Combined Spectral and Spatial Processing of ERTS Imagery Data

ROBERT M. HARALICK and K. SAM SHANMUGAM

University of Kansas Center for Research Inc., Remote Sensing Laboratory, Lawrence, Kansas 66044

In addition to spectral features, texture is an important spatial feature used in identifying objects or regions of interest in an image. Although texture is relatively easy for human observers to recognize and describe in empirical terms, it has been extremely refractory to precise definition and analysis by digital computers. This paper describes a procedure for extracting some easily computable features for the texture of blocks of digital image data and illustrates the applications of combined textural (spatial) and spectral features for identifying the land use categories of blocks of ERTS MSS (Earth Resources Technology Satellite Multi Spectral Scanner) data.

The land use classification algorithm based on textural and spectral features was developed and tested using 614 image blocks of 64 × 64 resolution cells derived from an ERTS image over the Monterrey Bay area of the California coast line. The algorithm was applied on a training set of 314 blocks and tested on a set of 310 blocks. The overall accuracy of the classifier was found to be 83.5% on seven land use categories.

I. Introduction

Spectral, textural, and context features are three fundamental pattern elements used in human interpretation of imagery data. Spectral features describe the band-to-band tonal variations in a multiband image set, whereas textural features contain information about the spatial distribution of tonal values within a band. Context features contain information derived from areas surrounding the sub-image region being analyzed. When small image areas are independently processed by a machine, only the textural and spectral features are available to the machine.

In much of the automated procedures for processsing image data from small areas, such as in crop classification studies, only the spectral features are used for developing a classification algorithm. While the "Per-field" and "sample" classification schemes do not ignore texture entirely, a comprehensive set of features for texture have not been defined and used in automated classification schemes. Because the areal characteristics of texture appear to carry valuable information, it is important to use the textural features in automated image processing schemes except in applications where the poor resolution of the imagery does not provide meaningful textural information.

Earlier image texture studies have employed autocorrelation functions (Kaizer, 1955), power spectral (Chavallier *et al.*, 1968) or restricted first- and second-order markov meshes (Bixby *et al.*, 1967). These had some degree of success, but we know little more about texture after finding out the results of these experiments than before because they did not try to specifically define, characterize, or model texture. They only used some general mathematical transformation which assigns numbers to the transformed image in a non-specific way. (For an excellent discussion of the pitfalls encountered in the indiscriminant use of Fourier transform, including second-order statistical problems, see Bremerman, 1968).

Darling and Joseph, in 1968, suggested a set of textural features based on the mean and variance of the image grey tones, the relative frequencies of occurrence of each grey level, and the entropies of the conditional distributions of grey levels in the horizontal and vertical directions. They classified different cloud formations and lunar landscapes using these measures. The main shortcoming of the features proposed by Darling and Joseph is that the features have to be computed without performing grey tone normalization of the imagery. Without grey tone normalization, two images of the same scene will produce different sets of textural features due to variations in the grey tones of the two images produced by variations in lighting, lens, film developer, and digitizer. Hence features of two images of the same scene will

produce a different classification result, which is undesirable.

Recent attempts to extract textural features have been limited to developing algorithms for extracting specific image properties such as coarseness and presence of edges. Many such algorithms have been developed and tried on special imagery. The subjective parameters, such as the selection of thresholds, associated with the techniques do not enable them to be generalized to imagery other than the ones processed by the authors.

Recently, Rosenfeld and his co-investigators have presented a set of procedures for extracting textural features for pictorial data. Rosenfeld and Troy (1970) use a procedure based on the grey-tone differences of adjacent image elements and the auto-correlation of the image grey-tone values for obtaining a measure of the texture "coarseness". Later, Rosenfeld and Thurston (1971) described a procedure for detecting boundaries separating regions which differ in texture coarseness. Texture measures based on neighborhood operations on the image grey tones have been proposed by Gerdes (1970), and Joseph and Vigilone (1971). Moore, in his 1971 paper, describes a set of complexity measures which are related to the image texture. Procedures for detecting textural properties such as lines and dots have also been suggested by other investigators, see, for example, the proceedings of Computer Image processing and Recognition (1972). Before applying these procedures to pictures other than the ones processed by the authors of the respective papers, the investigator has to make a choice of the method appropriate to the picture in question as well as the selection of parameters for the particular method.

We are presenting in this paper a general procedure for extracting textural properties of blocks of image data. These features are calculated in the spatial domain and the statistical nature of texture is taken into account in our procedure, which is based on the assumption that the texture information in an image I is contained in the overall or "average" spatial relationship which the grey tones in the image have to one another. We compute a set of spatial grey-tone dependence probability distribution matrices for a given image block and suggest a set of 32 textural features which can be extracted from each of these matrices. These features contain information about such image textural characteristics as homogenity, linear grey-tone dependencies (linear structure), contrast, amount and nature of boundaries present, and the complexity of the image. It is important to note that the number of operations required to compute anyone of these features is proportional to the number of resolution cells in the image block. It is for this reason that we call these features quickly computable.

