Shin-yi Hsu
Center for Scientific Study of Natural Resources
State University of New York
Binghamton, NY 13901

Texture-Tone Analysis for Automated Land-Use Mapping

A correct classification rate of 95 percent for the training set and 85 to 90 percent for the data property set was obtained with panchromatic images.

Abstract: *This paper summarizes the results of a computerized classification and mapping project utilizing a new texture measure sponsored by RADC, U.S. Air Force. The base data were digitized U-2 images. The texture feature extractor employs 3 by 3 and 5 by 5 pixels/windows referred to as Model I and Model II, respectively. For classification, only the center point of the pixel area is involved. In Model I, seventeen variables are extracted: (1) through (4), the four central moments; (5), the absolute deviation from the mean; (6), the contrast of the center point from its neighbor; (7), the mean brightness of the center point relative to the background; (8), the contrast between adjacent neighbors; (9), the squared value of (8); and (9) through (17), the mean area above and below three datum planes (50, 100, and 150). In Model II, an additional six wave-form variables scanned from x and y directions are employed: (1), the sum of the contrast values from peaks and troughs; (2), the sum of the peak positions from the origin; and (3), the number of peaks and troughs. All the 23 variables, except kurtosis, are proved to be statistically significant.*

The classifier employs linear discriminant functions based on the Mahalanobis **D²** *statistics derived from the generalized inverse of separate dispersion matrix for each group. In seven test areas in New York State, the preliminary results yield a correct classification rate of above 95 percent for the training sets, and 85 to 90 percent for the data property set, with five to ten land-use types.*

Introduction

In the last decade, there has been a trend toward remote sensing analysis, and for photo interpreters to employ more specialized and sophisticated imaging systems in pattern recognition. The traditional black-and-white aerial photos are gradually being phased out as part of modern remote sensing, as evidenced in the instructional course structure adopted by most of the colleges and universities in the United States. By convention, remote sensing is generally interpreted as the employment of photographic infrared, thermal infrared, multispectral, and side-looking radar images and the like in data processing. However, black-and-white and color photos still have a very important role to play in the field of land-use analysis, because they are more generally available and economical to the users. Moreover, the image processing techniques developed for black-and-white photos can be adopted by other imaging systems in a number of ways. First, compared to black-and-white photos, multispectral image data analysis can utilize the same texture-tone variables for all of the channels. For instance, if we have ten variables from the black-and-white photos, there will be 40 variables for analysis in the four channel multispectral system. Second, one can employ the ratio variables between channels as additional variables for the multispectral system. With four channels, there will be six possible ratio-sets, each set having the same number of variables used in analyzing the black-and-white photos. This paper is intended to develop an image processing technique with black-and-white photos from a texture analysis approach which should be applicable to other imaging systems. The support of Rome Air Development Center, U.S. Department of Air Force for this work under Contract F30602-76-6-0211 is gratefully acknowledged.

The Test Sites

The data set for this study was composed of eight scenes (four low-altitude and four high-altitude) from four test sites in the State of New York: Griffiss Air Force Base (GALA, GAHA), Verona (VPLA, PLHA), Stockbridge (SBLA, SBHA), and Utica (URLA, URHA). Their geographic locations, elevations, and flight heights are given in Table 1 (RADC, 1976a).

Using a vidicon digitization system, the resolutions of low- and high-altitude image data are 8.75 feet and 56.75 feet on the ground (per pixel), respectively. Stored on tape, each scene was composed of 256 by 256 pixels, with tonal densities ranging from 0 (black) to 255 (white).

For mapping purposes, seven terrain types were used although sub-classes such as Soil 1 and Soil 2 were often employed in pattern recognition with computerized processes. The definitions of these classes are given:

(1) Metal: metal roofs, oil tanks, and light colored cars.
(2) Pavements: cement roads, asphalt roads, and tarred surfaces such as paved roofs and parking lots.
(3) Water: deep and shallow water bodies
(4) Soil: bare soils and sands
(5) Cultivated fields: active and inactive farm lands.
(6) Vegetation: trees and bushes
(7) Composition: a mixture of several categories such as urbanized area.

It should be noted that terrain type other than those can be grouped into a "rejected" class, which may also represent "edges" between two classes.

Texture Analysis

In manual photo interpretation, texture means the apparent minute pattern of detail of a given area. It is ordinarily described by these terms: smooth, fine, rough, coarse, and the like. In digital data processing, texture means the spatial distributions of tones of the pixels; its attributes have to be specified by the investigator.

Texture analysis is a rather recent but rapid growing field of inquiry although its importance in visual perception was recognized by Gibson in 1950 (Gibson, 1950). Over the past 20 or more years, many texture measures have been proposed for characterizing and discriminating scenes. In a review of literature, Rosenfeld recognized that these approaches can be grouped into two broad categories: Fourier-based (power spectrum) features and statistical features (Rosenfeld, 1975).

In a more detailed classification, Haralick (1975) noted that there have been six basic approaches: auto-correlation functions (Kaizer, 1955), optical transforms (Lendaris & Stanley, 1970), digital transforms (Gramenopoulous, 1973; Hornung and Smith, 1973; Kirvida and Johnson, 1973), edgeness (Rosenfeld and Thurton, 1971) and related measures (Schachter *et al.*, 1977; Lev *et al.*, 1977), structural elements (Matheron, 1967; Serra, 1973) and spatial dependency probabilities (Haralick, 1970, 1973), and an extended method (Haralick, 1975). In general, Weszka and Rosenfeld (1975) concluded that statistical features perform much better

Table 1. The Data Set

Test Sites	Geographic Locations	Elevation (ft)	Flight Height (ft)
1. GALA	43° 14′N, 75° 25′W	515	15,500
GAHA			61,500
2. VPLA	43° 08′N, 75° 36′W	500	15,500
VPHA			60,500
3. SBLA	43° 02′N, 75° 39′W	1290	16,000
SBHA			60,500
4. URLA	43° 07′N, 75° 13′W	410	15,400
URHA			60,500

than grey tones co-occurrence, and Fourier-based features which is the poorest.

