Spectral Shape Classification of Landsat Thematic Mapper Imagery

Mark J. Carlotto

Abstract

A multispectral classifier based on an alternative spectral representation is described, and its performance over a full Landsat Thematic Mapper (TM) scene is evaluated. Spectral classes are represented by their spectral shape — a vector of binary features that describes the relative values between spectral bands. An algorithm for segmenting or clustering TM data based on this representation is described. After classes have been assigned to a subset of spectral shapes within training areas, the remaining spectral shapes are classified according to their Hamming distance to those that have already been classified. The performance of the spectral shape classifier is compared to a maximum-likelihood classifier over five sites that are fairly representative of the full Landsat scene considered. Although the performance of the two classifiers is not significantly different within a site, the performance of the spectral shape classifier is significantly better than the maximum-likelihood classifier across sites. Analysis of results suggest that the spectral shape classifier is relatively insensitive to seasonal changes between wetland and upland areas in the scene and is not affected by thin clouds over one of the sites. A full-scene spectral shape classifier is then described which combines spectral signature files that associate classes with spectral shapes derived over the five sites into a single file that is used to classify the full scene. The classification accuracy of the full-scene spectral shape classifier is shown to be superior to that of a stratified maximum-likelihood classifier.

Introduction

Accurate and reliable classification of multispectral imagery over extended areas is critical to the development of land-cover maps for a variety of applications in a timely and cost-effective manner. Yet after over 20 years, the generation of such maps from imagery in an operational manner remains a labor-intensive and costly process. Conventional statistical classifiers perform well over limited areas where spectral signatures do not vary greatly from those captured in the training data. However, as the size of area to be classified increases, the classification accuracy typically decreases due to environmental, topographical, and phenological factors. The most common method of classifying large heterogeneous regions is by spatial stratification whereby the scene is divided into regions (e.g., based on climate, topography, etc.), each region is classified separately, and the results are combined (Todd *et al.*, 1980; Hutchinson, 1982). An advantage of stratification is that regional knowledge can be used to significantly improve classification accuracy. By labeling each part of the scene separately and combining the results, the overall classification accuracy does not have to be sacrificed for area coverage. A disadvantage of stratification, however, is that it requires a certain amount of interactive processing as well as additional data (e.g., elevation matrices, maps) which add to the processing cost. Signature extension techniques (Henderson, 1974) and extendible classification algorithms (Carlotto, 1990) provide an alternative to stratification in which spectral signatures derived over limited portions of a scene are used to classify the remainder of the scene and, in some cases, other scenes as well.

An early rule-based multispectral classifier (Carlotto *et al.*, 1984) used qualitative knowledge and relative constraints for classifying general land-cover categories. Two kinds of rules were developed: those that defined classes relative to each other in terms of their spectral features (e.g., the greenness of vegetation is greater than the greenness of bare soil, etc.), and those that defined single classes in terms of the relative values between spectral bands. The spectral shape classifier (Carlotto and Tom, 1985) was an outgrowth of this work and addressed the problem of deriving a complete and consistent set of spectral classification rules from training data. In this paper, we describe a new method of classifying multispectral imagery based on a set of binary features that represent the relative values between spectral bands. We begin by describing the theoretical basis of the spectral shape representation. The spectral shape representation is compared to K-means clustering, and its use in multispectral classification is described. The software architecture of the spectral shape classification system is then outlined. Experimental results from a full Landsat scene are used to examine the classification accuracy of the spectral shape classifier within and across training areas and to compare its overall accuracy to that of a maximum-likelihood classifier.

Spectral Shape Representation

Instead of correcting for terrain and atmospheric effects systematically, we seek a representation (i.e., a set of invariant features) for classification that will be less sensitive to these effects, perhaps at the expense of loosing some spectral detail. Our method represents the shape of the spectral response in terms of the relative values between bands

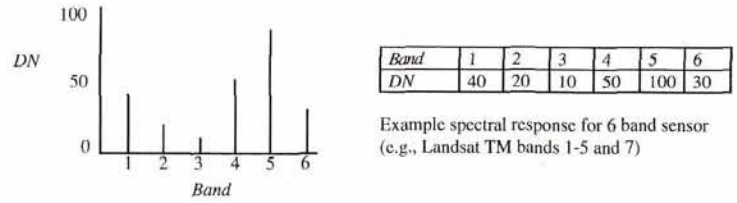
$$\phi(i,j,n,n') = \begin{cases} 1, & y(i,j,n) > y(i,j,n') \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

where $n = 1, 2, \ldots, N-1$ and $n' = n+1, n+2, \ldots, N$. In effect, the spectral shape representation converts the original multispectral image into an image of binary features that are used for classification. This set of features is equivalent to ranking the spectral bands in decreasing (or increasing) order by value; e.g.,

Pacific-Sierra Research Corporation, 1400 Key Boulevard, Suite 700, Arlington, VA 22209 (markc@psrw.com).

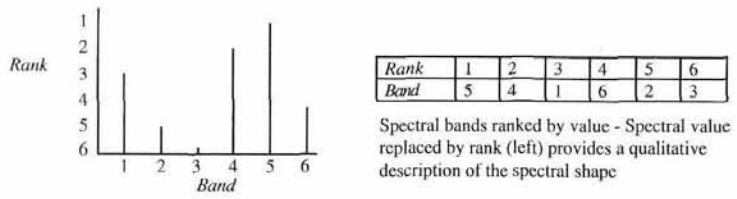
Photogrammetric Engineering & Remote Sensing, Vol. 64, No. 9, September 1998, pp. 905–913.

