Ross W. Leamer
Daniel A. Weber
Craig L. Wiegand*
USDA Agricultural Research Service
Weslaco, TX 78596

Pattern Recognition of Soils and Crops from Space†

Minimum-distance-to-mean and maximum-likelihood-ratio algorithms are both useful. The more variables used, the better the recognition results.

Introduction

This report discusses the relative effectiveness of some of the commonly used computer analysis techniques to extract land use (crop identification) information from digitized aerial photographs. Comparisons between minimum distance to the mean (MDM) and maximum likelihood ratio (MLR) algorithms show that either can successfully recognize land-use patterns. The number and combination of densities chosen to represent the land-use categories affects the relative ranking of the two algorithms. Final classification accuracy was not affected by the density units in the base data. Arbitrary digital counts from a density measuring system were as good as standardized optical density units so long as the two were linearly related. Conversion of densities of color film to analytical densities degraded the classification accuracy more often than it improved the results from integral densities. Combining optical density values from black-and-white films and from color film into a single data set resulted in the highest correct identification (84.9 per cent for the MDM and 100 per cent for the MLR technique).

Abstract: *Some commonly used computer analysis techniques to get land use information from digitized aerial photographs are compared. Density readings of color-IR film and of multispectral black-and-white films are used to make the comparisons. All methods increased in accuracy as more densities were included in the data set. The highest accuracy was obtained if density values from both color film and from black-and-white films were combined into a single data set. Using arbitrary units from density measuring equipment gave as good results as converting to standard optical density units or to analytical density values.*

Aerial photographs have been a standard tool of geologists, geographers, foresters, engineers, etc., for many years. The utility of such images has been well established. Remote sensing in general, and the space satellites in particular, have greatly increased the interest in digital analysis and interpretation of aerial photographs.

Latham[2] suggested in 1959 that electronic devices such as scanning densitometers could be used for the quantitative measurement and analysis of geographic features and

* Soil Scientist; Mathematics teacher, Donna Independent School District, formerly physical Science Technician; Research Leader, respectively.

† Contribution from Soil & Water Research, Southern Region, Agr. Res. Service, USDA, Weslaco, Texas. This study was supported in part by NASA under Contract No. R-09-038-002.

land use in photographs in a manner that would "permit the use of mechanical and/or electronic sorting and computing equipment for organizing and statistically evaluating the data." Rosenfeld[5], using the methodology introduced by Latham, conducted research into the possibilities of using a flying-spot scanner in identifying terrain types recorded on black-and-white film.

Because of the tremendous amount of photography and other imagery being generated, only a fraction of the images are being fully interpreted. The computer is a vital component of most of the analysis and interpretation systems currently being used. More efficient and effective use of computers will allow more of the information in remote sensing imagery to be extracted.

This report discusses the relative effectiveness of some commonly used computer analysis techniques to extract land use (crop identification) information from digitized aerial photographs.

Procedure

Multispectral terrain photographs of the Imperial Valley of California obtained by the Apollo-9 astronauts (experiment SO-65) served as the basic data set for this study. These data were used because they represent data generated by a low-resolution system, and thus they should be a rigid test for classification techniques. Extensive ground truth of the area collected by NASA and the University of Michigan has been reported by Spansail *et al.*[6]. The data set includes optical counts and integral and analytical film optical density readings from color-IR film (multiemulsion) and from multispectral black-and-white (multibase) films exposed simultaneously. The films were exposed in four 70-mm Hasselblad cameras with 80-mm focal-length lenses mounted to view the same area and connected so that the shutters were tripped simultaneously. Three of the cameras contained black-and-white films and had filters to obtain exposures to the green, the red, and the reflective infrared portions of the spectrum. The fourth camera contained color-infrared film which has three dye layers which are sensitive to approximately the same wavelength bands as those passing the filters on the black-and-white cameras.

Optical densities of specific areas having known ground truth were measured on the films by a Joyce, Loebl microdensitometer.*

* Trade names and company names are included for the benefit of the reader and do not imply an endorsement or preferential treatment by the U.S. Department of Agriculture of the product listed.