Recently texture analysis has also been approached from human perceptual point of view. Indeed, the human eyes coupled with the brain are very effective in analyzing imagery patterns although their data processing rate is slow with manual operations. While this system works empirically, the mechanism by which visual detection and recognition is achieved is still largely unknown, as noted by Barlow *et al.* (1972) and Julesz (1975).

Discrimination of textural pattern pursued by psychologists has been done by using mainly random dots (Pickett, 1967; Polit, 1976; Purks and Richards, 1977). Recently, Whitman Richards has been conducting experiments using a "generalized colorimetry" technique analogous to that used so successfully in studying human color vision. He concluded that most uniform textures can be simulated by three or four variables, provided that they contain the basic elemental token of the graphic display which automated image processing methods have not yet been able to achieve.

Recently Mitchell (1976) and Mitchell *et al.* (1977) proposed a new measure for texture classification based on the human visual system intuition where the important texture information is contained in the relative frequency of local extremes of various sizes of intensity. Thus, it is called a max-min descriptor.

A New Measure

During 1975 and 1976, U.S. Air Force/Rome Air Development Center (RADC) sponsored a study titled "Digital Image Processing Techniques for Automatic Terrain Classification for Generating Reference Maps from B/W Aerial Photography," conducted by Pattern Analysis & Recognition Corporation, Rome, N.Y. The task was carried out using RADC's image data processing system, called DICIFER, with a limited capability of texture analysis, since the parameters included only six measures: mean, standard deviation, range, median, high, and low. Because the correct classification rate was about 80 percent, the RADC personnel felt that another study was needed to improve the hit-rate by using a more powerful texture feature extractor. This led to a project, concluded by this author, using the same data set for the 1975-1976 study. A new texture measure was thus developed.

There are two essential elements in this texture measure: its variables are statistical features, which has been proved to be the best by Weszka and Rosenfeld (1975), and it is perceptually based. While the first element is obvious from the computational procedures, the second element needs empirical proof from psychophysical tests, which will be discussed below.

For automated crop-types identification, Hsu and Kedar extracted ten texture variables from the wave-form parameters of scan lines along both the x and y axes with digitized side-looking radar data as follows (Hsu and Kedar, 1972):

(1) total area above the datum,
(2) total area below the datum,
(3) sum of contrast values (from peaks to troughs),
(4) sum of peak positions from the origin, and
(5) sum of the number of peaks and troughs.

Based on one experiment, a hit-rate of 90 percent with respect to ground truth data was obtained.

Given this impressive and promising result from the 1972 experiment, a perceptual study was carried out to determine the degree of correlation between statistical discrimination using the ten-variable system and human visual discrimination utilizing density maps portrayed by grey-tone patterns. The experiment used four maps composed of 10 by 13 density cells. Using a normal distribution model, a discriminant analysis was carried out to measure the difference between pairs of maps in terms of Mahalanobis $\mathbf{D}^2$. In visual discrimination judgments, ten students were asked to judge the visual difference between pairs of maps according to a specific scale. Finally, a positive and significant coefficient of correlation of 0.98 (with raw data) was obtained between $\mathbf{D}^2$'s and the means of the perceived difference scores (Hsu, 1974). Later in 1975, the same kind of experiment was repeated utilizing color maps instead of grey-tone maps, yielding a correlation coefficient of 0.90.

These two experiments demonstrated that the outcome of a statistical discrimination using these ten variables in the classifier based on multivariate normal model is highly correlated with the outcome of human perceptual judgments.

To determine the applicability of this texture feature extractor coupled with a classifier for mapping detailed terrain types, a study sponsored by RADC was carried out by this author. Since the task required classification of pixels, the original ten-variable texture measure was modified to include summary statistics derived from either 3 by

3 or 5 by 5 pixels/windows, yielding a 17 variables system (Model I) and a 23 variables system (Model II), respectively. The difference between the 3 by 3 and the 5 by 5 design is in the number of texture variables as indicated in Table 2.

In Model I, the 17 variables are (1) through (4), the four central moments; (5), the absolute deviation from the mean; (6), the contrast of the center point from its neighbors; (7), the mean brightness of the center point; (8), the first neighbor contrast; (9), the squared value of (8); (10), the second neighbor contrast; (11), the sum of the squared value of (10); and (12) through (17), the mean area above and below the three datum planes having a tonal value of 50, 100, and 150, respectively. Note that these datum planes are derived from the digitization system using a scale of 0 for black and 255 for white. Although the values of the datum planes are arbitrarily determined, they do represent more or less the mean reflectance values of certain terrain types, such as vegetation, cultivated fields, and soil. Of course, if necessary, one may add another datum plane, say, the tonal value of 200.

In Model II, in addition to the above 17 variables, three measures were extracted to characterize the wave-form parameters of the scan lines obtained from both the x and y directions; thus, six variables are available for analysis. They are (1) sum of the contrast value from peaks to troughs, (2) sum of the distance of peak positions from the origin, and (3) sum of the number of peaks and troughs.