0099-1112/98/6409–905$3.00/0



Band	1	2	3	4	5	6
DN	40	20	10	50	100	30

Example spectral response for 6 band sensor (e.g., Landsat TM bands 1-5 and 7)



Rank	1	2	3	4	5	6
Band	5	4	1	6	2	3

Spectral bands ranked by value - Spectral value replaced by rank (left) provides a qualitative description of the spectral shape

Bands	1,2	1,3	1,4	1,5	1,6	2,3	2,4	2,5	2,6	3,4	3,5	3,6	4,5	4,6	5,6
Relative Value	1	1	0	0	1	1	0	0	0	0	0	0	0	1	1

Representation of spectral shape in terms of binary features

Figure 1. Example of spectral response and its spectral shape representation.

$$y(i,j,n_1) > y(i,j,n_2) \ldots \quad (2)$$

where n_k is the number of the kth largest band in value at a particular pixel, and using the rank-ordered band numbers as features. Figure 1 shows an example spectral response and its spectral shape representation, both in terms of binary features and rank-ordered band numbers. We choose the binary feature representation because it has a simple physical interpretation and can be compared using similarity measures such as the Hamming distance, as discussed later in the paper.

For N bands, there are $N \times (N-1) \ldots = N!$ possible ways to order the bands. Each ordering can be expressed by a unique combination of $N(N-1)/2$ binary features. Ordering bands by value using QUICKSORT (Sedgewick, 1983) requires approximately $N \log_2 N$ operations compared to $N(N-1)/2$ operations involved in comparing all distinct pairs of bands and is, thus, computationally more efficient as the number of bands increases. It should be clear, however, that because the number of features grows quadratically and the number of combinations grows factorially, the spectral shape representation is intended for multispectral sensors with a moderate number of bands such as Landsat TM and is not well-suited for sensors with a small number of bands like SPOT or for hyperspectral sensors.

The spectral shape representation effectively segments the spectral measurement space into $N!$ disjoint regions. Each spectral shape corresponds to a wedge-shaped region in this space, all of which touch the origin. Figure 2 illustrates the shape of the region in three-space corresponding to the set of features $\Phi = \{1,1,1\}$, i.e., the region where, for a sensor with three bands, $y(1) > y(2)$, $y(1) > y(3)$, and $y(2) > y(3)$. We show the construction of the region in stages for clarity. Without loss of generality, the original spectral response can be uniformly scaled to fit into the unit cube shown in Figure 2a. In Figure 2b, the cube is split in half along the diagonal $y(1) = y(2)$, and the half-space $y(1) > y(2)$ is retained. In Figure 2c, the previous region is split along the diagonal $y(1) = y(3)$, and the part where $y(1) > y(2)$ and $y(1) > y(3)$ is retained. Finally, in Figure 2d, the previous region is split along the diagonal $y(2) = y(3)$, and the region where $y(1) > y(2)$, $y(1) > y(3)$, and $y(2) > y(3)$ is retained. By integration, it can be shown that the volume of this or any of the other five wedge-shaped regions in three-space is $1/6$ $(= 1/N!)$.

The use of binary features provides a basis for comparing spectral shapes. The similarity between two spectral shapes u and v is defined to be equal to the number of binary features that are different and is given by the Hamming distance

$$d(u,v) = \sum_{n=1}^{N-1} \sum_{n'=n+1}^{N} \phi(u,n,n') \oplus \phi(v,n,n') \quad (3)$$

where $\oplus$ denotes exclusive-or. The Hamming distance satisfies the following properties:

(i) $d(u,v) \geq 0$

(ii) $d(u,v) = 0$, if $u = v$

(iii) $d(u,v) = d(v,u)$

(iv) $d(u,v) + d(v,w) \geq d(u,w)$

and is thus a valid metric.

Table 1 shows the Hamming distance between all six spectral shapes for $N = 3$ bands. Each spectral shape is shown along the top row and left column along with its set of binary features $\Phi = \{\phi_{12}, \phi_{13}, \phi_{23}\}$. As seen in Table 1, spectra that have a similar shape have a smaller Hamming distance between them than do spectra with different shapes.

Application to Clustering

Clustering is performed to identify regions with similar spectral properties in an image. Algorithms such as K-means and ISODATA (Tou and Gonzales, 1974) partition the underlying spectral measurement space into clusters where the parameters of the clusters are adjusted iteratively to minimize some objective function, typically the total squared error between the clusters and the data. Because the spectral shape representation effectively segments the measurement space, it functions like a clustering algorithm. An important difference between spectral shapes and clusters is that spectral shapes correspond to regions with fixed boundaries in the

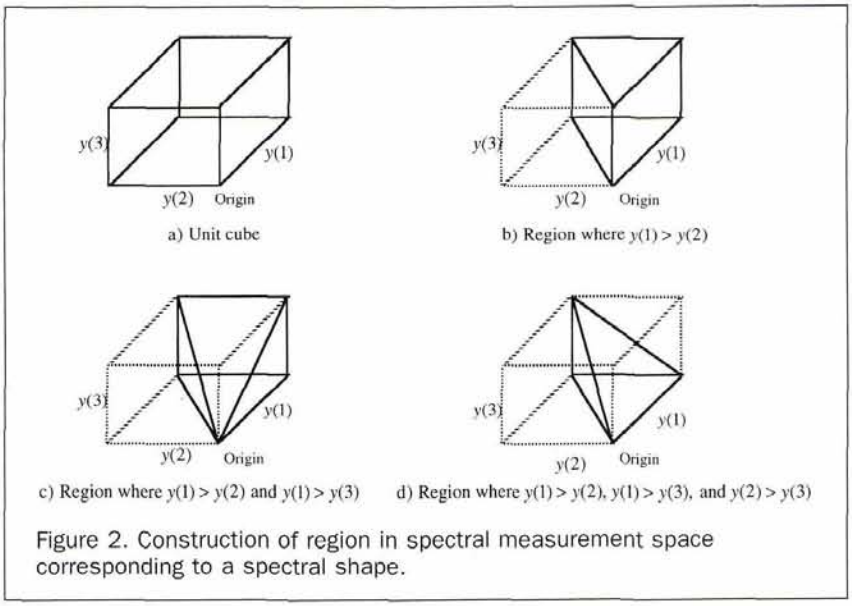


Figure 2. Construction of region in spectral measurement space corresponding to a spectral shape.

measurement space while the boundaries between clusters depend on the distribution of the data.

