components and the relative importance of the original wavelength bands in a set of multispectral scanner data. This seems to be true both for a general, multiclass situation and for individual cover classes. In other words, if the sample from which the covariance matrix is calculated is from a particular cover class, then the coefficients of the eigenvectors will indicate which original bands may best characterize that cover class. If, on the other hand, the sample includes many cover classes of varying a priori probabilities, then the eigenvector coefficients will indicate the overall relative importance of the original wavebands for the entire data set. The individual cover classes in this case may or may not be optimally represented in the eigenvectors, since those having high a priori probabilities and/or relatively large spectral variance may exert more influence on the calculation of the covariance matrix and, hence, the resulting eigenvectors.

VI. REFERENCES

1. Dean, M.E. and R.M. Hoffer. 1982. An Evaluation of Thematic Mapper Simulator Data for Mapping Forest Cover. Symposium on Machine Processing of Remotely Sensed Data, LARS/Purdue University, W. Lafayette, IN. 47906-1399.pp.300-307.

2. Jenson, S.K. and F.A. Waltz. 1979. Principal Components Analysis and Canonical Analysis in Remote Sensing. Proc. of the American Society of Photogrammetry, 45th Annual meeting, March 18 - 24, 1979. Washington, D.C. vol.1, pp.337-348

3. Kolman, B. 1970. Elementary Linear Algebra. Macmillan Publishing Co., Inc. New York. 255pp.

4. Landgrebe, D.A. 1978. Useful Information from Multispectral Image Data: Another Look. pp. 336-374. in Swain, P.H. and S.M. Davis (eds). Remote Sensing: The Quantitative Approach, McGraw-Hill Inter. Book Co., N.Y.

5. Morrison, D.F. 1967. Multivariate Statistical Methods. McGraw-Hill Book Co., N.Y. (2nd edition). 415pp.

6. Ready, P.J. and P.A. Wintz. 1972. Multispectral Data Compression Through Transform Coding and Block Quantization. LARS Information Note 050572, LARS/Purdue University, W. Lafayette, IN. 47906-1399. 148pp.

7. Schreier, H., L.C. Goodfellow and L.M. Lavkulich. 1982. The Use of Digital Multi-Date Landsat Imagery in Terrain Classification. Photogr. Eng. and Remote Sensing, Vol. 48(1), pp. 111-119.

8. Swain, P.H. and S.M. Davis (eds). 1978. Remote Sensing: The Quantitative Approach, McGraw-Hill Book Co., N.Y. 396pp.

9. Whitsitt, S.J. and D.A. Landgrebe. 1977. Error Estimation and Separability Measures in Feature Selection for Multiclass Pattern Recognition. LARS Publication 082377, LARS/Purdue University, W. Lafayette, IN. 47906-1399. 396pp.

10. Wiersma, D.J. and D.A. Landgrebe. 1980. Analytical Design of Multispectral Sensors. IEEE Transactions on Geoscience and Remote Sensing, vol. GE-18, No. 2, pp. 180-189.

AUTHOR BIOGRAPHICAL DATA

M. Ellen Dean is a research associate at the Laboratory for Applications in Remote Sensing (LARS), Purdue University, and has been primarily involved with the evaluation of analysis techniques for use with simulated and, currently, for actual Thematic Mapper data. She is currently pursuing an M.S. in Remote Sensing through the Department of Forestry and Natural Resources at Purdue and received a B.S. in Forestry from Virginia Polytechnic Institute and State University in 1979.

Roger M. Hoffer is Professor of Forestry and Leader of Ecosystems Research Programs, LARS, Purdue University. B.S. in Forestry, Michigan State University; M.S. and Ph.D. in Watershed Management, Colorado State University. Co-founder of LARS in 1966, Dr. Hoffer currently teaches three courses in Remote Sensing of Natural Resources. He has lectured, consulted, and participated in remote sensing activities in many countries throughout North and South America, Asia, and Europe; has served as a principal investigator on Landsat, Skylab, and other major remote sensing projects, and has authored more than 130 scientific papers and publications on remote sensing. Professor Hoffer's research has focused on the use and refinement of computer-aided analysis techniques for forestry applications and on the study of the spectral characteristics of earth surface features.