The reason for employing such a small window size as 3 by 3 or 5 by 5 is that the task of mapping detailed terrain types required the classification of individual pixels rather than a group of pixels or scenes. Moreover, with a larger window size for pixel classifications, the edge effect will be great, appearing as rejected pixels between two adjacent classes. In fact, from this author's experience, even with the 5 by 5 design, the edge effect is still substantial. This is the reason why the 3 by 3 system was utilized for generating the final decision maps and hit-rate analyses.

The Classifier

In order to classify an object into one of K types, one can employ a general discriminant analysis as described by Rao (1973). It is assumed that the spectral signature of the objects have density functions $P_i(Y), \ldots, P_k(Y)$, where $P_i(Y)$ is the density function for the objects in the i^{th} class. To classify an unknown object whose spectral signature is given as Y is to first compute

Table 2. The Texture-Tone Variables Model I

Code	Description or Computational Formula
1. MEAN	average (the four central moments)
2. STD	standard deviation (the four central moments)
3. SKEW	skewness (the four central moments)
4. KURT	kurtosis (the four central moments)
5. MDEVN	$\Sigma\lvert x_i - \bar{x}\rvert/n$, where x = tone value of individual pixels, $\bar{x}$ = mean
6. MPTCON	$\Sigma\lvert x_i - x_c\rvert/n$, where x_c = tone value of the center point
7. MPTREL	$\Sigma(x_c - x_i)/n$
8. MINCON	$\Sigma\lvert x_i - x_j\rvert/n$, i and j are adjacent pixels
9. MINSQR	$\Sigma(x_i - x_j)^2/n$
10. M2NCON	$\Sigma\lvert x_i - x_k\rvert/n$, x_c and x_k are second nearest neighbors
11. M2NSQR	$\Sigma(x_i - x_k)^2/n$
12. MADAT1	numerical calculation of mean area above datum 1 (50)
13. MADAT2	mean area above datum 2 (100)
14. MADAT3	mean area above datum 1 (50)
15. MBDAT1	mean area below datum 1 (50)
16. MBDAT2	mean area below datum 2 (100)
17. MBDAT3	mean area below datum 3 (150)
	Additional Variables—Model II
18. XCONT	(Distance from peaks to troughs) along x-axis
19. XPEAK	(Peak positions from the origin) along x-axis
20. XPANDT	(Number of peaks and troughs) along x-axis
21. YCONT	(Distance from peaks to troughs) along y-axis
22. YPEAK	(Peak positions from the origin) along y-axis
23. YPANDT	(Number of peaks and troughs) along y-axis

the numerical value of $P_i(Y)$ for each $i = 1, \ldots, k$, and then place the object into class i_0 for which $P_{i_0}(Y)$ is largest. In case that $P_{i_0}(Y)$'s are multivariate normal, this method leads to the usual discriminant function.

With respect to the remote sensing literature, the discriminant methods were reviewed by Nalepka (1970), and they were grouped into two broad categories: (1) maximum likelihood ratio decision rules based on the Bayesian formulation and (2) a class of linear discriminant functions.

In this study, the linear discriminant analysis approach was employed; and the maximum likelihood solution leads to a Mahalanobis classifier with

$$\mathbf{D}_i^2 = (\mathbf{Y} - \mathbf{U}_i)^T (\mathbf{Q}_i)^{-1} (\mathbf{Y} - \mathbf{U}_i) \qquad (1)$$

as a classification rule, i.e., assign $\mathbf{Y}$ to population i if $\mathbf{D}_i^2 = \min [\mathbf{D}_i^2, \ldots, \mathbf{D}_k^2]$ as given in Morrison (1976) and reviewed by Glick (1977). Here

- $\mathbf{Y}$ = spectral response (texture variables) of the unknown object,
- $\mathbf{U}_i$ = mean texture vector (centroid) of training set i, and
- $\mathbf{Q}_i$ = dispersion matrix of the training set i.

Conventionally, a pooled dispersion matrix ($\mathbf{Q}$) is used instead of individual dispersion matrices of each training set ($\mathbf{Q}_i$). The reasons may be that (1) computationally it is less time-consuming for using $\mathbf{Q}$; (2) $\mathbf{Q}_i$ tends to become singular when numerous texture-tone variables are involved; and (3) all of the methods for estimating the error probabilities associated with discriminant analyses (as I am aware of) are based on a pooled dispersion matrix (Glick, 1973; Lachenbruch and Mickey, 1968; McLachlan, 1976).

In fact, we have determined that on the average the rank deficiency in the 17-variable system is 5 with unpooled dispersion matrices. To solve this problem, the generalized inverse approach as given in Rao and Mitra (1970) was utilized. The above classifier becomes

$$\mathbf{D}_i^2 = (\mathbf{Y} - \mathbf{U}_i)^T (\mathbf{Q}_i)^{-} (\mathbf{Y} - \mathbf{U}_i) \qquad (2)$$

where $(-)$ (minus) stands for the generalized inverse.

Though the generalized inverse is not unique, any generalized inverse used in Equation 2 produces exactly the same classification.

With several experiments, it has been proved that the Mahalanobis classifier using the unpooled dispersion matrices (Equation 2) can improve the hit-rate of about 8 percentage points with the training set data. This will certainly translate into a substantial improvement in the final decision map.