We also investigate the usefulness of textural features for categorizing or classifying image blocks. We developed and tested a land-use classification algorithm using the spectal and textural features of 624 image blocks derived from an ERTS image over the California coast line. The land-use categories identified consisted of coastal forest, wood lands, annual grass lands, water bodies, urban areas, and small and large irrigated fields. Up to 70% of the image blocks were identified correctly based on the textural features alone compared to an accuracy of 74 to 77% for a scheme based only on the spectral characteristics of the 4-band MSS image set. When the combined textural and spectral features were used as inputs to the classifier, up to 83.5% of the image blocks were identified correctly.

II. TEXTURAL FEATURES

Spatial Grey-Tone Dependence Matrix

Let $L_x = \{1, 2, \ldots, N_x\}$ and $L_y = \{1, 2, \ldots, N_y\}$ be the x and y spatial domains and $L_y \times L_x$ be the set of resolution cells. Let $G = \{1, 2, \ldots, N_g\}$ be the set of possible grey tones. Then a digital image I is a function which assigns some grey tone to each and every resolution cell; $I: L_y \times L_x \rightarrow G$.

An essential component of our conceptual frame work of texture is a matrix, or more precisely, four closely related matrices from which all the features for texture are derived. These matrices are termed angular grey-tone spatial dependence matrices.

We assume that the texture information in an image I is contained in the overall or "average" spatial relationship which the grey tones in image I have to one another.

More specifically, we shall assume that this texture information is adequately specified by the matrix of relative frequencies $P(i,j,d,\theta)$ which two neighboring resolution cells separated by distance d and having an angular relationship θ occur on the image, one with grey tone i and the other with grey tone j (see Fig. 1). Such matrices of spatial grey-tone dependence frequencies are a function of the angular relationship between the neighboring resolution cells as well as a function of the distance between them. Formally, for angles quantized to 45° intervals the unnormalized frequencies are defined by:

$$P(i,j,d,0^\circ) = \#\{((k,1),(m,n)) \in (L_y \times L_x) \times (L_y \times L_x), k-m=0, |1-n|=d, I(k,1)=i, I(m,n)=j\}$$

$$P(I,j,d,45^\circ) = \#\{((k,1),(m,n)) \in (L_y \times L_x) \times (L_y \times L_x), (k-m=d, 1-n=-d) \text{ or } (k-m=-d, 1-n=d), I(K,1)=i, I(m,n)=j\}$$

$$P(i,j,d,90^\circ) = \#\{((k,1),(m,n)) \in (L_y \times L_x) \times (L_y \times L_x), |k-m|=d, 1-n=0, I(k,1)=i, I(m,n)=j\}$$

$$P(i,j,d,135^\circ) = \#\{((k,1),(m,n)) \in (L_y \times L_x) \times (L_y \times L_x), (k-m=d, 1-n=d) \text{ or } (k-m=-d, 1-n=-d), I(K,1)=i, I(m,n)=j\},$$

where the notation $\#\{\cdot\}$ denotes the number of elements in the set $\{\cdot\}$. These spatial grey-tone dependence matrices can be normalized by dividing each entry in the matrix by the total number of resolution cell pairs used in computing the entries. Details of the procedure used in computing the entries in the matrices may be found in Haralick (1971).

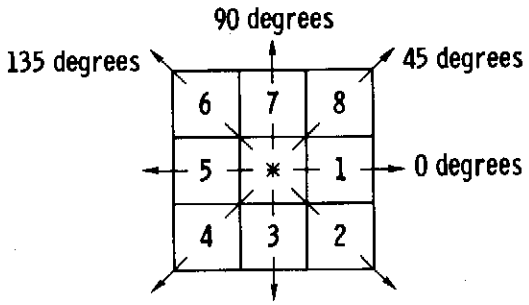


FIG. 1. Resolution cells nos. 1 and 5 are the 0-degree (horizontal) nearest neighbors to resolution cell'*', resolution cells nos. 2 and 6 are the 135-degree nearest neighbors, resolution cells 3 and 7 are the 90-degree nearest neighbors, and resolution cells 4 and 8 are the 45-degree nearest neighbors to'*'. (Note that this information is purely spatial and has nothing to do with grey-tone values.)

Before computing the entries in the P matrix, it is necessary that the image by quantized. Quantization serves two purposes. First, in a quantized image N_g will be small and hence all the entries in the P matrix can be meaningfully computed without using many more resolution cell pairs than the number contained in a small subimage. Secondly, quantization provides grey-tone normalization on the imagery. The data which the sensors or instruments produce are not always in the kind of normalized form with which it makes sense to work. For example, many sensors or measuring instruments produce relative measurements, i.e., the measurements are correct up to an additive or multiplicative constant. Despite calibration efforts, this is particularly true for the camera-film-digitizer system which produce the digital magnetic tape containing the digitized image. Variations in lighting, lens, film, developer, and digitizer all combine to produce a grey-tone value which is an unknown but usually monotonic transformation of the "true" grey-tone value. Under these conditions we would certainly want two images of the same scene, one image being a grey-tone monotonic transformation of the other, to produce the same results from the pattern recognition process. It is easy to show that normalization by equal probability

quantizing guarantees that images which are monotonic transformations of one another produce the same results. It should be realized that something is not gained for nothing. The normalization is achieved by sacrificing the detailed grey information. After quantizing to 16 levels, for example, an image which originally had 128 grey tones would only have 16 quantized grey tones and if equal probability quantizing were used, then the histogram of the quantized image would be uniform. Details of an algorithm which performs equal probability quantization may be found in Haralick (1971).

