P2mit

"Made available under NASA sponsorship in the interest of early and wide dissemination of Earth Resources Survey Program information and without liability for any use made thereof."

E7.4-10.236
CR-136556

AUTOMATED THEMATIC MAPPING AND CHANGE DETECTION OF ERTS-A IMAGES

Nicholas Gramenopoulos
Optical Systems Division
Itek Corporation
10 Maguire Road
Lexington, Massachusetts 02173

(E74-10236) AUTOMATED THEMATIC MAPPING AND CHANGE DETECTION OF ERTS-A IMAGES Interim Report, Feb. - Jul. 1973 (Itek Corp.) 73 p HC $5.75 CSCL 08B G3/13

N74-15041

Unclas 00236

August 1973
Interim Report for Period February - July 1973

Prepared for
GODDARD SPACE FLIGHT CENTER
Greenbelt, Maryland 20771

Original photography may be purchased from:
EROS Data Center
10th and Dakota Avenue
Sioux Falls, SD 57198

AUTOMATED THEMATIC MAPPING AND CHANGE DETECTION OF ERTS-A IMAGES

Nicholas Gramenopoulos
Optical Systems Division
Itek Corporation
10 Maguire Road
Lexington, Massachusetts 02173

August 1973
Interim Report for Period February - July 1973

Prepared for
GODDARD SPACE FLIGHT CENTER
Greenbelt, Maryland 20771

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle	5. Report Date
Automated Thematic Mapping and Change Detection of ERTS-A Images	August 1973
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Nicholas Gramenopoulos PR 504	
9. Performing Organization Name and Address Optical Systems Division Itek Corporation 10 Maguire Road Lexington, Massachusetts 02173	10. Work Unit No.
	11. Contract or Grant No. NAS5-21766
12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

For the recognition of terrain types, spatial signatures are developed from the diffraction patterns of small areas of ERTS-1 images. This knowledge is exploited for the measurements of a small number of meaningful spatial features from the digital Fourier transforms of ERTS image cells containing 32 x 32 picture elements.

Using these spatial features and a heuristic algorithm, the terrain types in the vicinity of Phoenix, Arizona were recognized by the computer with a high accuracy. Then, the spatial features were combined with spectral features and using the maximum likelihood criterion the recognition accuracy of terrain types increased substantially.

It was determined that the recognition accuracy with the maximum likelihood criterion depends on the statistics of the feature vectors. Nonlinear transformations of the feature vectors are required so that the terrain class statistics become approximately

(continued)

17. Key Words (Selected by Author(s))	18. Distribution Statement
Interpretation Techniques Development, Digital Information Extraction Techniques, Classification and Pattern Recognition	

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price*
Unclassified	Unclassified	73 ~~63~~	5.75

*For sale by the Clearinghouse for Federal Scientific and Technical Information, Springfield, Virginia 22151.

TECHNICAL REPORT STANDARD TITLE PAGE (CONTINUED)

ABSTRACT (CONTINUED)

Gaussian. It was also determined that for a given geographic area the statistics of the classes remain invariable for a period of a month but vary substantially between seasons.

PREFACE

A. OBJECTIVE

The main objective of the investigation is to develop digital interpretation techniques such that earth resources can be recognized automatically from ERTS-1 images. Emphasis has been placed on three characteristics of the interpretation techniques: accuracy of recognition, efficiency and practicality of generating thematic maps and automation of the interpretation functions.

B. SCOPE OF WORK

Under the investigation, multispectral and spatial pattern recognition techniques are being developed so that the information in the ERTS-1 images can be properly exploited.

Initial analysis of ERTS-1 computer compatible tapes showed that the data could not be used in its original form, and a major portion of the investigation's effort had to be directed into reformatting, resampling and filtering of the data. These preprocessing operations were time consuming and have been completed.

It also became necessary to develop new spatial signatures that take into account the ERTS-1 data characteristics and particularly its ground resolved distance. This work has been completed, and it has been demonstrated that the spatial features that have been selected carry the necessary spatial information to identify terrain types.

According to the Modified Data Analysis Plan, the spatial features were combined with spectral features, and the accuracy of terrain recognition increased substantially.

The effects of the maximum likelihood criterion on the recognition accuracy were analyzed, and nonlinear transformations of the feature vectors were devised which increase the accuracy of recognition results appreciably.

The variations in terrain class statistics with the time of the year were also explored in order to determine the need to have statistics for each ERTS-1 image.