To insure that pixels that do not actually belong to any of the training set categories are identified as "rejects," two probability values are also used in the classification process. They are (1) $P(G/x)$ or probability that the pixel with score x is a member of group G; and (2) $P(x/G)$ or probability that a member of group G would have a distance from G's centroid of x or more. The computational formulas for these two values are

$$P(G/x) = \frac{e^L}{\Sigma e^{-x_i}}, \text{ where } L \text{ is the largest } x_i, \text{ and} \qquad (3) \ \& \ (4)$$

$P(x/G)$ = the probability of $\mathbf{D}^2$ obtained from the chi-square table with the degree of freedom given by the rank of $\mathbf{Q}_i$ in Equation 2.

Whereas one can obtain a classification map with no rejects by setting both $P(G/x)$ and $P(x/G)$ equal to zero, the amount of rejects is produced by varying the combination of these two probability values. Since the value of $P(x/G)$ is much more sensitive that that of $P(G/x)$, one can employ only the former criterion (Equation 4). It is our experience that one needs to experiment with several values of $P(x/G)$ in order to obtain a reasonable decision map although one can set $P(x/G) = 0.01$ for the first trial.

Analysis of Training Sets

In general, there are two classes of classification procedures available to the users: supervised and unsupervised. Whereas the first method utilizes training sets determined by the analyst, the second method employs numerical clustering algorithms performed solely by the computer. In this study, a supervised approach is utilized.

Initially, the training sets were selected from a TV display of the image by the use of a joystick or cursor. The average sample size was about 200 pixels, ranging from 100 to 400 pixels depending on the size of scenes being investigated. To insure a high rate of correction classification, the validity of the training sets themselves should be examined first from a confusion matrix, obtained from a discriminant analysis on the training sets data. An example is given in Table 3.

In Table 3 the principal diagonal entries indicate the frequency of correct classifica-

TABLE 3. CONFUSION MATRIX (ACTUAL GROUP DOWN, PREDICTED GROUP ACROSS) FROM MODEL I

	Pavement	C Field 1	C Field 2	Vegetation	Metal	Soil 1	Soil 2	Edgepave
Pavement	300	0	0	0	0	0	0	0
C Field 1	0	442	8	0	0	0	0	0
C Field 2	0	0	248	2	0	0	0	0
Vegetation	0	0	0	100	0	0	0	0
Metal	0	0	0	0	90	0	0	0
Soil 1	0	0	0	0	0	41	2	2
Soil 2	0	0	0	0	0	1	60	3
Edgepave	0	0	0	0	0	0	3	171

Total correct classifications = 98.57%

tion, while the sum of the off-diagonal elements along a row is the omission error rate. If this error rate is too high, the original training sets having a high percentage of error should be replaced by new samples. This error-rate analysis should be repeated until a satisfactory hit-rate is obtained, say 90 percent.

The separation among group centroids also can be measured by the Mahalanobis $\mathbf{D}^2$. A matrix of $\mathbf{D}_{ij}^2$ among group centroids can be used to assess the significance of separation, as well as the desirability for new training sets. For instance, if the $\mathbf{D}_{ij}^2$ is so small that its significance level exceeds 0.01, group i and group j should be made into one category. On the other hand, one may have to increase the number of measures in the analysis, so that group i and group j can be separated further. Table 4 gives an example of the $\mathbf{D}^2$ matrix. It should be noted that the $\mathbf{D}^2$'s are obtained by using a separate dispersion matrices for each group and therefore the matrix is not symmetric.

As discussed earlier, there are 17 texture-tone variables in Model I with a 3 by 3 design and 23 variables in Model II with a 5 by 5 design. Using step-wise discriminant analysis procedures, 13 variables are determined significant in Model I and 19 variables in Model II, respectively. In general, the third and fourth order statistics are not significant, including skewness and kurtosis. Tables 3 and 4 are obtained from Model I, and the following Tables 5 and 6 are generated from Model II. From Tables 3 and 5, it can be concluded that the difference in the hit-rate is insignificant between Model I and Model II in this analysis.

GENERATION OF DECISION MAPS

Similar to classifying training sets data into one of the classes or a reject category, a decision may can be generated by using the pixels of the entire frame from which the sample data are selected. In this study, four decision maps were produced from each frame of 256 by 256 pixels: (1) Model I without a reject category, (2) Model I with rejects, (3) Model II without rejects, and (4) Model II with rejects.

At present, eight frames have been processed. From this experience, it can be concluded that, in general, Model I with 17 variables is sufficient for classifying land-use types. Since Model II is more powerful mathematically, it can be employed to further identify scenes with subtle differences. However, mis-classification of pixels along the edge between two categories may occur if no reject category is given due to a larger window (5 by 5) as compared to Model I (3 by 3). If a reject class is available, these edge pixels will be classified as rejects.

Therefore, it seems that the best strategy is to combine the results from Model I and Model II in the production of the final decision map. An alternative is to "edit" and remove the edge effect by utilizing an algorithm which can reclassify these edge points into one of the neighboring groups (RADC, 1976).

The format of the decision map can be either computer maps or color transparencies produced by a color printer from a magnetic tape (Hsu, 1977). The former is handy, but with scale distortions and without artistic appeal. The latter is more expensive, but without the drawbacks inherited in the computer printed maps. For investigational purposes, the computer map is preferred because the decision information regarding individual pixels is available. For publication of the decision map, the second method is highly recommended.