Optical density to white light was measured on the black-and-white films, and density to white, red, green, and blue light was measured on the color film. Computer pattern recognition of known crop and soil features using differences between optical densities read from these films have been reported by Wiegand *et al.*[7] Their classification and discriminations were based on the minimum distance to the mean (MDM) pattern-recognition technique. In this experiment, the same data were used to test the effectiveness of other forms of the data and other pattern-recognition techniques.

Each system designed to measure optical density of photographic film uses a different optical system and usually a different device to measure light passing through film. Some systems produce readings directly in optical density values, whereas others produce outputs that must be converted to optical density units. If measurements from one machine are to be compared with readings from another machine, readings from both machines must be converted to a common unit of measurement. Optical density units are used to compare the variables being studied in this report. Equation 1 converts the output from the Joyce, Loebl microdensitometer to optical density values:

$$OD = (OC - Base)\,(0.0082) \times 0.71 \qquad (1)$$

where OD is the optical (integral) density, OC is the optical count (machine units), $Base$ is 108.5 for White light, 106.1 for *R*ed light, 107.6 for *G*reen light, and 109.5 for *B*lue light.

Another objective of this experiment was to compare integral densities with analytical optical densities for identifying crop and soil conditions. Theoretically, analytical density quantifies the density of each dye layer independently of the other dye layers in a color film. Although each layer is primarily sensitive to energy of a certain wavelength band, each layer has an effect on the other layers, and all three layers affect each density measurement. Thus optical counts, or integral densities, measure the response to colored light passing through the three layers of the film; analytical densities express the independent contribution of each layer to the film density. Conversion from integral density to analytical density is designed to eliminate the overlap of response of the film layers. Two sets of equations were used to convert integral density of the color film to analytical density, one proposed by Kodak[1], the other by EG&G[3].

Kodak conversion of integral to analytical density,

$$R' = R(1.368) - G(0.321) + B(0.022) - 0.054$$

$G' = -R\ (0.120) + G\ (1.238) - B\ (0.147) + 0.006$

$B' = -R\ (0.039) - G\ (0.213) + B\ (1.154) + 0.016;$

EG&G conversion of integral to analytical density,

$R' = R\ (1.0188) - G\ (0.0258) - B\ (0.0101)$

$G' = -R\ (0.217) + G\ (1.1026) - B\ (0.0705)$

$B' = -R\ (0.0326) - G\ (0.2107) + B\ (1.1356)$

where R', G', and B' are the analytical densities of the film to red, green, and blue light, respectively, and R, G, and B are the respective integral densities for red, green, and blue light calculated in this experiment by Equation 1.

Optical counts were used for all discrimination tests reported here except for the tests comparing integral and analytical densities and those comparing optical counts with integral densities. Where integral densities were used, Equation 1 was employed to convert optical counts to integral densities.

The tests covered in this report can be grouped into four major categories:

- The MDM algorithm was compared with the maximum likelihood ratio (MLR) algorithm using both actual optical count data and the optical count differences. The two algorithms were also compared where the data were converted to the principal axis factor scores (PAFS) for the MLR algorithm.
- Optical counts (original machine units) were compared with integral optical densities in all algorithms.
- The effectiveness of analytical densities of colored film was compared with integral densities in discriminating among surface features.
- The effectiveness of combining density values from all films into a single data set was evaluated.

Information representing five crop and soil categories from 53 fields for which ground truth was known was the data set for this study. To eliminate any possible effect of field size in the discrimination programs, each field was represented by an average of several density readings. Consequently each density reading for each field is a mean for the field. Details on site selection and film density measurements are available in Wiegand *et al.*[7]

Computer programs were developed and modified to perform either the MDM or MLR pattern recognition algorithm within the capabilities of an IBM 1800 computer system. Each program calculated crop and soil standards from the ground truth, classified each sample into one of the five categories, gave an output vector showing the classification of each field along with a recognition matrix which identified correct and incorrect classifications and gave a per cent recognition figure. The programs gave recognition results using all possible combinations of densities and density differences for a field. PAFS cannot be calculated for single or pairwise comparisons because conversions to

TABLE 1. PER CENT CORRECT IDENTIFICATION BY MDM AND MLR ALGORITHMS OF 53 FIELDS BY OPTICAL COUNTS (OC) AND OPTICAL COUNT DIFFERENCES (OCD) FROM BOTH COLOR AND BLACK-AND-WHITE FILMS FROM SO-65 EXPERIMENT.