Spectral shapes and K-means clusters were extracted from a 226- by 236-pixel image over Gordonsville, Virginia (Figure 3). Figure 4 compares the structure of clusters in feature space extracted by the K-means algorithm to that of the fixed spectral shape regions for Landsat TM data. Instead of attempting to visualize the underlying six-dimensional space for Landsat TM (bands 1 to 5 and 7), the K-means clusters and spectral shape regions have been projected down into the two-dimensional space spanned by the first two tasseled cap features (Crist and Cicone, 1984). The boundaries between clusters and spectral shape regions are shown in the figure. Fifty-six spectral shapes were found in the Landsat TM imagery over the region shown in Figure 3. For comparison purposes, the same number of clusters were extracted using the K-means algorithm.

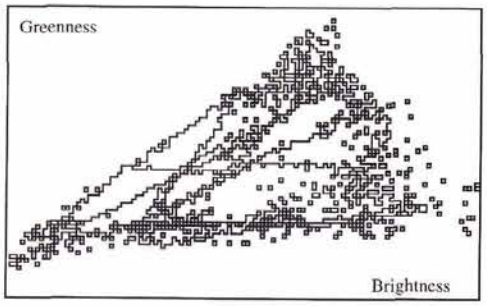
Comparing the two plots in Figure 4, K-means clusters are more evenly distributed in space than are spectral shapes (K-means attempts to minimize the squared error between the data and the clusters) and are more compact (K-means uses the Euclidean distance, which encourages the formation of compact regions). The boundaries between spectral shape regions, on the other hand, radiate out from the origin and have more of a tapered appearance. It has been observed that multispectral data tends to cluster in teardrop rather than elliptical distributions due to shading, shadowing, and pixel mixing (Craig, 1994; Crist and Cicone, 1984). The boundaries between spectral shapes also have this same general appearance.

Finally, it is noted that the amount of computation required to compute the spectral shape representation is significantly less than algorithms like K-means which have a complexity of $\approx O(KLN)$ floating point operations per pixel where L is the number of iterations. As a result, the spectral shape algorithm can be used to segment an entire data set

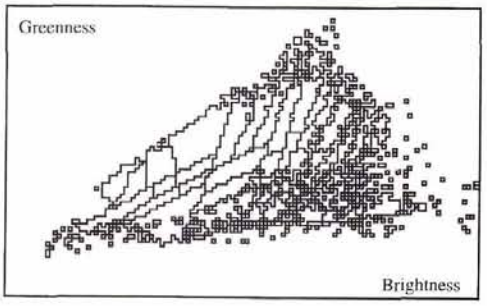
TABLE 1. HAMMING DISTANCE BETWEEN SPECTRAL SHAPES FOR $N = 3$ BANDS

	{1,1,1}	{1,1,0}	{0,1,1}	{0,0,1}	{1,0,0}	{0,0,0}
{1,1,1}	0	1	1	2	2	3
{1,1,0}	1	0	2	3	1	2
{0,1,1}	1	2	0	1	3	2
{0,0,1}	2	3	1	0	2	1
{1,0,0}	2	1	3	2	0	1
{0,0,0}	3	2	2	1	1	0

Figure 3. Landsat TM Band 3 over Gordonsville site.



(a)



(b)

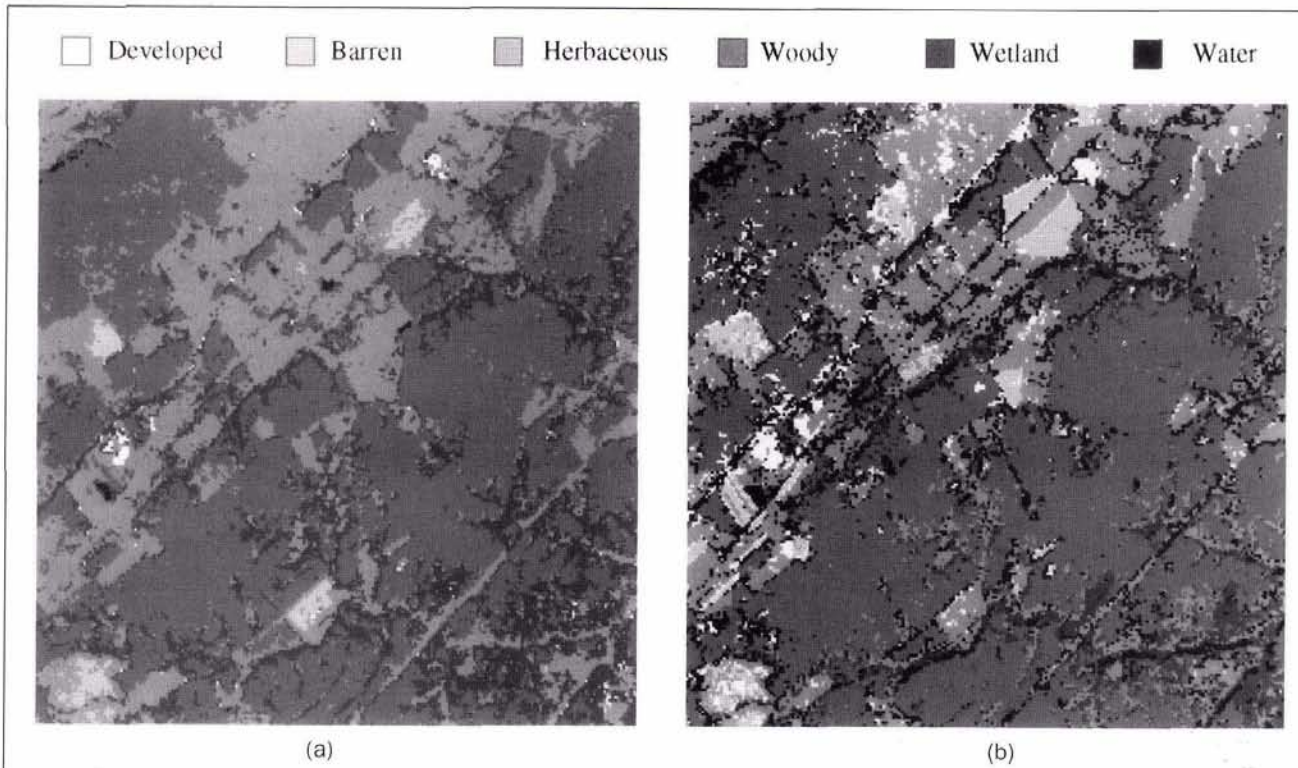
Figure 4. Comparison of spectral shape and K-means cluster boundaries in tasseled cap space. (a) Boundaries between spectral shape regions. (b) Boundaries between K-means clusters.

without having to first reduce its size, e.g., by sub-sampling, as is often required with K-means and similar clustering algorithms.