A new clustering algorithm was developed so that training data will not be required to determine the terrain class statistics.

C. CONCLUSIONS

Through the developed algorithms, it is possible to automatically recognize terrain types. By using the maximum likelihood criterion, it is possible to achieve high accuracy in the recognition process and some degree of automation since heuristic algorithms need not be developed for each georgraphical area.

The experimental results show that the accuracy of terrain recognition is affected by the statistics employed. The accuracy is best when the statistics are approximately Gaussian, and it is necessary to apply nonlinear transformations to the feature vectors to achieve approximately Gaussian distributions. The class statistics vary appreciably between seasons, such as summer and fall so that for each geographic area, it will probably be necessary to determine the seasonal variation in statistics. It seems, though, that once known, the statistics can be used repetitively for a number of years until substantial land use changes have occurred in the geographic area. The experimental results show that the spatial features selected contain the spatial information necessary to recognize terrain types.

The experimental results also showed that combining the spectral data with the spatial features increased the accuracy of the machine recognition.

D. RECOMMENDATIONS

The results of this investigation are applicable to the production of land-use maps. Anderson et al[10] have described three levels of land-use classification. The terrain classes recognized in this investigation fall in the first classification level. The second level requires the use of much higher resolution imagery than ERTS-1.

The machine recognition results of this investigation are very encouraging and suggest that the interpretation techniques have been developed can be directly applied to higher resolution data which will permit a subdivision of the major terrain classes and the generation of second level land-use maps.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES (CONTINUED)

LIST OF TABLES

1. INTRODUCTION

This is a Type II report for the period of February 1 - July 31, 1973. The purpose of the report is to describe the technical progress of the investigation for this period.

The main body of the report is divided into the following sections:

A. Input Data Operations, which describes the data handling operations necessary for the utilization of the data received from NASA. Particular attention is given to the ERTS Digital Tapes (Bulk CCT's) for which the digital preprocessing operations constitute a major part of the work done under this investigation.

B. Photointerpretation Tasks, which are necessary for the location of ground truth, familiarization with the test sites and the ERTS images and for evaluating the digital terrain type recognition operations.

C. Terrain Spatial Signatures. This section describes the work done to isolate signatures for the various types of terrain. The signatures that have been developed and means for efficiently measuring them are also described.

D. Terrain Type Recognition. This is an extensive section describing two algorithms for recognizing terrain and the results obtained when these algorithms were employed. The algorithms utilized both multispectral and spatial data.

E. Conclusions. This section describes the conclusions that can be drawn from the work performed.

F. Program for the Completion of the Investigation. The work planned for the remaining time interval of the investigation is presented in this section.

G. Recommendations.

2. INPUT DATA OPERATIONS

2.1 ERTS Images

Since February 1, 1973, a substantial number of 9-1/2 inch positive and color transparencies has been delivered to the project by NDPF. A system of checking and cataloging these images has been developed so that they can be utilized by the project efficiently.

Through the standing orders, the project receives black and white positive transparencies on 9-1/2 inch film. The images may contain a portion or an entire test site and are the first indication that a test site has been acquired on a specific date. Upon receipt of the transparencies, an overlay is produced photographically. The overlay contains all the annotations and coordinate marks of the MSS 5 (Red Band) image. Georgraphic coordinates and the test site coordinates are plotted on the overlay. (See Figure 2-1.) Using the overlay, it is possible to determine the percentage coverage of the test site.

At this point, a decision is made whether to submit a request for digital tapes. The decision is based on cloud cover, haze, percentage of test site photographed, as well as on previous acquisitions of the test site.

The ERTS images are also being utilized for the photointerpretation tasks (section 3) and for the development of terrain spatial signatures (section 4). Both of these operations require changes in scale which are accomplished very efficiently photographically. The photointerpretation tasks require enlargements while the spatial signatures require high resolution reductions in scale in order to obtain large diffraction patterns.

2.2 ERTS Digital Tapes

A number of digital operations are required to utilize the data available in the bulk Computer Compatible Tapes (CCT's). Each computer compatible tape delivered by NDPF contains one-quarter of an ERTS image. The four spectral bands (for an MSS image) are interleaved. In addition, the sampling interval (in meters) along a scan line is smaller than the sampling interval between scan lines. In other words, there are more picture elements along a kilometer on the surface of the earth parallel to a scan line (approximately east-west direction) than a kilometer normal to a scan line (approximately north-south direction). Furthermore, successive lines are slightly shifted to each other, owing to the earth's rotation in relation to the satellite velocity.