Comparison	Single-Level Comparisons MDM	Single-Level Comparisons MLR
	Per cent correct identification	
Color film		
Red minus blue (R-B)	62.2	60.4
Red minus green (R-G)	52.6	58.5
Green counts (G)	50.9	56.6
White minus red (W-R)	49.0	60.4
Blue counts (B)	49.0	50.9
Red counts (R)	49.0	49.0
White counts (W)	45.3	45.3
White minus blue (W-B)	34.0	39.6
Green minus blue (G-B)	30.2	41.5
White minus green (W-G)	34.0	34.0
Black-and-white films		
Red band (R)	66.0	71.7
Green band (G)	64.2	45.3
IR minus red (IR-R)	62.3	66.0
Green minus IR (G-IR)	60.4	41.5
IR band (IR)	39.6	49.0
Green minus red (G-R)	37.7	26.4

TABLE 2. PER CENT CORRECT IDENTIFICATION BY MDM AND MLR ALGORITHMS OF 53 FIELDS BY PAIRWISE COMPARISONS OF OPTICAL COUNTS (OC) AND OPTICAL COUNT DIFFERENCES (OCD) FROM BOTH COLOR AND BLACK-AND-WHITE FILMS FROM SO-65 EXPERIMENT

Comparison	Pairwise Comparisons MDM	MLR
	Per cent correct identification	
Color film		
R, B	73.6	62.3
R, G	71.7	71.7
R-B, G-B	67.9	71.7
R-G, G-B	66.0	71.7
R-G, R-B	66.0	71.7
W-R, R-G	67.9	69.8
W-R, R-B	67.9	66.0
W-G, R-G	66.0	69.8
W, R	64.2	67.9
W-R, W-G	62.3	69.8
W-R, R-B	60.4	66.0
W-B, R-G	60.4	64.2
W-R, W-B	58.5	66.0
W-B, R-B	56.6	66.0
G, B	52.8	69.8
W-R, G-B	52.8	66.0
W, G	52.8	56.6
W, B	49.0	54.7
W-G, G-B	45.3	58.5
W-B, G-B	43.4	58.5
W-G, W-B	41.5	58.5
Black-and-white films		
G, R	73.6	71.7
IR, R	67.9	73.6
G-IR, IR-R	67.9	67.9
G-R, IR-R	67.9	67.9
G-IR, G-R	60.4	67.9

PAFS requires a minimum of three original densities to generate the two scores required to use PAFS in the MLR algorithm.

RESULTS

MDM VS MLR COMPARISONS

The MDM algorithm was tested against the MLR algorithm using both optical counts and optical count differences as the original variables. Tests were run on individual comparisons between optical counts obtained from white light through the three black-and-white films; these were compared with the three possible differences of these optical counts. Similar tests were made of the three optical counts from the color film and the three differences in optical counts. Table 1 gives the summary of the per cent correct identification of each single level comparison. For the color film, MLR yields the higher recognition in six out of ten trials, whereas for the six black-and-white film comparisons, each of the MDM and MLR methods is better than the other in three instances.

Table 2 summarizes the results of trials made with all possible pairwise combinations of optical counts and optical count differences, both from color film and three black-and-white films.

Table 3 gives per cent correct identification for all three level combinations of optical counts and optical count differences except combinations of the three differences that were linearly dependent on the other two: for instance, in the set of differences *W-R*, *W-G*, and *R-G*, *W-R* = *(W-G)* − *(R-G)*. The reason for excluding such sets is that the MLR algorithm recognizes the dependence of one of the variables on the others and either does not calculate PAFS or, in trials without PAFS generation, creates covariant matrices whose determinates are less than, or very close to, zero, thereby rendering any discrimination results useless.

The three level comparisons of Table 3 show that for color film, the optical counts gave better identification than optical count differences for both the MDM and MLR tech-

TABLE 3. PER CENT CORRECT IDENTIFICATION BY MDM AND MLR (WITH AND WITHOUT PAFS) ALGORITHMS OF 53 FIELDS BY THREE-LEVEL, FOUR-LEVEL, AND SEVEN-LEVEL COMPARISONS OF OPTICAL COUNTS AND OPTICAL COUNT DIFFERENCES FROM BOTH COLOR AND BLACK-AND-WHITE FILMS FROM SO-65 EXPERIMENT.