Textural Features

From each of the grey-tone dependency matrices we extract a set of 32 textural features. The equations which define these features are given in the appendix of this paper. For illustrative purposes we define three of these features here:

$$f_1 = \sum_{i=j}^{N_g}\sum_{j=1}^{N_g}\left(\frac{P(i,j)}{R}\right)^2;$$

$$f_2 = \sum_{n=0}^{N_g-1} n^2 \left\{\sum_{|i-j|=n}\left(\frac{P(i,j)}{R}\right)\right\}; \qquad (2)$$

$$f_3 = \frac{\sum_{i=j}^{N_g}\sum_{j=1}^{N_g}\frac{ijP(i,j)}{R} - \mu_x\mu_y}{\sigma_x\sigma_y}$$

where, R is the number of resolution cells pairs, and μ_x, μ_y and σ_x, σ_y are the means and standard deviations of the marginal distribution P_x and P_y obtained by summing the rows and columns of $P(i,j)/R$. The notation $P(i,j)$ denotes the $(i,j)^{\text{th}}$ entry in one of the angular spatial grey-tone dependence matrices $P(i,j,d,\theta)$. For different values of d and θ, we get a set of values for each of the above features.

To explain the significance of these features, let us consider the values they take on for two different land-use category images. Figure 2 shows the digital print-out of two sub-images from the California frame. The image shown in 2(a) belongs to the grass land category and image shown in Figure 2(b) is mostly water. Values of the features f_1, f_2, and f_3 are also shown for these images in Fig. 2.

The angular second moment features (ASM), f_1, is a measure of homogeneity of the image. In a homogeneous image, such as shown in 2(b), there are very few grey-tone transitions occurring with high relative frequencies. Hence, the P matrix for this image will have fewer entries of large magnitude. For an image like the one shown in Figure 2(a), the P matrix will have a large number of small entries and hence the ASM feature which is the sum of squares of the entires in the P matrix will be smaller. A comparison of the ASM values given below the images in Fig. 2 shows the usefulness of the ASM feature as a measure of the homogeneity of the image.

The contrast feature f_2 is a difference moment of the P matrix and is a measure of the contrast or the average amount of local variations in the image. Since there is a large amount of local variations in the image 2(a) compared to the image shown in 2(b), the contrast feature for the grassland image has consistently higher values compared to the water body image.

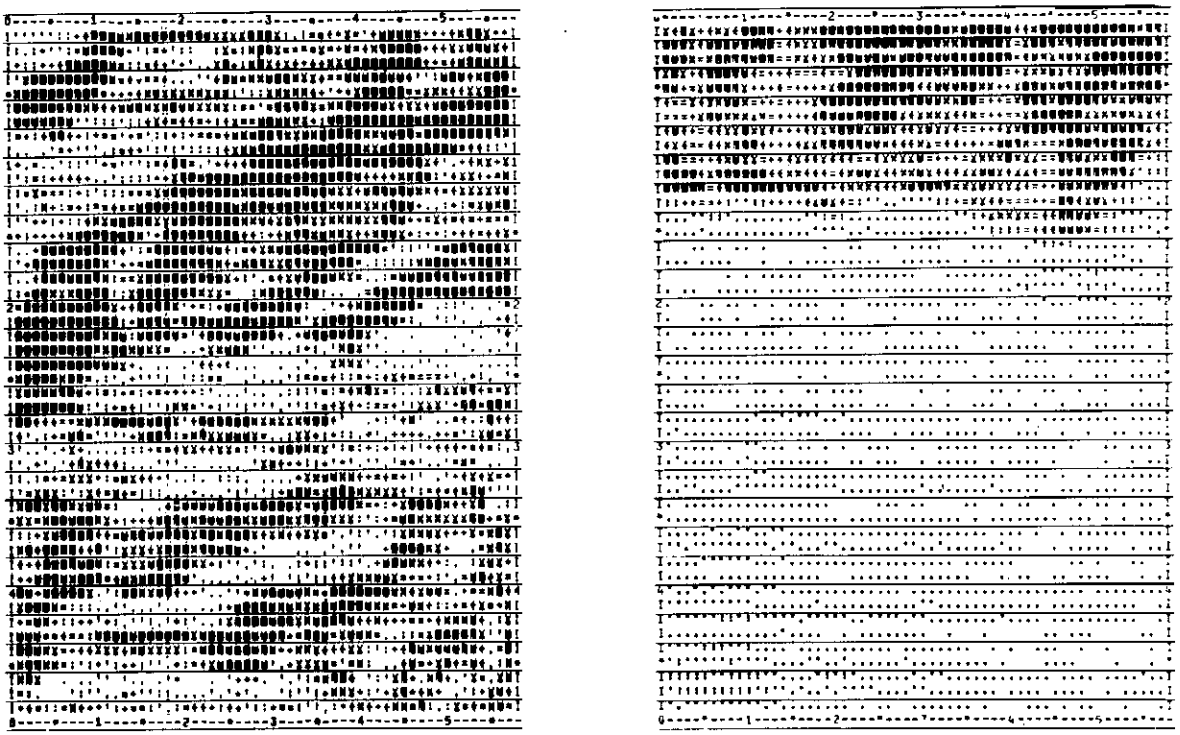
The correlation feature, f_3, is a measure of linear grey-tone dependencies in the image. For both the images shown in Fig. 2, the correlation feature is somewhat higher in the horizontal (0°) direction, along the line of scan. The water body image consists mostly of a constant greytone value for the water plus some additive noise. Since the noise is mostly uncorrelated, the correlation feature for the water body image have lower values compared to the grassland image. Also, the grassland image has a considerable amount of linear structure along 45° lines across the image, and hence the value of the correlation feature is higher along this direction compared to the values for 90° and 135° directions.

