SHIN-YI HSU
State University of New York
Binghamton, NY 13901

Automated Pattern Recognition with Thermal IR Imagery *

An analytical model is employed which can classify objects with an accuracy rate of 90 per cent.

INTRODUCTION

DURING THE PAST two decades, most photographic and imagery interpretation has been performed manually. Recently, the development of the false-color enhancement technique (for imagery tonal densities) has made a semi-automatic interpretation process possible. However, due to the massive amounts of data from satellites such as *ERTS* and *Skylab* and from low-and-high-altitude aircraft, automated interpretation processes utilizing computers have become increasingly important. This is particularly true when objects are to be identified from multispectral imagery. However, automated imagery interpretation systems currently installed at the major centers for remote sensing research, such as *LARS* of the Purdue University, are rarely cost-effective. For instance, it costs $50,000 yearly for Indiana State University to have a remote terminal link-up with the *LARS* system (excluding costs for computing times). Low-cost computerized interpretation methods are urgently needed. Turinetti and Mintzer have reported that such low-cost systems could be obtained by utilizing digital computers in the analysis of densitometer scanned data obtained from thermal infrared imagery (1973) and non-registered, multiformat, multispectral imagery (1974). Their 1973 analysis demonstrated that heat differentials of terrestrial objects derived from daytime and nighttime infrared images (*TIRD* and *TIRN*) can be used to identify objects by computers at an accuracy rate of over 40 per cent, and their 1974 analysis demonstrated that the accuracy rate can be improved up to over 90 per cent by utilizing multiple formats of thermal in-

ABSTRACT: *This paper proposes the employment of both day-and-night thermal-infrared imagery density values in a multivariate model for automatic pattern recognition. The analytical model is discriminatory analysis, utilizing the discriminatory function, derived from the day-and-night-density vectors, for pattern recognition and mapping purposes. For mapping purposes, two models are developed further. First, to identify and map single objects, such as houses, roads, and water, a point-classification system is used. The discriminant function is derived from only two vectors. Second, to identify scenes, such as residential areas, factories and croplands, an areal-classification model is used. The discriminant function is then derived from 10 parameters (5 from* X-Z *axes, and 5 from* Y-Z *axes) extracted from the density surface constructed by densitometer scanning and computer graphics. The results indicated that an accuracy rate of well over 85 per cent can be accomplished for automatic mapping purposes. A low-cost automatic imagery interpretation system can be obtained.*

* Presented at the Fall Technical Meeting, American Society of Photogrammetry, Washington, D.C., September 1974.

frared imagery, equivalent to using multispectral imagery. The purpose of this paper is to derive an automatic pattern recognition method by which an accuracy rate of 90 per cent can be obtained by using day- and night-time thermal infrared imagery alone.

The analytical model is discriminatory analysis, using the discriminant function derived from the day- and night-time density vectors, for classifying unknown objects into one of the calibration samples. Two models have been designed for the analysis: Model I classifies objects according to the spectral response of tones (points), and Model II recognizes scenes according to the spectral response of texture (spatial distribution of tones).

METHODOLOGY

RESPONSE VARIABLES OF SPECTRAL SIGNATURE

Object-identification in remote sensing is usually done by two integrated processes: (1) laboratory analysis and identification of spectral responses of terrestrial objects; and (2) matching of these spectral signatures with imagery characteristics of the same objects. Spectral signatures obtained from laboratory analysis can be used only as a guide for imagery interpretation, because laboratory conditions will not be identical to the field conditions under which the imagery was taken. However, we are able to obtain comparable spectral signatures from imagery data by utilizing multiband imagery. This process involves converting imagery tonal densities from continuous bands, representing spectral reflectance and/or emittance, to numerical scales utilizing densitometer scanning, or false color enhancement using equipment such as I²S manufactured by International Imaging Systems, or matching against paper graytone step wedge, and then plotting them on a two-dimensional diagram as in Figure 1.

In Figure 1 the x coordinate represents spectral bands; the number depends on the scanner system. The X coordinate represents tonal densities scaled into steps:

> 32 steps for I²S;
> 10 steps for gray-tone wedge; and many steps for densitometer data, the number to be specified by the investigator.

The curve in the graph is called the spectral signature, each curve representing a terrestrial object.