Comparison	MDM	MLR w/o PAFS	MLR w/PAFS
	Per cent correct identification		
Three level comparisons			
Color film			
W, R, G	75.5	71.7	71.7
R, G, B	77.4	69.8	66.0
W, R, B	73.6	67.9	60.4
W-R, R-G, G-B	69.8	64.2	62.3
W-G, R-G, R-B	69.8	64.2	62.3
W-G, R-G, G-B	69.8	64.2	62.3
W-R, R-G, R-B	67.9	64.2	62.3
W-G, W-B, R-B	67.9	64.2	62.3
W-B, R-G, R-B	67.9	64.2	62.3
W-B, R-B, G-B	66.0	64.2	62.3
W-R, W-G, W-B	66.0	64.2	62.3
W-R, W-G, R-B	66.0	64.2	62.3
W-B, R-G, G-B	64.2	64.2	62.3
W-G, W-B, R-G	64.2	64.2	62.3
W-R, R-B, G-B	64.2	64.2	62.3
W-R, W-B, R-G	62.3	64.2	62.3
W-R, W-B, G-B	60.4	64.2	62.3
W-B, R-B, G-B	60.2	64.2	62.3
W-R, W-G, G-B	58.5	64.2	62.3
W, G, B	50.9	64.2	60.4
Black-and-white films			
G, IR, R	73.6	75.5	75.5
Four level comparisons			
Color film			
W, R, G, B	73.6	73.6	71.7
Seven level comparisons			
All films			
W, R, G, B, G, R, IR	84.9	100.0	81.1

niques and that principal axis factor score pre-processing tended to decrease correct identifications using the MLR.

The optical counts for the green, red, and infrared-sensitive black-and-white films and the integral densities of the color film to red, white, green, and blue light yield equal identifications (73.6 to 75.5 per cent) regardless of recognition algorithm used.

Considering the results of the 89 trials comparing MDM vs. MLR algorithms, the MLR method showed a slight advantage in being more accurate in 51.1 per cent of the trials. However, in the 23 instances where three or more densities were used, the MDM method proved superior 64.0 per cent of the time. In the 64 cases where only one or two densities were used, the MLR method gave better results 68.8 per cent of the time. If three or more density values were used, identifications were better with both MLR and MDM than if one or two variables were used.

COMPARISONS INCLUDING PAFS

For the MDM VS. MLR (PAFS) comparison, 23 tests were run using optical counts taken from color film with white, red, green, and blue light. Table 3 includes the results obtained using all four readings from the color film and using all seven densities for each field from both the color film and the three black-and-white films. The MDM algorithm using the seven densities correctly identified 45 of the 53 fields for a correct identification percentage of 84.9. Generation of PAFS in the MLR algorithm with seven densities resulted in correctly identifying 43 of the 53 fields for a correct identification percentage of 81.1. However, where the MLR algorithm was used without conversion to PAFS, all 53 fields were identified correctly. The identifications obtained using all seven film densities for each field indicated that the information content of black-and-white, and color-infrared films complement each other.

TABLE 4. CLASSIFICATION RESULTS OBTAINED USING THE MDM PATTERN RECOGNITION ALGORITHM ON COLOR-FILM DATA COMPARING IDENTIFICATION FROM OPTICAL COUNTS WITH IDENTIFICATION FROM INTEGRAL OPTICAL DENSITIES. BOTH TESTS USED THE PAIRWISE DIFFERENCES OF THE DENSITIES OF COLOR FILM TO WHITE, RED, GREEN, AND BLUE LIGHT FOR THE CATEGORIES ALFALFA (1), BARLEY (2), SUGARBEETS (3), BARE SOIL (4), AND SALT FLATS (5).