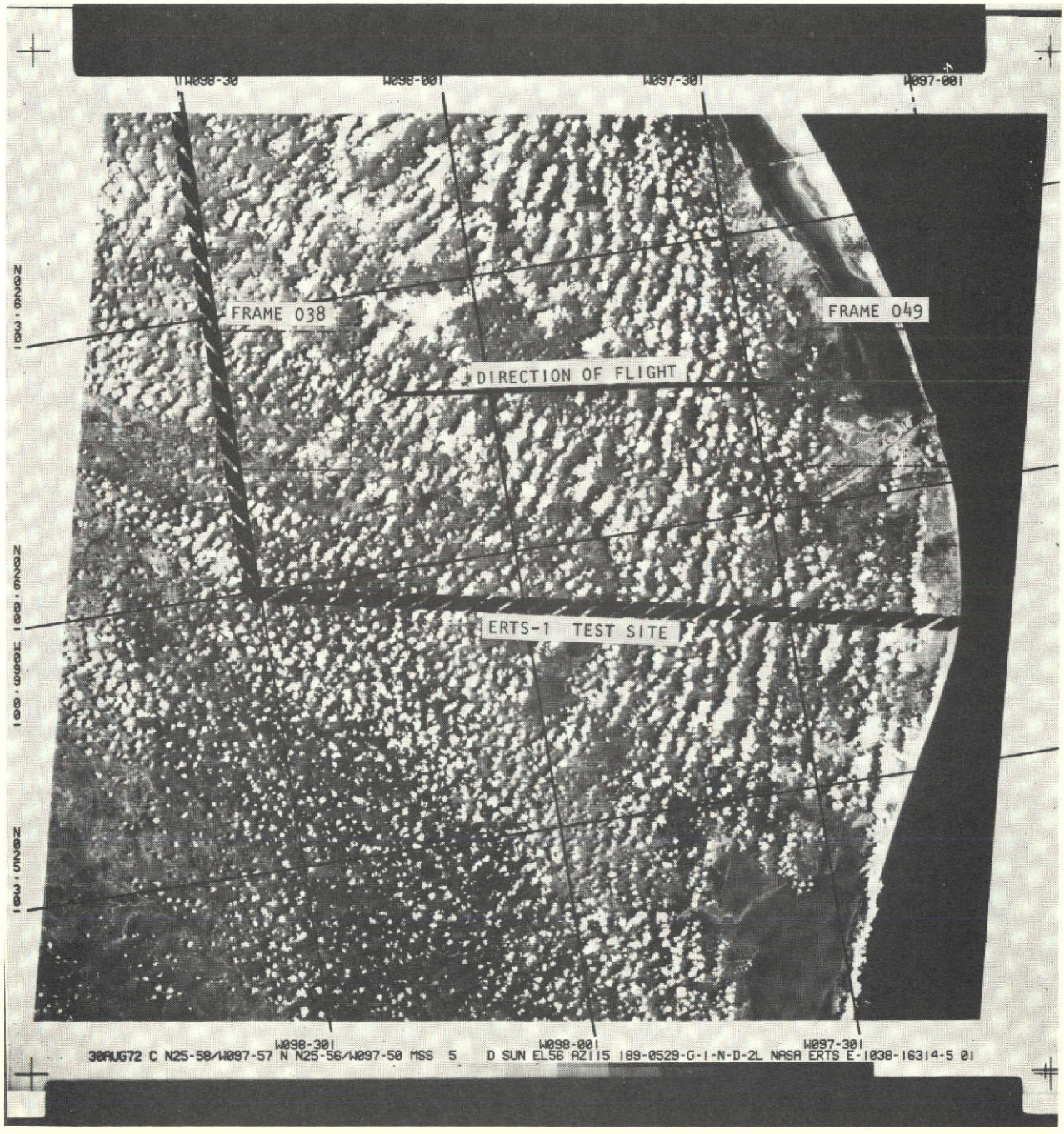


Fig. 2-1 — Annotated ERTS-1 image 1038-16314

The first preprocessing operation requires the separation of the MSS 4, 5, and 7 spectral bands into separate digital images consisting of 810 samples per scan line and 1,000 scan lines each. The 1,000 scan lines are selected from each CCT (which contains 2,340 lines) so that they cover approximately the north-south extent of a test site. To cover the east-west extent of a test site, it is necessary to combine data from two adjacent CCT's. So, the first preprocessing operation involves stripping data from two CCT's and producing six images (two in each spectral band), containing 810 x 1,000 samples each.