A HIT-RATE ANALYSIS

To assess the performance of the developed texture measures in Model I, a hit-rate analysis of a test site has been carried out. The procedures include (1) placing a 10 by 10 grid onto both the computer decision map and the photo print of the test area, and (2) estimating and enumerating the per-

TABLE 4. SQUARE OF THE MAHALANOBIS DISTANCE BETWEEN EACH PAIR OF GROUP CENTROIDS RELATIVE TO ROW GROUP DISPERSION FROM MODEL I

	Pavement	C Field 1	C Field 2	Vegetatn	Metal	Soil 1	Soil 2	Edgepave
Pavement	0.0 (P < 1.000)	(P < 0.000)	(P < 0.000)	(P < 0.000)	73115.26 (P < 0.000)	(P < 0.000)	(P < 0.000)	(P < 0.000)
C Field 1	4370.07 (P < 0.000)	0.0 (P < 0.000)	50.56 (P < 0.000)	270.03 (P < 0.000)	7484.28 (P < 0.000)	1705.49 (P < 0.000)	1492.94 (P < 0.000)	1316.92 (P < 0.000)
C Field 2	385.52 (P < 0.000)	14.02 (P < 0.000)	0.0 (P < 0.000)	15.61 (P < 0.000)	3144.59 (P < 0.000)	495.81 (P < 0.000)	391.65 (P < 0.000)	580.19 (P < 0.000)
Vegetatn	1347.83 (P < 0.000)	148.71 (P < 0.000)	39.37 (P < 0.000)	0.0 (P < 0.000)	2200.87 (P < 0.000)	893.77 (P < 0.000)	785.18 (P < 0.000)	782.04 (P < 0.000)
Metal	530.90 (P < 0.000)	92681.51 (P < 0.000)	169500.01 (P < 0.000)	275864.13 (P < 0.000)	0.0 (P < 0.000)	349.65 (P < 0.000)	464.65 (P < 0.000)	849.90 (P < 0.000)
Soil 1	532.87 (P < 0.000)	1309.25 (P < 0.000)	2423.30 (P < 0.000)	3972.90 (P < 0.000)	1368.96 (P < 0.000)	0.0 (P < 0.000)	89.87 (P < 0.000)	51.54 (P < 0.000)
Soil 2	727.52 (P < 0.000)	6521.19 (P < 0.000)	11703.00 (P < 0.000)	18517.45 (P < 0.000)	875.21 (P < 0.000)	15.52 (P < 0.000)	0.0 (P < 0.000)	36.67 (P < 0.000)
Edgepave	921.56 (P < 0.000)	554.00 (P < 0.000)	1009.35 (P < 0.000)	1573.01 (P < 0.000)	1325.54 (P < 0.000)	35.95 (P < 0.000)	4.25 (P < 0.000)	0.0 (P < 0.000)

TABLE 5. CONFUSION MATRIX (ACTUAL, GROUP DOWN, PREDICTED GROUP ACROSS) FROM MODEL II

	Pavement	C Field 1	C Field 2	Vegatatn	Metal	Soil 1	Soil 2	Edgepave
Pavement	300	0	0	0	0	0	0	0
C Field 1	0	446	4	0	0	0	0	0
C Field 2	0	0	249	1	0	0	0	0
Vegatatn	0	0	1	99	0	0	0	0
Metal	0	0	0	0	90	0	0	0
Soil 1	0	0	0	0	0	45	0	0
Soil 2	0	0	0	0	0	1	62	1
Edgepave	0	0	0	0	0	0	0	174

Total correct classifications = 99.46

centage of all terrain type classes in each cell. The ground truth information was obtained by manual operations, whereas the decision map information was obtained by a computerized method. Note that these two processes were carried out independently by two groups of research assistants. The choice of the 10 by 10 grid evaluation cell was made according to the physical size of the high-resolution U-2 images for interpreting the ground truth; that is, with this grid system, the size of each cell is about half an inch, which is convenient for determining the percentage of each class within the cell using manual estimations. The hit-rate is computed as

Hit rate = 1 − Difference between photo-interpretation and computer-decision map/photo-interpretation (in terms of total area of each class) or = 1 − omission error rate.

The result is given in Table 7.

This analysis shows that, in general, a hit-rate of 85 percent to 90 percent (except for small areas) has been achieved by Model I. The error rate in discriminant analysis given in Table 7 has been judged to be insignificant based on the fact that the size of the training sets were very large (an average of 200 pixels for each class). This argument is derived from Foley's principle: for a valid discriminant analysis, the minimum size of the training set should be three times as large as the variables used in the discriminant functions (Foley, 1971).

It should be noted that the photo interpretation of the ground truth was obtained from a high-resolution aerial photo rather than low-resolution images from which the computer decision map was derived.

We have tried to map water bodies. In general, we can identify them; however, the hit-rate is statistically meaningless because the area is too small for a valid numerical analysis. It was also found that digitizing errors (with a vidicon system) existed in the high-altitude images (GAHA, VPHA, URHA); thus, the hit-rate for these frames must be obtained by sub-groups within categories. For instance, in VPHA, two types of cultivated fields were used in the training sets. Some detailed analysis of the hit-rate and false alarm rate for each frame is given below.

GALA. The analysis shows that a hit-rate over 90 percent (except for soil) has been achieved by Model I. The author has investigated further the problem regarding the soil class using the output from Model II. It was first thought to be the "edge effect." However, since the mis-classification of the soil pixels was largely eliminated in Model II, it was therefore determined to be "resolution effect," which was purposely induced into the images during the process of digitization.

SBLA. In general, the overall terrain pattern came out very well in the decision map. The metal-objects were correctly identified using the reject category, as intended. Similar to GALA, "resolution" effect exists at the "edge" of two distinctive classes and at certain vegetation areas.

The rejects were about 10 percent of the total area. There was no significant difference between the "reject" pattern determined by $P(x/G) = 0.01$ and that by $P(x/G) = 0.001$. This means that the pixels being rejected were really different from the design sets.