The various features which we suggest are all functions of distance and angle. The angular dependencies present a special problem. Suppose image A has features a,b,c,d for angles 0°, 45°, 90°, and 135° and image B is identical to

A except that B is rotated 90° with respect to A. Then B will have features c,d,a,b, for angles 0°, 45°, 90°, and 135°, respectively. Since the texture of A is the same as the texture of B, any decision rule using the angular features a,b,c,d must produce the same results for c,d,a,b or for that matter b,c,d,a (45° rotation) and d,a,b,c (135° rotation). To guarantee this, we do not use the angularly dependent features directly. Instead, we use two functions of a,b,c,d, their average, and their range which are invariant under rotation. The textural features used in our classification study were computed for four angles and for a distance of one.

III. Spectral Features

The spectral features used in our study con-



a. Grassland

Angle	ASM	Contrast	Correlation
0°	.0128	3.048	.8075
45°	.0080	4.011	.6366
90°	.0077	4.014	.5987
135°	.0064	4.709	.4610
Avg.	.0087	3.945	.6259

b. Water Body

ASM	Contrast	Correlation
.1016	2.153	.7254
.0771	3.057	.4768
.0762	3.113	.4646
.0741	3.129	.4650
.0822	2.863	.5327

FIG. 2. Textural features for two different land use category images.

sisted of the mean and variance of the greytone values within each image block. These features were computed for each of the four MSS bands, thus yielding a set of 8 spectral features for each image block. The mean and variance of the spectral data have been used extensively in the past for crop classification studies based on MSS data.

IV. Land Use Classification Studies

Data Set

All of the data used in our present study were derived from parts of the ERTS-I image frame 1002-18134. The date of the flight was July 25, 1972, and the center coordinates of the frame were 37.291N, 120.935W. The area of coverage includes segments of Oakland, San Jose Urban area, and the Monterray Bay on the coast line of California. Of the four image strips in this frame, all of strip one and one-half of strip three were digitally processed. The second MSS band (MSS-5) image and the area processed are shown in Fig. 3.

As a first step in the digital processing, the image area processed was divided into a total 648 sub-images of size 64 X 64 resolution cells (each sub-image covering a ground area of 7.5 square miles). For each of the sub-images, the digital spectral data from the second MSS band (MSS-5) were extracted. The imagery data for each sub-image was then normalized using an equal probability quantizing algorithm, and the 32 textural features for each sub-image were then computed. The spectral features for each image block were computed using the digital data from all four sensor bands.

Ground truth for each of the sub-images were obtained with the help of photointerpreters working on the MSS images and the color composite image. A total of seven land use categories could be identified reasonably well on the image, and for 624 of the 648 sub-image blocks we could determine a distinct land-use category. The set of 7 land-use categories used in our study consist of the following.

Categories of Land Use

(1) Coastal Forest: This area can be typified as a dense temperate rainforest sited on the windward side of the Coast Ranges of California. It is composed of needleleaf and broadleaf evergreen trees which generally form a continuous cover.

(2) Woodlands: These are areas to the leeward side of the Coast Ranges. At higher elevations, these woodlands consist of oaks, mostly evergreen, with varying but less continuous cover. At lower elevations these oak woods grade into chapparal which provides decreasing cover with decreasing elevation until it gives way to annual grasslands.

(3) Annual Grasslands: This is an area of non-native annual grasses which have already completed their life cycle by the end of July (turned brown). This area is primarily, like the preceding two categories, quite mountainous. This is also the natural vegetation which would be found in the Great Valley of California except for anthropogenic vegetation.