Multispectral Classification

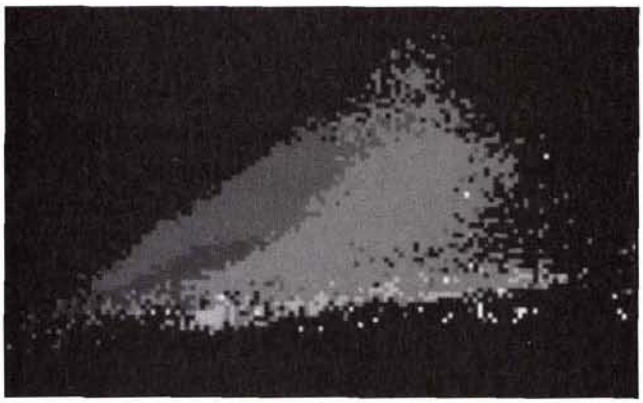
The spectral shape approach can be used for multispectral classification by segmenting an image into spectral shapes as described in the previous section, assigning a class to a subset of spectral shapes (e.g., those within a training area), and classifying the remaining spectral shapes according to their Hamming distance from those that have already been assigned a class. Figure 5 compares spectral shape and maximum-likelihood classifications over the Gordonsville site shown in Figure 3. The spectral shape classifier was trained by visually assigning a class to a subset of the spectral shapes. Remaining spectral shapes were classified by assigning the class of the spectral shape with the smallest Hamming distance. A maximum-likelihood classifier was trained by first clustering the image using K-means as described in the previous section. Classes were assigned by visually assigning a class to a subset of the K-means clusters. Mean vectors and covariance matrices were then computed from the clusters and used to classify the full image. (A more detailed discussion of training and accuracy assessment is contained in the next section.)

Six general surface categories are depicted in Figure 5: built-up, barren, herbaceous (grassland and agriculture), woody, wetland, and open water. In general, the spectral shape classification appears to contain somewhat less detail than the maximum-likelihood classification, although the latter appears to confuse shadows and water and to contain more built-up pixels, many of which are scattered within woody and agricultural areas.

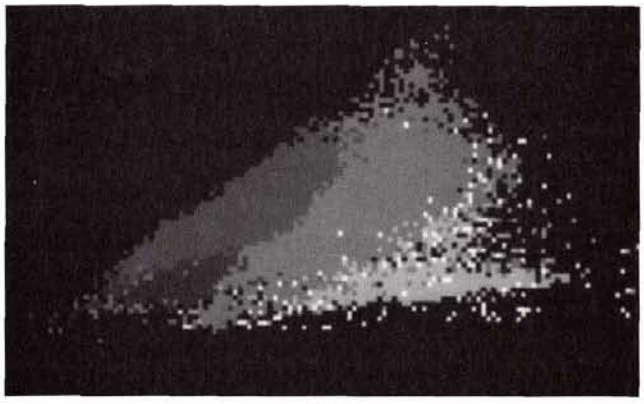


(a) (b)

Figure 5. Visual comparison of spectral shape and maximum-likelihood classification results. (a) Spectral shape classification. (b) Maximum-likelihood classificaton.



(a)



(b)

Figure 6. Spectral shape and maximum-likelihood classifications plotted in tasseled cap space. (a) Spectral shape classification. (b) Maximum-likelihood classification.

Figure 6 plots the classification results contained in Figure 5 into tasseled cap space as a function of their brightness and greenness values (the shades of gray match those used in Figure 5). Even though different classifiers were used, the class boundaries in the two distributions are similar. The boundaries between spectral classes in Figure 6 are more like the boundaries between spectral shapes (Figure 4a) than the boundaries between K-means clusters (Figure 4b). Over-clustering and merging K-means clusters into classes appears to create decision regions that are very similar in appearance to the boundaries between spectral shape regions.

Software Implementation

Figure 7 depicts the software architecture of the spectral shape classification system. The system contains four major functions:

- *Compute Spectral Shape Representation.* Segments the input Landsat TM image into spectral shapes. The spatial extent of each spectral shape is identified by a unique value in the label image. A file lists the relative frequency of each label along with the binary feature vector describing the corresponding spectral shape. (In the present implementation, the largest 255 spectral shapes are extracted, i.e., those accounting for most of the image area. When there are more than 255 spectral shapes in an image, pixels whose spectral shapes are not retained are assigned the nearest spectral shape among the 255 spectral shapes that were retained. Typically, for a full Landsat scene, this amounts to only a few pixels total. Most TM scenes processed to date contain only about 100 to 200 spectral shapes out of a possible 6! = 720.)
- *Training.* Associates classes derived from ground truth data with spectral shapes. The output of the training process is a classification file that specifies the class most frequently associated with each spectral shape in the training set along with its relative frequency.
- *Merge Classification Files.* Combines data from multiple training areas into a single classification file. Spectral shapes that have been assigned more than one class in different training areas are assigned the class that has been assigned most frequently to the spectral shape overall.
- *Minimum Hamming Distance Classifier.* Assigns the class associated with spectral shapes in the classification file to spectral shapes present in the image being classified. Spectral shapes not in the classification file can be assigned the nearest class.

Table 2a is a classification file computed by training over one of the scenes (Chickahominy) discussed in the next section. It lists all spectral shapes (in base 10), their assigned class, and probability (fraction). Table 2b is part of the classification file for another scene (Gordonsville). Table 2c is the result of merging the two files. The class assignments for spectral shape 1760 are underlined in these tables. In Chickahominy, this spectral shape is assigned class 13 (emergent wetland) while in Gordonsville it is assigned class 14 (woody wetland). The conflict is resolved during merging by assigning the spectral shape class that occurs more frequently, here by assigning class 14 because it occurs more often in Gordonsville (0.055) than class 13 does in Chickahominy (0.027) for the same spectral shape.