VPLA. The overall terrain pattern in the decision map was good in the sense that essential types were correctly identified. In terms of a detailed hit-rate analysis, the correct classification rate is about 85 percent (excepting pavement). Two factors caused the error rate: (1) asphalt-paved roads could not be differentiated from fields used for recreational purposes; and (2) a new concrete road was being built at the time the image was taken—many types of "pavement" were present at this section of the image. If cultivated field and pavement were treated as one group, the hit-rate will be over 95 percent.

To achieve a correct classification of this

TABLE 6. SQUARE OF THE MAHALANOBIS DISTANCE BETWEEN EACH PAIR OF GROUP CENTROIDS RELATIVE TO ROW GROUP DISPERSION FROM MODEL II

	Pavement	C Field 1	C Field 2	Vegatatn	Metal	Soil 1	Soil 2	Edgepave
Pavement	0.0 (P < 0.000)	153885.26 (P < 0.000)	212962.32 (P < 0.000)	284292.49 (P < 0.000)	739.02 (P < 0.000)	29725.14 (P < 0.000)	31618.34 (P < 0.000)	38254.80 (P < 0.000)
C Field 1	8979.62 (P < 1.000)	0.0 (P < 0.000)	84.66 (P < 0.000)	444.37 (P < 0.000)	19560.23 (P < 0.000)	2351.88 (P < 0.000)	2022.89 (P < 0.000)	2004.62 (P < 0.000)
C Field 2	362.74 (P < 0.000)	18.53 (P < 0.000)	0.0 (P < 1.000)	34.33 (P < 0.000)	8368.47 (P < 0.000)	1161.90 (P < 0.000)	1184.07 (P < 0.000)	1019.67 (P < 0.000)
Vegatatn	1978.49 (P < 0.000)	362.28 (P < 0.000)	102.22 (P < 0.000)	0.0 (P < 1.000)	3195.81 (P < 0.000)	1360.89 (P < 0.000)	1374.05 (P < 0.000)	1357.40 (P < 0.000)
Metal	1155.31 (P < 0.000)	39390.28 (P < 0.000)	134969.04 (P < 0.000)	0.0 (P < 0.000)	0.0 (P < 1.000)	1085.48 (P < 0.000)	1320.20 (P < 0.000)	660.97 (P < 0.000)
Soil 1	3819.53 (P < 0.000)	5390.93 (P < 0.000)	11994.95 (P < 0.000)	21753.31 (P < 0.000)	7589.26 (P < 0.000)	0.0 (P < 1.000)	122.50 (P < 0.000)	71.01 (P < 0.000)
Soil 2	24736.62 (P < 0.000)	41504.21 (P < 0.000)	86500.74 (P < 0.000)	150277.45 (P < 0.000)	25203.98 (P < 0.000)	100.96 (P < 0.000)	0.0 (P < 1.000)	77.35 (P < 0.000)
Edgepave	4628.38 (P < 0.000)	3604.24 (P < 0.000)	34144.28 (P < 0.000)	117667.75 (P < 0.000)	41260.38 (P < 0.000)	125.31 (P < 0.000)	38.56 (P < 0.000)	0.0 (P < 1.000)

frame, four types of cultivated field were used in the training sets to cover significant local variations. In terms of the training set itself, a hit-rate of 98.4 was achieved. However, in terms of the test set, the hit-rate is much lower due to significant local variations.

URLA. The URLA was a more complicated frame; thus, an iterative process was utilized to generate the decision maps. The more obvious classes, such as metal, pavement, composition, etc., were processed first and the "uncertain" and insignificant (in terms of aereal coverage) vegetation were left out. The "reject" area thus represents mixed water, vegetation and cultivated fields, etc. At both 0.01 and 0.001 probability reject levels, the area showing "rejects" is very small, corresponding to a potential area of mixed water and vegetation.

SBHA. This was the only frame in the high-altitude image group that had few digitization problems. The generation of the decision map was therefore rather straightforward due to less complexity in the terrain configuration, and a very high hit-rate was achieved (over 95 percent).

VPHA. Image digitization error existed in the frame; specifically, the upper one-third is much lighter than the lower two-thirds portion of the frame. Using the RADC DICIFER system, it was determined that a 30-point difference existed between these two portions of the frame for cultivated field category.

GAHA and URHA. The same digitization problem caused the fact that the NE corner of GAHA is much lighter than the same terrain types in the SW corner. To process this frame, two artificial types of cultivated fields had to be used in the design sets. Since vegetation and cultivated field classes were really confused by this digitization effect, they were grouped as one class in the hit-rate analysis.

We were unable to obtain a reliable hit-rate for URHA due to the same digitization problem. However, we were able to produce a fairly good decision map in terms of the overall terrain pattern.

CONCLUSION AND DISCUSSIONS

In this paper two models of texture analysis were discussed. Model I with 17 spatial-tone measures derived from a 3 by 3 data matrix was determined as very effective in classifying general land use types. With six additional wave-form parameters, Model II is developed specifically to discriminate objects and scenes of subtle differences. A

TABLE 7. HIT-RATE OF MODEL I

Classes	Vegetation	Cultivated Field	Metal	Soil	Pavement	Water	Composition
Low Altitude							
Frame 1	88.4%	98.46%	90%	53.13%	92.28%	—	—
Frame 2	89.81%	89.46%	—	87.13%	—	—	—
Frame 3	—	—	80.50%	45.90%	85.24%	—	87.4%
Frame 4	90%	85.5%	95.0%	86%	—	—	—
High Altitude							
Frame 5	88.51%		—	85.53%	72%	—	75.9%
Frame 6	99%	95%	—	98%	—	—	95%
Frame 7	60%	84.1%	—	93.7%	70%	—	85.1%

hit-rate of 85 to 90 percent has been determined regarding the classification of general land-use types with panchromatic images. This methodology can easily be adopted for multispectral systems.