(4) Urban: Centers of human activity are extremely important as a landscape component within this frame. The primary urban agglomerations within the frame are segments of Oakland, San Jose urban area, and parts of the Sacramento-Stockton urban areas. Unlike the preceding categories, one of the main features of urban areas is non-uniformity. A variety of textural patterns as well as a variety of tones, are associated with urban complexes.

(5) Large Irrigated Fields: for practical purposes all field agriculture in the Great Valley is based on irrigation. The areas assigned to this category would most likely contain field crops such as cotton, alfalfa, and other crops readily adaptable to high mechanization. Most of the tree crops and vineyards would also be in this category.

(6) Small Irrigated Fields: These areas would contain high value low mechanization crops typified by vegetables.

(7) Water: Although almost self-explanatory as a term, the category in this instance is used to include: (a) ocean, (b) lakes – natural and man-made, and (c) standing water (in fields and on flood plains).

Out of the 648 sub-images in the frame, the photointerpreters helped us to find a distinct land use ground truth category for a total of 624 sub-images. Due to cloud cover and other

ambiguities the ground truth for 24 sub-images could not be positively identified.

Classification Algorithm

The problem of developing an algorithm for identifying the land use categories of sub-image blocks from an ERTS MSS image set can be stated as follows. A set of N measurement pairs (X_1,θ^1), (X_2,θ^2), ... (X_N,θ^N) are given as learning observations. A vector measurement (pattern) X_i where the components of X_i are the spectral and textural features of a sub-image block, comes from an image block whose land use category θ^i is known. θ^i is one of the R land use categories ($R = 7$, in our study) c_1, c_2, ... c_R. Based on the set of learning observations, we want to develop an algorithm for identifying the land use category of a sub-image block based on the measurement (pattern) vector X it produces.

In a widely used algorithm (Fukunaga 1972), Fu and Mendel (1972), Miesel (1972), the pattern space is partitioned into a number of