The spectral shape classifier is currently implemented in C and is able to classify a full Landsat TM scene in about an hour on a SUN SPARC 10 workstation with 128 Megabytes of RAM.

Experimental Results

A series of experiments was performed to assess the accuracy of the spectral shape classifier over a full Landsat TM scene (Path 15 Row 34) acquired on 23 October 1993. Five sites approximately 25 km² in size within the scene were selected for study (Figure 8). Two of the sites — Fort Eustis and Chickahominy — are wetland sites containing mostly open water, grassy and forested wetlands, and forests. The other three sites — Fort A.P. Hill, Gordonsville, and Prince Edward State Forest — are upland sites containing forested and agricultural areas. With the exception of large built-up areas, the five sites are fairly representative of the overall scene.

Site selection was based on the availability of M7 imagery, which was used as image truth for classifier development and testing. The M7 was acquired within several days of the Landsat imagery (M7 is a 12-channel multispectral im-

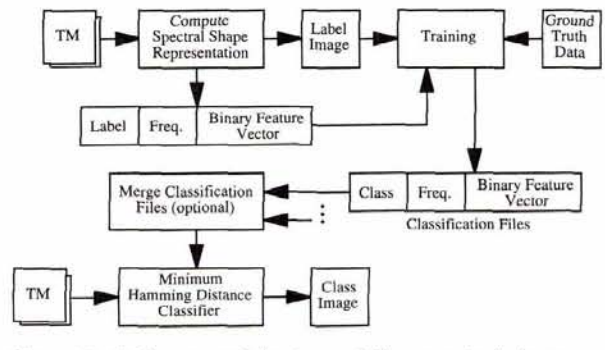


Figure 7. Software architecture of the spectral shape classification system.

TABLE 2a. CLASSIFICATION FILE FOR CHICKAHOMINY

Spectral Shape	Class	Probability
0	15	2.78E-02
1	15	7.20E-05
32	15	2.62E-03
96	15	4.32E-04
224	15	2.76E-03
512	15	8.88E-04
576	15	9.60E-05
736	15	3.60E-04
1536	15	2.45E-03
1600	15	8.88E-04
1728	7	1.18E-01
1760	13	2.07E-02
2016	11	5.93E-03
2020	11	3.84E-04
2028	11	5.47E-03
3776	7	5.68E-01
3780	7	2.40E-05
4032	7	4.79E-02
4036	7	4.32E-04
4076	11	7.44E-04
4096	12	2.64E-04
4128	15	4.80E-05
4256	15	7.20E-05
4320	15	6.96E-04
5120	14	9.60E-05
5248	12	4.80E-05
5344	12	9.36E-04
5632	12	1.44E-04
5760	12	9.60E-05
5824	7	1.21E-02
5856	13	4.15E-02
5864	13	9.60E-04
5868	11	1.92E-04
6112	13	1.09E-02
6120	11	8.76E-03
6124	5	3.64E-02
6126	5	4.49E-03
7872	7	3.96E-02
8128	7	2.31E-02

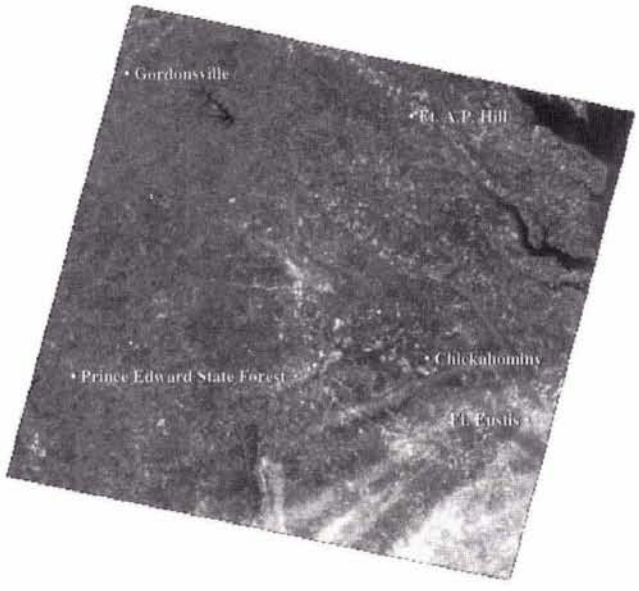


Figure 8. Landsat TM (path 15 row 34) acquired on 23 October 1993 (Band 3) showing site locations.

TABLE 2b. PART OF CLASSIFICATION FILE FOR GORDONSVILLE

Spectral Shape	Class	Probability
0	15	2.62E-04
32	15	1.90E-05
96	15	1.90E-05
224	15	1.69E-04
512	15	3.70E-05
576	15	1.90E-05
736	15	9.40E-05
1536	15	1.50E-04
1600	15	3.37E-04
1728	8	8.92E-02
1732	7	5.60E-05
1760	14	5.51E-02
1764	14	1.18E-03
1772	14	1.39E-03
1774	10	1.90E-05
2016	7	4.06E-02
2020	6	5.08E-03
2028	6	2.59E-02
2030	10	2.81E-04
3776	7	1.60E-01

TABLE 2c. PORTION OF MERGED CLASSIFICATION FILE

Spectral Shape	Class	Probability
0	15	1.65E-02
1	15	4.20E-05
32	15	1.55E-03
96	15	2.65E-04
224	15	1.72E-03
512	15	5.43E-04
576	15	6.80E-05
736	15	2.67E-04
1536	15	1.53E-03
1600	15	7.20E-04
1728	7	6.95E-02
1732	7	3.30E-05
1760	14	3.24E-02
1764	14	6.94E-04
1772	14	8.15E-04
1774	10	1.10E-05
2016	7	2.39E-02
2020	6	2.99E-03
2028	6	1.52E-02
2030	10	1.65E-04

aging system with an instantaneous field of view of 2 mrad (Slater, 1985)). The effective ground resolution of the imagery used in this study was about 5 metre/pixel.) Over each site, the 25-metre Landsat TM data were resampled to 5 metres using nearest-neighbor interpolation. M7 imagery was then registered to the TM.