The solution algorithms for Model I and Model II are presently programmed in FORTRAN language. The processing time, therefore, is long. It takes about 90 minutes CPU time for Model I and 130 minutes for Model II with respect to processing 256 by 256 pixels with the IBM 370-158 system. The time can be shortened to one-fourth of this amount by using the assembler language. The alternative is to use only the most effective variables in the analysis coupled with a better classifier, which is currently being investigated by this author under the sponsorship of the U.S. Air Force/Rome Air Development Center.

In addition to the feature extractor and the classifier, the hit-rate and false alarm rate also depend on the factors regarding sample size, the location, and the number of the training sets.

The minimum sample size problem has been investigated by Foley (1971). His principle states that for a valid analysis the minimum sample size is three times as large as the number of the variables used. For instance, if one employs ten texture variables in the analysis, the minimum number of each training set is 30. It is also our experience that the Forley principle is valid and that empirically the sample size of each training set should be greater than 30 pixels in general.

Improper training sets generally lead to a low hit-rate. To avoid such an error, one should first employ the confusion matrix (from the training sets) to identify confused classes and to locate mis-classified pixels on the (preliminary) decision map. Then, one should change the location of the training sets in order that the "pure" training sets can be obtained. This is an iterative process, and it can be done manually or by the operator using interactive graphics, i.e., using a cursor on the color monitor with a terminal control. Once the correct classification rate in the design set reaches a level of 90 percent or over, the investigator can proceed to classify the test set data.

To classify the test set, one can classify each group at a time, or classify many groups in one process. Theoretically, the first method will yield a lower hit-rate because there is only one probability value for each pixel to be used in the classification, which may not be maximum once other groups are introduced. Most likely, this method will produce overlapping groups; that is, an individual pixel may belong to several groups.

To insure that the test sets are properly classified, all the desired groups should be introduced in the design set. Furthermore, if local differences exist within one group, sub-groups should be introduced. These sub-groups can be labeled as one group only after the decision map is produced. It is our experience that a sufficient number of groups should be used in the design sets; otherwise, mis-classifications or rejects will be substantial.

An immediate application of this texture analysis will be in the development an Automatic Feature Extraction System (AFES) intended to produce maps with detailed areal and linear information from digitized imagery data. This is currently pursued by the U.S. Air Force/Rome Air Development Center for the U.S. Defense Mapping Agency (RADC, 1976b).

REFERENCES

Barlow, R. M., R. Narasimahan, and A. Rosenfeld. 1972. "Visual Patterns Analysis in Machines and Animal" *Science*, 177, pp. 567-575.

Foley, D. 1971. "The Probability of Error in the Design Set as a Function of the Sample Size and Dimensionality," Technical Report, RADC-TR-71-171.

Gibson, J. J. 1950. *The Perception of the Visual World,* New York: Houghton Mifflin.

Glick, N. 1973. "Sample-based Multi-nomial classification" *Biometrica* 29, pp. 241-256.

_________ 1977. "Additive Estimators for Probabilities of Correct Classification" IEEE, *Pattern Recognition and Image Processing,* pp. 304-317.

Gramenopoulos, N. 1973. "Terrain type Recognition Using ERTS-1 MSS Images," *Symposium on Significant Results Obtained from the Earth Resources Technology Sattelite,* NASA SP-327, pp. 1229-1241.

Haralick, R. M. 1970. *Texture Tone Study with Applications to Digitized Imagery,* Technical Report 181-1, Center for Research, Inc., University of Kansas, Lawrence, Kansas.

_________ 1975. "A Resolution Preserving Textural Transform for Images" *IEEE Proceedings of the Conference on Computer Graphics, Pattern Recognition & Data Structure,* pp. 51-61.

Haralick, R. M., K. Shanmagam, and I. Dinstein. 1973. "Textural Features for Image Classification" *IEEE Transactions on Systems, Man, and Cybernetics,* Vol. SMC-3, No. 6, pp. 610-621.

Hornung, R. J., and J. A. Smith. 1973. "Application of Fourier Analysis to Multispectral/Spatial Recognition," *Management and Utilization of Remote Sensing Data ASP Symposium,* Sioux Falls, South Dakota.

Hsu, S. 1972. "Application of the Discriminant Function in Automated Pattern Recognition of Side-Looking Radar Imagery" (with Kedar) in *Imaging Techniques For Testing and Inspection,* SPIE, pp. 145-148.

_________ 1974. "A Mathematical Method for Measuring Visual Differences in Multiple Map Comparisons," *Proceedings, American Congress on Surveying and Mapping,* pp. 116-32.

_________ 1977a. "A Texture-Tone Analysis for Landuse Mapping with Panchromatic Images" *Proceedings, American Society of Photogrammetry,* pp. 203-215.

_________ 1977b. "Texture Tone Feature Extraction and Analysis," Rome Air Development Center, Air Force Systems Command, Griffiss Air Force Base, Rome, New York. (RADC-TR-77-279).

_________ 1978. "Automation of Thematic Maps with Electro-Optical Colormation Method" (in Press) *The Canadian Cartographer.*

Julesz, B. 1975. "Experiments in the Visual Perception of Texture" *Scientific America,* 232(4), pp. 34-43.