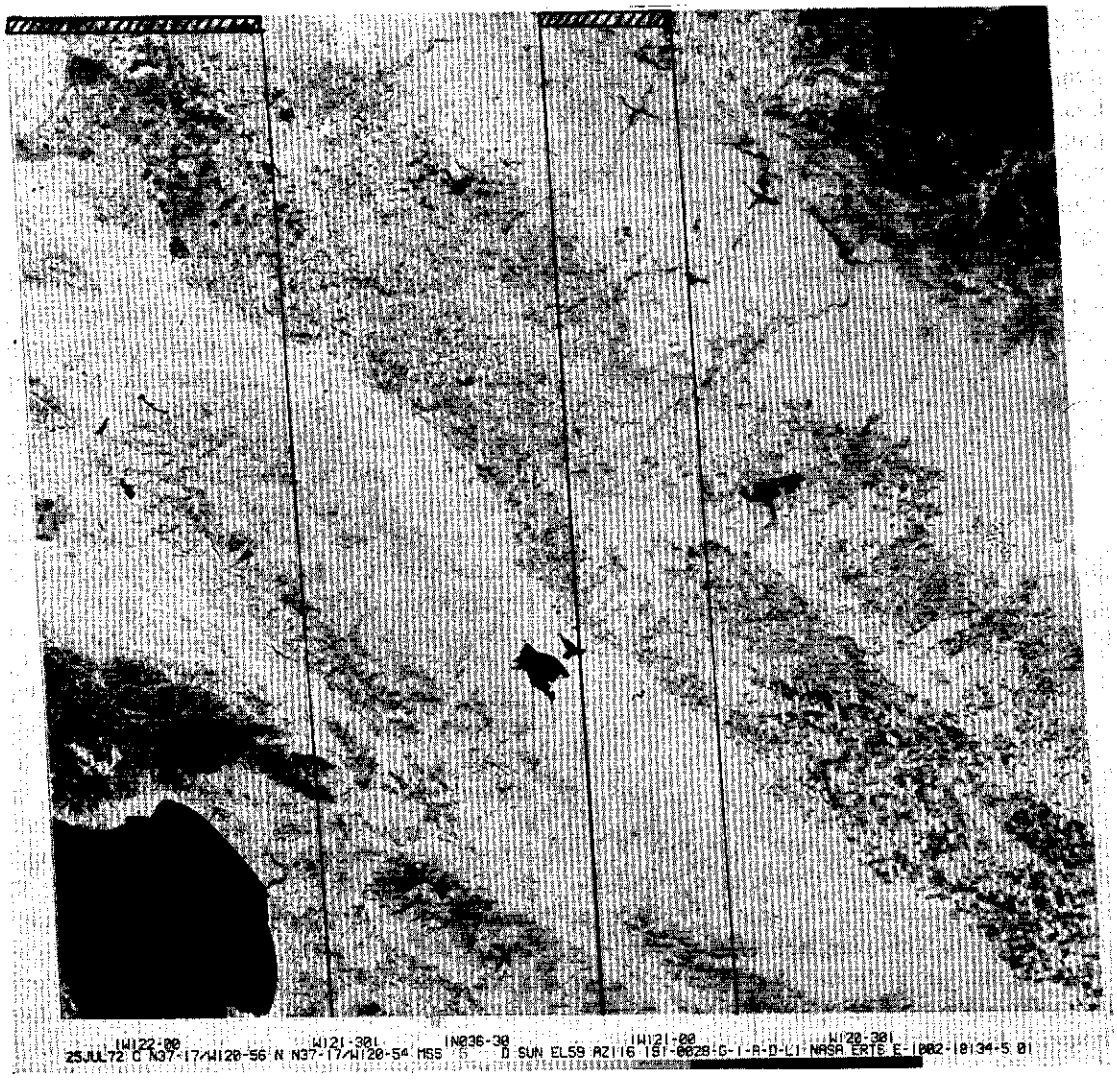


FIG 3. Two strips whose boundaries are shown on the top were processed from this ERTS MSS image set (1002-18134).

regions using a set of hyperplanes (decision boundaries) whose locations are determined by the sample patterns. Each region is dominated by sample patterns of a particular category. When a new pattern is presented for identification, it is assigned a category depending on the region in which it belongs. If the new pattern X is located in a region dominated by sample patterns of category c_j, then X is classified as coming from category c_j.

For the multicategory problem involving N_R categories, a total of $N_R(N_R - 1)/2$ hyperplanes are used to partition the pattern space. These hyperplanes are defined by a set of weight vectors $W_{ij}(i = 1,2, \ldots, N_R, j = 1,2 \ldots, N_R\ j>i)$ which separates the sample patterns belonging to the i^{th} and j^{th} categories. A least-square algorithm given in Fukunaga (1972), (chapter 4, pp 100-101) was used to obtain the weight vectors. After the location of the hyperplanes are determined, the classification of new patterns is done as follows. For each category c_i, the number of hyperplanes, V_i, which gives a positive response when the new pattern X is presented are determined using

$$V_i = \sum_{\substack{j=1 \\ j\neq i}}^{N_R} \frac{|W_{ij}^T Z| + W_{ij}^T Z}{2|W_{ij}^T Z|}, \qquad (3)$$

$$i = 1, 2, \ldots, N_R$$

where Z is the augmented pattern vector obtained by adding a component of value 1 to X, i.e.,

$$Z = \begin{bmatrix} 1 \\ X \end{bmatrix} \qquad (4)$$

X is assigned to category c_j if $V_j = \max\{V_i\}$. If there is a tie between categories c_m and c_n, then X is assigned to c_m if $W_{mn}^T Z < 0$ or to c_n if $W_{mn}^T Z \geqslant 0$. This classification algorithm has the invariance property that scale factor changes (transformation of the pattern space using a diagonal transformation matrix) do not affect the results of classification. Several modifications of the linear discriminant function method and a multitude of other classification procedures may be found in the references cited.

Land-Use Classification Experiments

The textural and spectral feature vectors for each of the 624 sub-images were randomly divided into training and test sets. The classification algorithm was developed using the information contained in the training set and the samples in the test set were assigned to one of the seven possible land use categories. The accuracy of classification was obtained by comparing the category assigned by the classifier with the known ground truth category. The standard deviation, σ, associated with the estimate of the accuracy of a classifier based on the classification of N test samples is given by

$$\sigma = \sqrt{\frac{\epsilon(1-\epsilon)}{N}}, \qquad (5)$$

where ϵ is the true classification accuracy.

Experiment No. 1: In this experiment, a land use classification scheme was developed using the 32 textural features of the sub-images on MSS-5 band. Fifty per cent of the samples were arbitrarily selected and used for developing the classification algorithm and the algorithm was tested on the remaining samples. There were seven land use categories and the overall identification accuracy of the classifier on the test samples was 67.5%. With 70% of the samples used for training, the accuracy of the classifier was 70.5%.

Experiment No. 2: In this experiment, the classification algorithm was developed and tested using the 8 spectral features. With 50% of the samples in the test set (N = 310), the accuracy of the classifier was 77% on the test set and with about 30% of the samples in the test set (N = 190), the accuracy of the classifier was 74%. It may be somewhat surprising to see the accuracy of the classifier decrease as the number of training samples is increased. This is mainly due to the randomness of the estimators of the classifier accuracy. The standard deviation of the estimator can be obtained from equation (5) as 2.45% for N = 310, and 3.15% for N = 190, if the true value of the classifier accuracy is assumed to be 75%.