The two-level classification scheme shown in Table 3, derived from a land-cover classification system under development by Bara (personal communication, 1993), was used in our evaluation. This classification scheme may not be totally exhaustive in general, but was suitable for the area considered in this study.

Our evaluation was based on comparing the overall accuracy (fraction correct) of the spectral shape classifier to that of a maximum-likelihood classifier. Over each site, both classifiers were trained using the M7 imagery as truth. For the spectral shape classifier, spectral shapes were first extracted over the site. A subset of the spectral shapes were assigned a class visually by an image analyst using the M7 imagery as reference. The remaining classes were then as-

TABLE 3. DEFINITION OF LAND-COVER CLASSES

Level 1 Class	Criteria	Level 2 Class	Criteria
Developed	> 50% man-made	High Intensity	> 80% man-made
		Low Intensity	50-80% man-made
Herbaceous Land	> 50% herbaceous	Cropland	Managed
		Grassland	Unmanaged
Woody	> 50% woody	Deciduous	> 67% deciduous
		Evergreen	> 67% evergreen
		Mixed	
Barren	< 50% vegetated		
Wetland	Anderson Level 1	Shore	< 50% vegetated
		Emergent	> 50% herbaceous
		Woody	> 50% woody
Open Water	Anderson Level 1		

signed the class of the nearest spectral shape previously assigned a class based on the Hamming distance. The classification file was retained for use in classifying other sites.

For the maximum-likelihood classifier, the imagery was first clustered using the K-means algorithm into 30 clusters. This number of clusters was judged by an image analyst to be adequate to separate the classes of interest within each of the sites. A subset of the clusters was assigned a class, again using the M7 imagery as a reference. The mean vectors and covariance matrices were then computed and used to classify the entire image. The means and covariances were then retained for use in classifying other sites.

In order to obtain an unbiased estimate of the relative performance of the two classifiers, a random sampling scheme was used to generate 60 sample points within each of the five sites. The points were assigned a class using the M7 imagery by a second image analyst. The points were then used to measure the accuracies of the two classifiers at Levels 1 and 2 as defined in Table 3.

Tables 4 to 7 summarize the classification results over the five study areas: Fort A.P. Hill (APH), Chickahominy (CHI), Fort Eustis (FTE), Gordonsville (GOR), and Prince Edwards State Forest (PED). Tables 4 and 5 give the spectral shape classification results at Levels 1 and 2, respectively. Each entry gives an overall classification accuracy (fraction correct) for the spectral shape classifier developed over the training site (row) and measured over the evaluation site (column); i.e., the classification file derived from the training site was used to classify the imagery of the evaluation site. Tables 6 and 7 are maximum-likelihood classification results at Levels 1 and 2, respectively, provided for comparison purposes. Here the training statistics (class means and covariance matrices) from the training site were used to classify the imagery of the evaluation site. The 95 percent confidence intervals based on 60 points are the values in the tables ± 10 percent, approximately.

First, we consider the overall accuracy of the two classifiers trained and evaluated over the same sites (these are the entries along the main diagonal of the tables). The accuracies of the spectral shape classifier averaged over the five sites were 0.84 and 0.63 (Levels 1 and 2, respectively). The corresponding accuracies of the maximum-likelihood classifier averaged over the five sites were 0.76 and 0.48. For 5 by 60 or 300 points, the 95 percent confidence intervals are the above values ± 5 percent, approximately. The performance of the spectral shape classifier is thus not significantly different from that of the maximum-likelihood classifier at Level 1, but does appear to be significantly better at Level 2. However, if we leave out the FTE site which has thin cloud cover, neither Level 1 nor Level 2 results are significantly different.

Next, we compare the overall accuracies of the two classifiers trained over one study area but evaluated over a different area (these are the off-diagonal entries in the tables). This will provide an indication of the signature extendibility of the classifiers over the scene. For the five sites, there are 20 possible combinations of different training and evaluation sites. The accuracies of the spectral shape classifier averaged over the 20 combinations were 0.79 and 0.57 (Levels 1 and 2, respectively). The corresponding accuracies of the maximum-likelihood classifier averaged over the 20 combinations were 0.57 and 0.31. The performance of the spectral shape classifier is thus significantly better than that of the maximum-likelihood classifier at Levels 1 and 2. The spectral shape classification accuracies averaged over the 12 combinations of different training and evaluation sites that exclude FTE are 0.81 and 0.58 (Levels 1 and 2, respectively). The corresponding maximum-likelihood accuracies are 0.76 and 0.41. If we leave out the FTE site, the Level 1 results are not significantly different but the spectral shape classifier accuracy is significantly better than that of the maximum-likelihood classifier at Level 2.

Within wetland sites, the Level 1 and 2 accuracies of the spectral shape classifier — 0.90 and 0.74 — are significantly better than the corresponding accuracies of the maximum-likelihood classifier — 0.63 and 0.35. This is due largely to

TABLE 4. SPECTRAL SHAPE CLASSIFICATION RESULTS - LEVEL 1

	Evaluation Site				
Training Site	APH	CHI	FTE	GOR	PED
APH	0.800	0.903	0.560	0.938	0.904
CHI	0.775	0.900	0.917	0.645	0.754
FTE	0.655	0.833	0.957	0.633	0.672
GOR	0.775	0.916	0.902	0.725	0.852
PED	0.741	0.866	0.821	0.707	0.806

TABLE 5. SPECTRAL SHAPE CLASSIFICATION RESULTS - LEVEL 2

	Evaluation Site				
Training Site	APH	CHI	FTE	GOR	PED
APH	0.523	0.587	0.360	0.733	0.606
CHI	0.606	0.672	0.835	0.500	0.539
FTE	0.491	0.606	0.847	0.419	0.428
GOR	0.557	0.573	0.680	0.531	0.619
PED	0.409	0.590	0.684	0.621	0.571