		Recognition vector		
Integral optical densities				
1115111214	3212222222211	3323223333	44144114144	515554455
Alfalfa	Barley	Sugarbeet	Bare soil	Salt flat
Optical counts				
1115111214	3212222222211	3323223333	44144114144	555554455
Alfalfa	Barley	Sugarbeet	Bare soil	Salt Flat

OPTICAL COUNT VS OPTICAL DENSITY COMPARISON

An experiment was run to see if the conversion of optical count data from the Joyce, Loebl microdensitometer to optical density values affected the results obtained by the various identification algorithms. The conversion involves three constants: base count, wedge factor, and step count (equation 1). Eight trials were run to compare identification from optical counts with identification using integral optical densities in the MDM algorithm. In each of the trials, the identification from optical count data was identical to that from the corresponding optical density data, both on fields correctly identified and on a point-by-point basis. Thus, it was concluded that the conversion to optical density is unnecessary where data from only one machine are used.

Table 4 shows the output recognition vector obtained by Wiegand *et al.*[7] for color film compared with the corresponding output vector obtained in this study. Both used all six possible pairwise differences of the four readings from each field. The only difference between the data for the two tests is that one used optical densities and the other used optical counts. The results are identical except for one point. This is presumed to be due to the fact that different computers were used for the two trials and a round-off error occurred.

Wiegand *et al.* used optical density differences and obtained an overall classification accuracy of 68 per cent from color film and 72 per cent from black-and-white films (Table 5). In this study, classification accuracy was increased to a maximum of 77 per cent using optical counts from the color film with the MDM algorithm. Optical counts with the MLR algorithm increased the per cent correct classification from black-and-white films to 75 from 72 per cent.

INTEGRAL DENSITY VS ANALYTICAL DENSITY COMPARISONS

Both the Kodak and the EG&G conversions from integral density to analytical density were tested against integral densities using all noninterdependent one-, two-, and three-level combinations of both these den-

TABLE 5. NUMBER AND PER CENT CORRECT IDENTIFICATION BY CROP AND SOIL CATEGORIES FOR BOTH COLOR AND BLACK-AND-WHITE FILMS FROM SO-65 EXPERIMENT (After Wiegand *et al.*, 1971.)

Crop	No. of fields	Ektachrome IR film Correctly identified Number	Per cent	Black-and-white films Correctly identified Number	Per cent
Alfalfa	10	7	70	4	40
Barley	13	9	69	8	62
Sugarbeet	10	7	70	9	90
Bare soil	11	7	64	9	82
Salt flats	9	6	67	8	89
Total	53	36		38	
Overall per cent			68		72

sities and their pairwise differences in the MLR algorithm. The outcome of the 26 trials was very nearly the same regardless of the conversion used (Table 6). All results were scored on the basis of one point for the conversion having the greater per cent of fields identified correctly with ties scored as one-half of a win and one-half as a loss. On this basis, integral densities came out ahead in 17 trials, behind in seven, and tied in two, for a winning percentage of 70.8. Here, too, the best identification was obtained with the larger number of density variables.

Conclusions

This study showed that both the MDM and MLR algorithms were useful for pattern recognition. Each algorithm had advantages under certain circumstances. The MDM algorithm was slightly more accurate where three or more variables were used, but the MLR algorithm proved to be better where less than three variables were used. The MDM algorithm takes considerably fewer formula steps and consequently less computer time[4] so its use should be considered seriously in any pattern recognition work where three or more variables are available and computer time is a factor.

This study also pointed out that the more variables used in the classification algorithm, the better the recognition results. Where density values from both multispectral black-and-white films and color film are available, all values should be used. Together they contained enough information to raise previously reported overall recognition accuracy of 68 per cent (color-IR film) and 72 per cent (black-and-white films) involving four and three independent optical density differences to 84.9 per cent (using all seven optical counts in the MDM algorithm) and to 100 per cent (using all seven optical counts in the MLR algorithm).

It made no difference to the final classification accuracy what density units were used as the base data. Arbitrary digital counts from a single density measuring system were as good as standardized optical density units so long as the two are linearly related.

Even conversion of densities measured on color film to analytical densities does not improve classification accuracy. Conversion to analytical densities in this study degraded the classification results more often than it improved the results from integral densities.