Experiment No. 3: Eight textural features (consisting of the mean and range of features

TABLE 1

Contingency Table for Land Use Classification of Satellite Imagery. Number of Training Samples = 314. Number of Test Samples = 310. Accuracy of Classification on Training Set = 84.0%. Accuracy of Classification on Test Set = 83.5%

	ASSIGNED CATEGORY							
TRUE CATEGORY	Coastal forest	Woodlands	Annual grasslands	Urban area	Large irrigated fields	Small irrigated fields	Water	Total
Coastal Forest	23	1	2	0	0	0	1	27
Woodlands	0	17	10	0	1	0	0	28
Annual Grasslands	1	3	109	1	1	0	0	115
Urban area	0	3	10	13	0	0	0	26
Large Irrigated Fields	1	2	6	0	37	2	0	48
Small Irrigated Fields	0	0	4	0	3	24	0	31
Water	0	0	0	0	0	0	35	35
Total	25	26	141	14	42	26	36	310

f_1, f_2, f_3 and f_6) were combined with eight spectral features and the algorithm developed and tested using the combined features. The overall accuracy of the land-use classification algorithm was 83.5% for test set sizes of 30% and 50%. The contingency table for the classification is shown in Table 1. The results of this experiment show that an improvement in the classification accuracy can be obtained by using a combination of spectral and spatial features.

V. Discussion

The use of spectral, spatial (textural), and contextual features in combination has been discussed in numerous situations by various investigators. However, this type of classification using the combined features has not been pursued sufficiently in the past. We have presented an approach to bring together spectral and spatial processing of remotely sensed image data. The results of our study shows the usefulness of using both spectral and textural characteristics of ERTS MSS data for developing classification procedures. The textural features we have developed had been used in a variety of other image classification tasks including land use classification from aerial photographs (Haralick, Shanmugam and Dinstein, 1972), classification of photomicrographs of sand stones (Haralick and Shanmugam, 1972), and the classification of chest radiographs (Krueger *et al.*, 1972).

The 83.5% accuracy we have achieved on automatic land-use mapping may be less than the accuracy achievable by human interpreters. The difference in accuracy between the automated procedure and human interpreters can be attributed to the lack of contextual features as inputs to the automated classifier, since image interpreters rely heavily on contextual features, in addition to the spectral and textural features of the image. Further research is needed to develop a combination of spectral, textural, and contextual features for the automatic classifier.

The authors would like to thank NASA for providing the financial support for this study under Contract No. NAS5-21822 to the Remote Sensing Laboratory of the University of Kansas Center for Research Inc.

REFERENCES

Bell, C. B. (1962), Mutual Information and Maximum Correlation Measures of Dependence. *Annals of Mathematical Statistics, Vol. 43, pp. 587-595.*

Bixby, R., G. Elerding, V. Fish, J. Hawkins, and R. Loewe (May, 1967), *Natural Image Computer Final Technical Report,* Vol. 1, Philco-Ford Corporation Aeronutronic Division, Newport Beach, California, Publication No. C-4035.

Bremerman, H. J. (July, 1968), Pattern Recognition, Functionals, and Entropy, *IEEE Trans. on Biomedical Engineering,* Vol. BME 15.

Chevallier, R., A. Fontanel, G. Gray, and M. Guy (July, 1968), "Application Du Filtrage Optique a L'Etude Des Photographies Aeriennes," *XI Congress International De Photogrammetric.*

Darling, E. M., and R. D. Joseph (1968), "Pattern Recognition from Satellite Altitudes," *IEEE Trans. on Systems Science and Cybernetics,* Vol. SSC-4, pp. 38-47.

Fu, K. S., and J. M. Mendel (1972), *Adaptive Learning and Pattern Recognition Systems,* Academic Press, New York.

Fukunaga, K. (1972), *Introduction to Statistical Pattern Recognition,* Academic Press, New York.

Gerdes, J. (1970). *Automatic Target Recoenition Device (ATRD).* Technical Report, Rome Air Development Center, Griffiss AFB, Rome, New York.

Haralick, R. M. and D. Anderson, (November, 1971), *Texture-Tone Study With Applications to Digitized Imagery,* CRES Technical Report No. 182-2, University of Kansas Center for Research, Inc., Lawrence, Kansas.

Haralick, R. M., K. Shanmugam, and I. Dinstein (August, 1972), "On some quickly Computable Features for Texture," *Proc. Symp. Computer Image Processing and Recognition.* University of Missouri, Vol. 2, pp. (12-2-1)-(12-2-8).

Haralick, R. M. and K. Shanmugam (1973), "Computer Identification of Reservoir Sand Stones," to appear in the *IEEE Trans. Geosci. Electron.*

Joseph, R. D., and S. S. Vigilone (1971), "Interactive Image Analysis," in the *Proc. Symp. on Two-Dimensional Image Processing* (Columbia, Mo.) pp. (7-1-1)-(7-1-16).

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

Kruger, R. P., W. B. Thompson and A. F. Turner (1972), *Automated Computer Diagnosis of Pneumoconiosis from the Standard Posterior Anterior Chest Radiographs,* University of Southern California, School of Elect. Eng., Technical Report.

Linfoot, E. H. (1957), "An Information Measure of Correlation," *Information and Control,* Vol. 1, pp 85-89.

Miesel, W. (1972), *Computer Oriented Approaches to Pattern Recognition Systems,* Academic Press, New York.

Moore, D. J. H. (1971), "A Theory of Form," on *Intern. J. Man-Machine Studies,* Vol. 3, pp 31-59.