TABLE 6. MAXIMUM-LIKELIHOOD CLASSIFICATION RESULTS - LEVEL 1

	Evaluation Site				
Training Site	APH	CHI	FTE	GOR	PED
APH	0.719	0.960	0.074	0.700	0.784
CHI	0.540	0.919	0.900	0.604	0.666
FTE	0.063	0.109	0.611	0.126	0.109
GOR	0.829	0.905	0.339	0.716	0.879
PED	0.735	0.765	0.447	0.800	0.819

TABLE 7. MAXIMUM-LIKELIHOOD CLASSIFICATION RESULTS - LEVEL 2

	Evaluation Site				
Training Site	APH	CHI	FTE	GOR	PED
APH	0.515	0.539	0.050	0.524	0.475
CHI	0.366	0.634	0.463	0.311	0.315
FTE	0.029	0.047	0.253	0.047	0.081
GOR	0.347	0.500	0.279	0.476	0.475
PED	0.149	0.415	0.300	0.500	0.508

the FTE site, which contains thin clouds that adversely affect the maximum-likelihood classifier. In the uplands sites, the accuracy of the spectral shape classifier is significantly better than that of the maximum-likelihood classifier at Level 2 (0.57 compared to 0.441) but is comparable at Level 1 (0.81 compared to 0.76). The average accuracy of the spectral shape classifier within wetland and upland sites is 0.84 (Level 1). Between wetland and upland sites, it falls slightly to 0.76 (Level 1). (The within wetland and upland cases involve all combinations in which the training and evaluation site are either both wetland or both upland. The between wetland and upland cases involve all combinations in which the training site is wetland and the evaluation site is upland, or vice versa.) The average accuracy of the maximum-likelihood classifier within wetland and upland sites is 0.73 (Level 1), but between wetland and upland sites falls dramatically to 0.47 (Level 1).

We then constructed a spectral shape classifier for the full scene by merging classification files from each study area. In cases where spectral shapes were assigned different classes in different sites, the class with the highest probability across sites was selected and assigned to the spectral shape globally. The resultant classification file was then used to classify the full scene. The Level 1 classification results are summarized by class in Table 8 based on 60 points per site for a total of 300 points. The overall accuracies of the full-scene spectral shape classification were 0.85 and 0.65 at Levels 1 and 2, respectively.

To gain some further insight into the classification performance, Level 2 results for woody and wetlands categories are shown in Tables 9a and 9b. (In these tables, the category "other" represents all other Level 2 categories combined.) The Level 1 accuracy for woody areas is high because most of the errors are between deciduous and evergreen categories. On the other hand, the Level 1 accuracy for wetlands is low because a large fraction of wetlands have been misclassified as other (woody and herbaceous). The spectral shape classifier appears to have some trouble in discriminating between evergreen and coniferous trees but is effective in discriminating trees from other categories. It also appears to have trouble in discriminating woody wetlands from other woody areas.

Discussion

Phenological and environmental variations within a scene limit the extendibility of a classifier. From general ground surveys conducted at the time, trees in the upland areas had already begun to shed their leaves while those in the wetland sites had just begun to change color. Based on the marked decrease in Level 1 classification accuracy of the maximum-likelihood classifier from 0.73 within wetland and upland sites down to 0.47 between wetland and upland sites compared to the relatively stable accuracy of the spectral shape classifier (0.84 down only to 0.76), it appears that the spectral shape classifier is less sensitive to the phenological differences between upland and wetland sites at this particular time of year in this scene. Analysis of the confusion matrices, however, indicates some difficulty in discriminating between Level 2 vegetation categories. It is conjectured that what the spectral shape representation looses in discriminating power (e.g., its ability to separate Level 2 vegetation categories), it gains in signature extendibility across certain kinds of phenological variations within the scene.

Earlier, we noted that the accuracy of the spectral shape classifier was significantly better than that of the maximum-likelihood classifier within the FTE site and between FTE and other sites. As seen in Figure 8, cloud cover south of Richmond partially obscures this site. Thin clouds tend to behave as non-selective scatterers; i.e., the wavelength dependence of the scattering is $\approx \lambda^0$ (Schanda, 1986). Scattering by thin clouds (as well as dust and fog) can be modeled by the linear relation, $y(n) = ax(n) + c$, where $x(n)$ is proportional to the surface radiance in the nth band, and the constants a and c account for a uniform decrease in path transmittance and increase in path radiance across all bands. Clearly, the spectral shape representation is not affected by a constant gain and offset across bands because the features

$$\phi(n,n') = \begin{cases} 1, & ax(n) + c > ax(n') + c \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

do not depend on a or c. It should be noted, however, that this invariance does not hold for areas affected by smoke and haze where the scattering is wavelength dependent, nor for areas obscured by thick clouds.

The full-scene spectral shape classifier combines classification files derived from several training sites that are representative of the spectral diversity of the full scene and combines them into a single file that is applied to the full data set without stratification. The overall accuracies of the full-scene spectral shape classification were 0.86 and 0.65 (Levels 1 and 2, respectively). This compares well with the average spectral shape classification accuracies computed over all sites. We noted above that the average accuracy of the maximum-likelihood classifier across wetland and upland sites was considerably lower than the accuracy within those sites. If we assume that the scene contains two strata — wetlands and upland areas — the accuracies of a stratified maximum-likelihood classifier would be 0.73 and 0.41 (Levels 1 and 2). Of course, this presupposes that it is possible to accurately delineate the boundaries between strata in order to develop such a classifier in the first place.