References

1. Fritz, Norman L., Private communication with John D. Tallant. May 4, 1972.
2. Latham, James P., Possible Application of Electronic Scanning and Computer Devices to the Analysis of Geographic Phenomena, *Report No. 1, NR 387-023* (Washington Office of Naval Research), August 1959.
3. Linnerud, Harold J., Private communication with Craig L. Wiegand, May 4, 1972.
4. Richardson, A. J., Torline, R. J., Weber, D. A., Leamer, R. W., and Wiegand, C. L., Computer Discrimination Procedures Using Film Optical Densities, *SWC Research Report 422* (USDA, Weslaco, Texas), December 1970.
5. Rosenfeld, Azriel, Automatic Recognition of Basic Terrain Types from Aerial Photographs,

Table 6. Per Cent Correct Identification by MDM and MLR Algorithms of 53 Fields by Optical Counts, Integral Optical Densities, and Analytical Optical Densities of Color Film from SO-65 Experiment.

	MDM	MLR			
				Analytical Density	
Comparison	Optical Counts	Optical Counts	Integral Density	EG&G	Kodak
		Per cent correct identification			
R, G, B	77.4	69.8	69.8	73.6	71.7
R, B	73.6	62.3	62.3	71.7	67.9
R, G	71.7	71.7	71.7	67.9	67.9
R-B, G-B	67.9	71.7	71.7	62.3	69.8
R-G, G-B	66.0	71.7	71.7	66.0	64.2
R-G, R-B	66.0	71.7	71.7	67.9	69.8
R-B	62.3	60.4	60.4	53.5	62.3
G-B	52.8	69.8	69.8	54.7	50.9
R-G	52.6	58.5	58.5	60.4	50.9
G	50.9	56.6	56.6	49.0	47.2
B	49.0	50.9	50.9	50.9	50.9
R	49.0	49.0	49.0	39.6	50.9
G-B	30.2	41.5	41.5	28.3	37.7

Photogrammetric Engineering, 27:115-132 (January 1962).

6. Spansail, N. *et al.* (1969), Imperial Valley Ground Truth for Apollo 9 Overflight of March 1969, *Univ. Mich. Inst. Sci. Technol. Rept. 2264-7-X*.
7. Wiegand, C. L., Leamer, R. W., Weber, D. A., and Gerbermann, A. H., Multibase and Multiemulsion Space Photos for Crops and Soils, *Photogrammetric Engineering*, 37: 147-156 (February 1971).

A.S.P. ANNOUNCES ORTHOPHOTO WORKSHOP III

ORTHOPHOTO WORKSHOP III, latest in the ASP series of symposia on the state-of-the-art in orthophotography, is scheduled for June 4-6, 1975.

Sponsored by the American Society of Photogrammetry, this year's event will be held at the El Tropicano Motor Hotel, San Antonio, Texas. The Society's Texas-Louisiana Region will host the workshop.

Several new orthophoto devices have come onto the market since the last workshop and many projects are underway or now complete, in which orthophotography plays a major role.

Richard T. Church, Workshop Chairman, indicates that workshop objectives are to 1) identify the state-of-the-art, 2) provide a forum for users of orthophoto equipment, and 3) to supply the buyer or potential buyer of orthophotos a clear understanding of the fundamentals and advantages of orthophotography and its many uses.

The Third workshop is to consist of six (6) technical sessions, (two each day) during which the invited technical papers will be discussed informally, along with a limited number of unsolicited papers. According to Dr. Robert T. Turpin and Dr. Robert Baker, Program Co-Chairmen, all accepted papers will be published in a bound volume.

No formal call for papers is planned. Anyone wishing to prepare an un-solicited paper should submit the following information:

1. The paper's title
2. Author's name, address and telephone number
3. An abstract of approximately 200 words

This information should be mailed to:

Dr. Robert T. Turpin
Civil Engineering Department
Texas A&M University
College Station, Texas 77843

The Co-Chairmen have indicated that invited papers will include the fundamentals, history, recent technical progress, user procedures and project descriptions of orthophotographic endeavours.

An exhibit area including both commercial and noncommercial exhibits will be open throughout the show. Manufacturers will be exhibiting new orthophoto equipment and recent projects will be featured in the noncommercial area.

Several other national organizations are cooperating in presenting the workshop. Members will receive more detailed information on this important technical meeting at a later date.

ISP Congress Newsletter

To ensure that your name is on the mailing list for the 1976 Congress Newsletter of the International Society of Photogrammetry, write to Mrs. A. Savolainen, Institute of Photogrammetry, Technical University of Helsinki, Otaniemi, Finland.