Kaizer, H. 1955. "A Quantification of Textures on Aerial Photographs," Boston University Research Laboratories, Technical Note 121, AD 69484.

Kirvida, L., and G. Johnson. 1973. "Automatic Interpretation of Earth Resources Technology Satellite Data for Forest Management," *Symposium on Significant Results Obtained from the Earth Resources Technology Satellite,* NASA SP-327, pp. 1076-1082.

Lachenbruch, P. A., and R. R. Mickey. 1968. "Estimation of Error Rates in Discriminant Analysis" *Technometrics* 10, pp. 1-10.

Lendaris, G. G., and G. L. Stanley. 1970. "Diffraction-Pattern Sampling for Automatic Pattern Recognition," *Proceedings of the IEEE,* Vol. 58, No. 2, pp. 198-216.

Lev, A., S. W. Zucker, and A. Rosenfeld. 1977. "Iterative Enhancement of Noisy Images" *IEEE Transactions on Systems, Man, and Cybernetics,* Vol. SMC-7, No. 6, pp. 435-442.

Matheron, G. 1967. *Elements Pour Une Theorie des Milieux Poreux,* Masson, Paris.

McLachlan, G. J. 1976. "The Bias for the Apparant Error Rate in Discriminant Analysis" *Biometrica,* 63,2, pp. 239-44.

Mitchell, O. R. 1976. "Texture Edge Detection and Classification Using Max-min Descriptors" *Image Understanding and Information Extraction,* T. S. Huang & K. S. Fu, Contract No. F3-6-2-75-C-0150, Griffiss Air Force Base, Rome, New York.

Mitchell, O. R., C. R. Myers, and W. Boyne. 1977. "A Max-min Measure for Image Texture Analysis" *IEEE Transaction on Computers,* pp. 408-414.

Morrison, O. F. 1976. *Multivariate Statistical Methods,* New York: McGraw Hill, (2nd Edition).

Nalepka, R. 1970. *Investigation of Multi-spectral Discrimination Techniques,* Report No. 2264-12-F. Willow Run Laboratories, Institute of Science and Technology, University of Michigan.

Pickett, R. M. 1967. "The Perception of Random Visual Texture", in *Models for the Perception of Speech and Visual Form,* ed. W. Wathem-Dunn, MIT Press, Cambridge, Mass., pp. 224-232.

Polit, A. 1976. "Computer Generations of Textures", *Behavior Research Methods and Instrumentation,* 8(4), pp. 367-368.

Purks, S. R., and W. Richards. 1977. "Visual Texture Discrimination Using Random-dot Patterns" J. *Opt. Soc. Am.,* 67, pp. 765-771.

RADC/Rome Air Development Center. 1976a. *Digital Image Processing Techniques for Automatic Terrain Classification for Generating Reference Maps from B/W Aerial Photography.* RADC-TR-76-196 Final Technical Report by Pattern Analysis and Recognition Corporation. Rome, New York.

_________ 1976b. "Statement of Work for Automatic Feature Extraction System", PR1-7-4700. Rome, New York.

Rao, C. R. 1973. *Linear Statistical Inference and Applications* (2nd Edition), New York: Wiley.

Rao, C., and S. Mitra. 1971. *Generalized Inverse of Matrices and its Applications,* pp. 408-414, New York: John Wiley & Sons.

Richards, W. 1977. *Experiments in Texture Perception,* Annual Report Contract F44620-74-0076, Air Force Office of Scientific Research, Bolling Air Force Base, Sponsor: ARPA Order No. 2765.

Rosenfeld, A. 1975. *Visual Texture Analysis: An Overview*, Technical Report Series (TR-406, F44620-72C-0062) University of Maryland, College Park, Maryland.

Rosenfeld, A., and M. Thurton. 1971. "Edge and Curve Detection for Visual Scene Analysis" *IEEE Transactions on Computers*, Vol. C-20, No. 5, pp. 562-569.

Schachter, B. J., A. Lev, S. W. Zucker, and A. Rosenfeld. 1977. "An Application of Relaxation Methods to Edge Enforcement" *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-7, pp. 813-816.

Serra, J., and G. Verchery. 1973. "Mathematical Morphology Applied to Fibre Composite Materials," *Film Science and Technology*, Vol. 6, pp. 141-158.

Weszka, J., and A. Rosenfeld. 1975. "A Comparative Study of Texture Measures for Terrain Classification" IEEE Proceedings of the Conference on *Computer Graphics, Pattern Recognition & Data Structure*, pp. 62-64.

(Received November 7, 1977; revised and accepted June 29, 1978)

Notice to Contributors

1. Manuscripts should be typed, double-spaced on $8\frac{1}{2} \times 11$ or $8 \times 10\frac{1}{2}$ white bond, on *one* side only. References, footnotes, captions—everything should be double-spaced. Margins should be $1\frac{1}{2}$ inches.
2. Ordinarily *two* copies of the manuscript and two sets of illustrations should be submitted where the second set of illustrations need not be prime quality; EXCEPT that *five* copies of papers on Remote Sensing and Photointerpretation are needed, all with prime quality illustrations to facilitate the review process.
3. Each article should include an abstract, which is a *digest* of the article. An abstract should be 100 to 150 words in length.
4. Tables should be designed to fit into a width no more than five inches.
5. Illustrations should not be more than twice the final print size: *glossy* prints of photos should be submitted. Lettering should be neat, and designed for the reduction anticipated. Please include a separate list of captions.
6. Formulas should be expressed as simply as possible, keeping in mind the difficulties and limitations encountered in setting type.