TABLE 8. FULL-SCENE CLASSIFICATION RESULTS (LEVEL 1)

	Ground Truth Classes:					
Classified as:	Developed	Herbaceous	Woody	Barren	Wetland	Water
Developed	0	0	0	1	0	0
Herbaceous	1	32	13	0	3	0
Woody	3	9	192	0	4	0
Barren	0	1	0	0	0	0
Wetland	6	0	2	0	15	1
Water	0	0	0	0	2	15

TABLE 9a. LEVEL 2 CLASSIFICATION RESULTS FOR WOODY AREAS

	Ground Truth Classes:		
Classified as:	Deciduous	Evergreen	Mixed
Deciduous	136	20	0
Evergreen	21	15	0
Mixed	0	0	0
Other	14	1	0

TABLE 9b. LEVEL 2 CLASSIFICATION RESULTS FOR WETLANDS

	Ground Truth Classes:		
Classified as:	Shore	Emergent	Woody
Shore	0	0	0
Emergent	0	14	1
Woody	0	0	0
Other	0	5	4

Summary

A new method of classifying multispectral imagery based on a set of binary features that represent the relative values between spectral bands was described, and its performance over a full Landsat scene was evaluated. The overall classification accuracy was found to be comparable to that of a maximum-likelihood classifier over individual sites. However, between sites, the spectral shape classifier out-performed the maximum-likelihood classifier. A full-scene spectral shape classifier was developed by combining classification files from the five sites into a single classification file. The accuracy of the resultant classifier tested over the five sites was significantly better than the accuracy of a stratified maximum-likelihood classifier, assuming wetlands and uplands strata. Preliminary results suggest that the spectral shape representation can provide a greater degree of signature extendibility but at the expense of reduced discrimination at Level 2.

Additional work in several areas is either underway or planned. Further testing over other areas is being performed in order to obtain a more representative estimate of the accuracy of the spectral shape classifier. Other methods for combining classification files derived from individual training sites and applying them to the full scene are being explored that retain regional class-cluster relationships. Finally, experiments are planned to measure the extent to which the spectral shape classifier is sensitive to topographic effects.

Acknowledgments

The author wishes to thank Greg Heberle and Ernie Carroll for registering the Landsat TM and M7 imagery, and for assisting in the accuracy assessment experiments.

References

Carlotto, M.J., 1990. Towards the unsupervised interpretation of outdoor imagery, *J. Visual Communication and Image Representation*, 1(2):158–175.

Carlotto, M.J., V.T. Tom, P.W. Baim, and R.A. Upton, 1984. Knowledge-based multispectral image classification, *Proceedings SPIE*, San Diego, California, 504:47–53.

Carlotto, M.J., and V.T. Tom, 1985. Surface material classification based on spectral shape, *Proceedings of IGARSS*, Amherst, Massachusetts, pp. 703–708.

Craig, M.D., 1994. Minimum volume transforms for remotely sensed data, *IEEE Trans. Geoscience and Remote Sensing*, 32(3):542–552.

Crist, E.P., and R.C. Cicone, 1984. A physically-based transformation of thematic mapper data - The TM Tasseled Cap, *IEEE Trans. Geoscience and Remote Sensing*, 22(3):256–263.

Henderson, R.G., 1976. Signature extension using the MASC algorithm, *IEEE Trans. Geoscience Electronics*, GE-14(1):34–37.

Sedgewick, R., 1983. *Algorithms*, Addison-Wesley, Reading, Massachusetts, pp. 103–108.

Salter, P.N., 1985. Survey of multispectral image systems for earth observations, *Remote Sensing of Environment*, 17:85–102.

Schanda, E., 1986. *Physical Fundamentals of Remote Sensing*, Springer-Verlag, Berlin Heidelberg, pp. 116–117.

Todd, W.J., D.G. Gehring, and J.F. Haman, 1980. Landsat wildland mapping accuracy, *Photogrammetric Engineering & Remote Sensing*, 46(4):509–520.

Tou, J.T., and R.C. Gonzalez, 1974. *Pattern Recognition Principles*, Addison-Wesley, Reading, Massachusetts, pp. 94–104.

(Received 07 July 1995; revised and accepted 19 August 1997; revised 19 February 1998)

Forthcoming Articles

John E. Anderson, **Spectral Reflectance and Detection of Iron-Oxide Precipitates Associated with Acidic Mine Drainage.**

Michel Boulianne and *Clément Nolette*, **Virtual Reality Applied to User Interfaces for Digital Photogrammetric Workstations.**

Daniel G. Brown and *Alan F. Arbogast*, **Digital Photogrammetric Change Analysis as Applied to Active Coastal Dunes in Michigan.**

Giles M. Foody, **The Continuum of Classification Fuzziness in Thematic Mapping.**

A. Goller, M. Gelautz, and *F. Leberl*, **Parallel Image Processing Applied to Radar Shape-from-Shading.**

Peter L. Guth, **Contour Line "Ghosts" in USGS Level-2 DEMs.**

Graham Horgan, **Wavelets for SAR Image Smoothing.**

Daniel Lebel and *Ronald Da Roza*, **An Innovative Approach Using Digital Photogrammetry to Map Geology in the Porcupine Hills, Southern Alberta, Canada.**

Rongxing Li, **Potential of High-Resolution Satellite Imagery for National Mapping Products.**

C.P. Lo and *Lee J. Watson*, **The Influence of Geographic Sampling Methods on Vegetation Map Accuracy Evaluation in a Swampy Environment.**

Stuart R. Phinn, Douglas A. Stow, and *David Van Mouwerik*, **Remotely Sensed Estimates of Vegetation Structural Characteristics in Restored Wetlands, Southern California.**

John E. Pinder III, **Indications of Relative Drought Stress in Longleaf Pine from Thematic Mapper Data.**

Ken Rutchey and *Les Vilchek*, **Air Photo-Interpretation and Satellite Imagery Analysis Techniques for Mapping Cattail Coverage in a Northern Everglades Impoundment.**

Liora Sahar and *Amnon Krupnik*, **Semiautomatic Extraction of Building Outlines from Large-Scale Aerial Images.**

Terry L. Sohl, **Change Analysis in the United Arab Emirates: An Investigation of Techniques.**

Jaan-Rong Tsay and *Bernhard P. Wrobel*, **A New Algorithm for Surface Determination Based on Wavelets and Its Practical Application.**

Jung-Sup Um and *Robert Wright*, **The Analog-to-Digital Transition and Implications for Operational Use of Airborne Videography.**

David Wladis, **Automatic Lineament Detection Using Digital Elevation Models with Second Derivative Filters.**

Huang Youcai and *Robert M. Haralick*, **Testing Camera Calibration with Constraints.**