To distinguish one object from another, several response variables (or attributes) have to be extracted from the curve and used in pattern recognition. Since each band constitutes an imagery, tone density values from all the bands can be used as attributes for imagery discrimination. In addition, we can extract attributes from the distributional characteristics of the curve as a whole. Five response variables (attributes) can be extracted: (1) total area above the datum (arbitrarily assigned), (2) total area below the datum, (3) sum of contrast values, (4) sum of absolute high values, and (5) sum of the number of peaks and troughs. Therefore, we have two sets of response variables: (1) spectral response from bands, and (b) spectral response from the curve (Figure 1).

To make a statistical test valid and meaningful, a larger sample (objects) has to be used in the analysis. This means that we should use at least 20 similar objects (such as 20 houses) to characterize the spectral response of *house*. Assuming that we want to classify unknown objects into one of the ten known categories, the data matrix using five multiband images is as follows:

Object 1
Response Variables

$$\begin{array}{cc}
\text{Samples} & \begin{array}{cccccc} 1 & 2 & \text{----} & 5 & 6 & 7 \text{----} 10 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ \cdot \\ \cdot \\ \cdot \\ 20 \end{array} &
\begin{bmatrix}
y^{(1)}_{11} & y^{(2)}_{11} & \cdots & y^{(5)}_{11} & y^{(6)}_{11} & y^{(7)}_{11} \cdots y^{(10)}_{11} \\
y^{(1)}_{12} & y^{(2)}_{12} & \cdots & y^{(5)}_{12} & y^{(6)}_{12} & y^{(7)}_{12} \cdots y^{(10)}_{12} \\
\cdot & \cdot & & & \cdot & \cdot \quad \cdot \\
\cdot & \cdot & & & \cdot & \cdot \quad \cdot \\
\cdot & \cdot & & & \cdot & \cdot \quad \cdot \\
y^{(1)}_{1,20} & y^{(2)}_{1,20} & \cdots & y^{(5)}_{1,20} & y^{(6)}_{1,20} & y^{(7)}_{1,20} \cdots y^{(10)}_{1,20}
\end{bmatrix} \\
\text{Mean} & \begin{bmatrix} \bar{y}^{(1)}_{1} & \bar{y}^{(2)}_{1} & \cdots & \bar{y}^{(10)}_{1} \end{bmatrix}
\end{array}$$

Object 10
Response Variables

$$\begin{array}{cc}
 & \begin{array}{cccc} 1 & 2 \text{---} & & 10 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ \cdot \\ \cdot \\ 20 \end{array} &
\begin{bmatrix}
y^{(1)}_{10,1} & y^{(2)}_{10,1} & \cdots & y^{(10)}_{10,1} \\
\cdot & \cdot & & \cdot \\
\cdot & \cdot & & \cdot \\
\cdot & \cdot & & \cdot \\
y^{(1)}_{10,20} & y^{(2)}_{10,20} & \cdots & y^{(10)}_{10,20}
\end{bmatrix} \\
\text{Mean} & \begin{bmatrix} \bar{y}^{(1)}_{10} & \bar{y}^{(2)}_{10} & \cdots & \bar{y}^{(10)}_{10} \end{bmatrix}
\end{array}$$

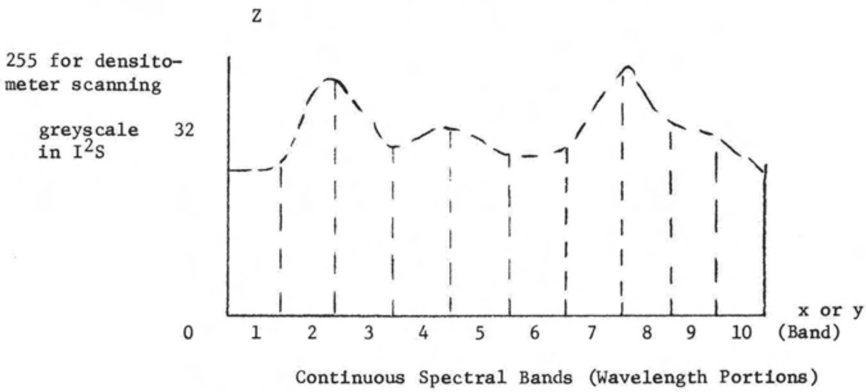


FIG. 1. Two-dimensional spectral signature diagram.