Proc. Symp. Computer Image Processing and Recognition, University of Missouri, Columbia, August 24-26, 1972.

Rosenfeld, A. and E. Troy (June, 1970) *Visual Texture Analysis,* University of Maryland Computer Science Center, Technical Report 70-116.

Rosenfeld, A., and M. Thurston (September, 1970), *Visual Texture Analysis,* 2, University of Maryland Computer Science Center, Technical Report 70-129.

Received March 1, 1973; revised June 12, 1973

APPENDIX A: Textural Features

We suggest a set of 32 textural features which can be extracted from each of the spatial greytone dependance matrices. The following equations define these features:

Notation

$p(i,j) - (i,j)^{\text{th}}$ entry in a normalized spatial grey-tone dependence matrix. $(p(i,j) = P(i,j)/\sum_i \sum_j P(i,j)$

$p_x(i) -$ i^{th} entry in the marginal probability matrix obtained by summing the rows of $p(i,j)$.

$$(p_x(i) = \sum_{j=1}^{N_g} p(i,j)$$

$$p_y(j) = \sum_{i=1}^{N_g} p(i,j)$$

$N_g =$ number of distinct grey levels in the quantized image.

$$p_{x+y}(k) = \sum_{\substack{i=1 \\ i+j=k}}^{N_g} \sum_{j=1}^{N_g} p(i,j); k = 2, 3, \ldots 2N_g,$$

$$p_{x-y}(k) = \sum_{\substack{i=1 \\ |i-j|=k}}^{N_g} \sum_{j=1}^{N_g} p(i,j); k = 0, 1, \ldots, N_g - 1.$$

$\sum_i$ and $\sum_j$ denote $\sum_{i=1}^{N_g}$ and $\sum_{j=1}^{N_g}$ respectively.

Textural Features

(1.) Angular second moment:

$$f_1 = \sum_i \sum_j \{p(i,j)\}^2.$$

(2.) Contrast:

$$f_2 = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{\substack{i=1 \\ |i-j|=n}}^{N_g} \sum_{j=1}^{N_g} p(i,j) \right\}.$$

(3.) Correlation:

$$f_3 = \frac{\sum_i \sum_j ijp(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y},$$

where μ_x, μ_y, σ_x, and σ_y are the means and standard deviations of p_x and p_y.

(4.) Sum of squares: Variance

$$f_4 = \sum_i \sum_j (i-\mu)^2 p(i,j).$$

(5.) Product moment: Covariance

$$f_5 = \sum_i \sum_j (i-\mu)(j-\mu)p(i,j).$$

(6.) Inverse Difference Moment:

$$f_6 = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j).$$

(7.) Difference Moment:

$$f_7 = \sum \sum (i-j)^2 p(i,j).$$

(8.) Sum Average:

$$f_8 = \sum_{i=2}^{2N_g} ip_{x+y}(i).$$

(9.) Sum Variance:

$$f_9 = \sum_{i=2}^{2N_g} (i-f_8)^2 p_{x+y}(i).$$

(10.) Sum Entropy:

$$f_{10} = -\sum_{i=2}^{2N_g} p_{x+y}(i) \log \{p_{x+y}(i)\}.$$

(11.) Entropy:

$$f_{11} = -\sum_i \sum_j p(i,j) \log p(i,j).$$

(12.) Difference Variance:

f_{12} = variance of p_{x-y}

(13.) Difference Entropy:

$$f_{13} = -\sum_{i=0}^{N_g-1} p_{x-y}(i) \log \{p_{x-y}(i)\}.$$

(14, 15.) Information measures of Correlation:

$$f_{14} = \frac{HXY - HXY1}{\max\{HX, HY\}}$$

$$f_{15} = \sqrt{1 - \exp[-2.0(HXY2 - HXY)]},$$

where

$$HXY = -\sum_i \sum_j p(i,j) \log p(i,j)$$

HX, HY = entropies of p_x and p_y

$$HXY1 = -\sum_i \sum_j p(i,j) \log \{p_x(i)p_y(j)\}$$

$$HXY2 = -\sum_i \sum_j p_x(i)p_y(j) \log \{p_x(i)p_y(j)\}.$$

(16.) Maximal Correlation Coefficient:

$$f_{16} = \sqrt{\text{Second largest eigen value of } Q}$$

where

$$Q(i,j) = \sum_k \frac{p(i,k)p(j,k)}{p_x(i)p_y(k)}.$$

These measures of correlation have some desirable properties which are not brought out in the rectangular correlation measure f_3. For details, see references Linfoot (1957) and Bell (1962).

In the set of 16 measures we have defined above, some of the measures are strongly correlated with each other. A feature selection procedure selection procedure may be applied to select a subset or linear combinations of the above 16 measures. Referring to equation (1) in the text, for a given distance d we have 4 angular spatial grey-tone dependency matrices. Hence we obtain a set of 4 values for each of the above 16 measures. The mean and range of each of these 16 measures, averaged over the 4 values, comprise the set of 32 features which are used as inputs to the classifier.