Here the first five attributes are obtained from the response of the five bands, while the second five variables are from the response of the curve.

The superscript stands for attributes; the first subscript stands for group (objects) identification, and the second subscript sample point.

MULTIVARIATE DISCRIMINANT ANALYSIS (PATTERN RECOGNITION OF OBJECTS)

In the analysis, the data matrix can be simplified in vector form:

$$\text{Object 1} \quad [\underline{y}_1^{(1)} \ \text{---}\ \underline{y}_1^{(10)}] \qquad \text{Object 10} \quad [\underline{y}_{10}^{(1)} \ \text{---}\ \underline{y}_{10}^{(10)}]$$

Then we test ten hypotheses simultaneously to see whether there are significant differences between and among the group means

$$Ho \quad \bar{y}_1^{(1)} = \bar{y}_2^{(1)} = \dots = \bar{y}_{10}^{(1)}$$

$$\vdots$$

$$\bar{y}_1^{(10)} \quad \bar{y}_2^{(10)} = \dots = \bar{y}_{10}^{(10)}$$

Or in vector form

$$Ho \quad \underline{\bar{y}}_1 = \underline{\bar{y}}_2 \text{---} = \underline{\bar{y}}_{10}$$

Alternative: one of the equalities does not exist.

The approach is to combine the ten response variables into a "weighted total"

$$Z = a_1\underline{y}^{(1)} + a_2y^{(2)} + \text{---} + a_{10}\underline{y}^{(10)}$$

and choose $a_1, a_2, \text{---}, a_{10}$ in such a way that the distance between the groups is maximized. The discriminant function is in fact the estimated values of the above function

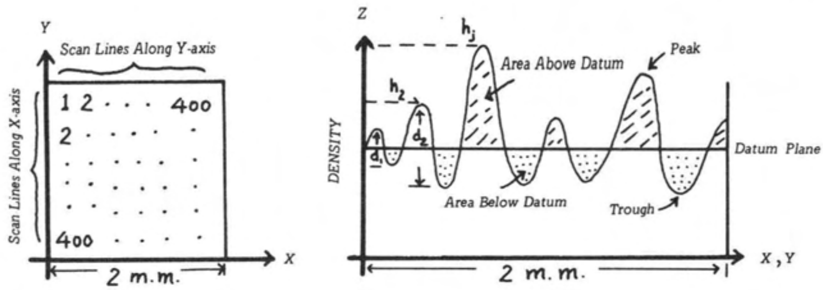
$$\hat{Z} = \hat{a}_1\underline{y}^{(1)} + \hat{a}_2\underline{y}^{(2)} + \text{---} + \hat{a}_{10}\underline{y}^{(10)}$$

Utilizing the UCLA BMD 0 7M program for the analysis, the distance between pairs of objects is Mahalanobis D^2. Once we know there are some values for the *a's* (maximum is used), the discriminant function is determined. It will then be used to classify any unknown features into one of the known objects called calibration samples.

PATTERN RECOGNITION OF SCENES

What I have discussed so far only concerns pattern recognition of objects in terms of point-by-point classification. Another aspect of pattern recognition involves classifying associated objects distributed over an area called a scene, such as a residential area, airport, industrial complex, and so on. Meaningful characteristics of scenes can be extracted only from the spatial distribution of tonal densities (responses from objects) in the *scene*, not by point classifications. The structural difference between the object and the scene is that the former is a two-dimensional space (with density values *(Z)* distributed on the *X* or *Y* plane) whereas the latter is a three-dimensional statistical surface of Z values distributed on the *X-Y* plane. Figure 2 is a perspective representation of the surface.

For the purpose of numerical analysis, the surface can be separated into two compo-



1) TOTAL AREA ABOVE THE DATUM PLANE.

2) TOTAL AREA BELOW THE DATUM PLANE.

3) SUM OF THE CONTRAST VALUE OR Σd_i

4) SUM OF THE ABSOLUTE HIGH DENSITY VALUES OR Σh_j

5) SUM OF THE NUMBER OF PEAKS AND TROUGHS

(ALONG BOTH X AND Y AXES)

FIG. 2. Three-dimensional spectral signature diagram.

nents: Z-values on the X-axis and Z-values on the Y-axis.

Furthermore, the spectral signature, or the distribution of Z-values, can be characterized by the above-mentioned five response variables: (1) total area above the datum, (2) total area below the datum, (3) sum of contrast values, (4) sum of absolute high density values, and (5) sum of the number of peaks and troughs (Figure 2). Therefore, for both X and Y axes, there will be altogether ten response variables that can be used to characterize the spectral response surface of a *scene*.

The discriminant function thus becomes

$$\hat{Z} = \hat{a}_1\underline{y}_1 + \hat{a}_2\underline{y}_2 + \cdots + \hat{a}_{10}\underline{y}_{10}$$

The technique of automatic pattern recognition utilizing the discriminant function in scenes is the same as that used in classifying objects. It has been proved very effective by this investigator in the pattern recognition of crop types utilizing microdensitometer scanned side-looking radar imagery (Kedar and Hsu, 1972). The combination of the two interpretation methods, i.e., object identification and scene identification, should be very effective in a computerized pattern recognition system. If we want to identify small objects, such as houses, trucks, and so on, we should employ the object identification method by the point and line scanning procedure. Again, if we are to identify a residential area, the second method should be employed. The procedures are to define one unit area consisting of perhaps 10 scan lines on both the X and Y axes, to design a moving grid system of 10-by-10 scan lines, and (3) pick up *overlapping* areas for analysis. The amount of overlapping can be specified by the analyst. Using a computerized system, several options should be provided, such as two-thirds and one-half overlapping systems as shown in Figure 3.

The reason for using this area overlapping system is to guard against misclassification involving split scenes. For instance, the first cell picked up for analysis may include only half the missile site; therefore, it may not be recognized as a missile site. The second cell, which picked up, for instance, half of the first cell and half of the next regular cell, should include the whole of the missile site. It should then be classified as a missile site.

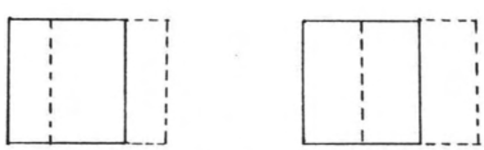


FIG. 3. Two-thirds and one-half area overlapping systems.

This system is very versatile. The unit size for a scene can vary according to specific purposes and the amount of overlapping between cells can also vary depending on accuracy requirements.

Case Studies Utilizing Thermal IR Imagery

The methodologies discussed previously have been successfully applied to the automatic classification of objects and scenes utilizing daytime and nighttime thermal infrared imagery. The material was supplied by Rome Air Development Center. The gray tones were determined by matching against paper graytone step wedge. They can be determined more accurately, of course, by means of densitometer or TV scanning processes. The discriminant function was derived from only two vectors: tonal densities of TIRD and TIRN. To identify objects, BMD computer program 07M (developed by UCLA), is utilized. Calibration samples (groups) are placed (stored) first, and unknown objects are listed together as the last group. All of the unknown objects will be classified into one of the groups. In our experiment (with five independent groups), an accuracy rate of about 90 per cent is obtained in differentiating among roads, water, houses, and fields.

To identify scenes, UCLA BMD program 04M is to be utilized. The program is designed to distinguish groups rather than cases in the groups. We can define a scene, such as a residential area, as a group consisting of several scan lines. In the analysis, these scan lines are comparable to "cases." The program obtains a discriminant function for each pairwise (double-group) comparison. Unknown scenes can then be classified as one of the known groups by means of the discriminant functions. The results are expected to be better than those obtained in the object classification study because the discriminant function in scene classification is more powerful — it includes more response variables.

References

Hsu, S. Y., (co-author with E. Y. Kedar), "Applications of the Discriminant Function in Automatic Pattern Recognition of Side-Looking Radar Imagery," *Imaging Techniques for Testing and Inspection, Proceedings of the Society of Photo-Optical Instrumentation Engineers*, Vol. 29, pp. 145-148. (1972)

Turinetti, J. D., and O. W. Mintzer, "Computer Analysis of Imagery," *Photogrammetric Engineering*, Vol. XXXIX, No. 5 (1973), pp. 501-505.

________, "Low Cost Computerized Land Use Classification," a paper given at the American Society of Photogrammetry, Fall Meeting, 1973, at Orlando, Florida.

Articles for Next Month