The second preprocessing operation combines adjacent images, so one is left with three images (one in each spectral band) containing 1,620 x 1,000 samples each. These images have a major distortion which is an enlargement of the images in the scanning direction (east-west, approximately) and results from the unequal sample spacing in the scanning and raster directions. Visually, the images look like oblique rather than vertical photographs. This image distortion causes artificial variations in the spatial signatures of the terrain types. For example, a square farm would appear as a rectangle or a skewed paralellogram depending on its orientation to the scanning direction. It is necessary, therefore, to resample each image in the scanning direction in order to equalize the sample spacing in the two orthogonal directions.

The third preprocessing operation involves resampling of the images in the scanning direction and reducing the samples from 1,620 samples per line to 1,170 samples per line. After this resampling operation, the scale of the images is approximately the same in all directions.

The fourth preprocessing operation records the digital images as transparencies in a laser beam recorder for visual examination. This is necessary since many images display missing lines and a few images have pronounced line structure. The line structure with periodicity of six lines seems to arise from inaccurate calibration of the MSS detectors (there are six detectors per spectral band).

The fifth preprocessing operation involves printing the digital data from selected scan lines to ascertain that the artifacts are actually present in the data.

The sixth preprocessing operation is the correction of the artifacts. Missing lines are replaced by the average of the two adjacent lines. The line structure is corrected by recalibrating the detectors. Each detector contributes one sixth the data (every sixth line) of each image. By integrating the data from each detector over the entire image, an estimate of the calibration errors can be made. First, the samples along each scan

line are summed up. Let the sum of samples along the jth line be s(j) where:

$$j = 1, 2, \ldots . \quad (996) \tag{2-1}$$

Then, one forms the six sums

$$S_i = \sum_{k=0}^{165} s(6k + i) = s(i) + s(6 + i) + \ldots s(990 + i) \tag{2-2}$$

where i = 1, 2, 3, 4, 5, and 6.

S_i is the sum of all the samples contributed by the ith detector. If all detectors were correctly calibrated, then:

$$S_1 = S_2 = S_3 = S_4 = S_5 = S_6 \tag{2-3}$$

To correct the detector calibration, the samples of the scan lines numbered (6k + i) where:

$$k = 0, 1, 2, \ldots 166 \text{ and } i = 1, 2, \ldots 6 \tag{2-4}$$

are multiplied by the ratio (S_1/S_i).

The correction in the calibration achieved by this method is only approximate but has worked very satisfactorily whenever it was applied.

It is obvious that the preprocessing operations on the digital data are quite extensive, involved and time consuming. According to the Modified Data Analysis Plan, data from nine ERTS images has been preprocessed and is stored on magnetic tapes in the form of 27 separate images. These

are being used for the terrain type recognition described in section 5.

The preprocessing operations have been completed for the following ERTS-1 images:

1031-17325 from Phoenix, Arizona
1049-17324 from Phoenix, Arizona
1103-17332 from Phoenix, Arizona

1040-18201 from Cascade Mountains
1077-18260 from Cascade Mountains

1015-17415 from Salt Lake, Utah
1069-17420 from Salt Lake, Utah

1070-17495 from Salton Sea, California
1034-17500 from Salton Sea, California

Figure 2-2 is an example of a preprocessed image that was recorded as a transparency.

2.3 Aircraft Photography

Low altitude photography (10,000 feet) and high altitude photography (~60,000 feet) from all the test sites has been delivered by NASA (the Johnson Spacecraft Center or the Ames Research Center). In general, the photography is very good and is being fully utilized in the photointerpretation tasks (section 3).

The handling operations on the aircraft images are all photographic and involve making prints, enlargements and transparencies. (See, also, reference 1.)

The aircraft photography is the primary source of ground truth for the test sites. Without it, it would be impossible to determine the accuracy of the digital terrain recognition results.

3. PHOTOINTERPRETATION TASKS

Each test site area has been subjected to photointerpretation for the purpose of determining ground level conditions. Aircraft photography is the primary data base for the photointerpretation tasks, and it is supplemented by maps of the test sites.

The photointerpretation involves making overlays and comparing the ERTS-1 images to the high altitude and the low altitude photography.

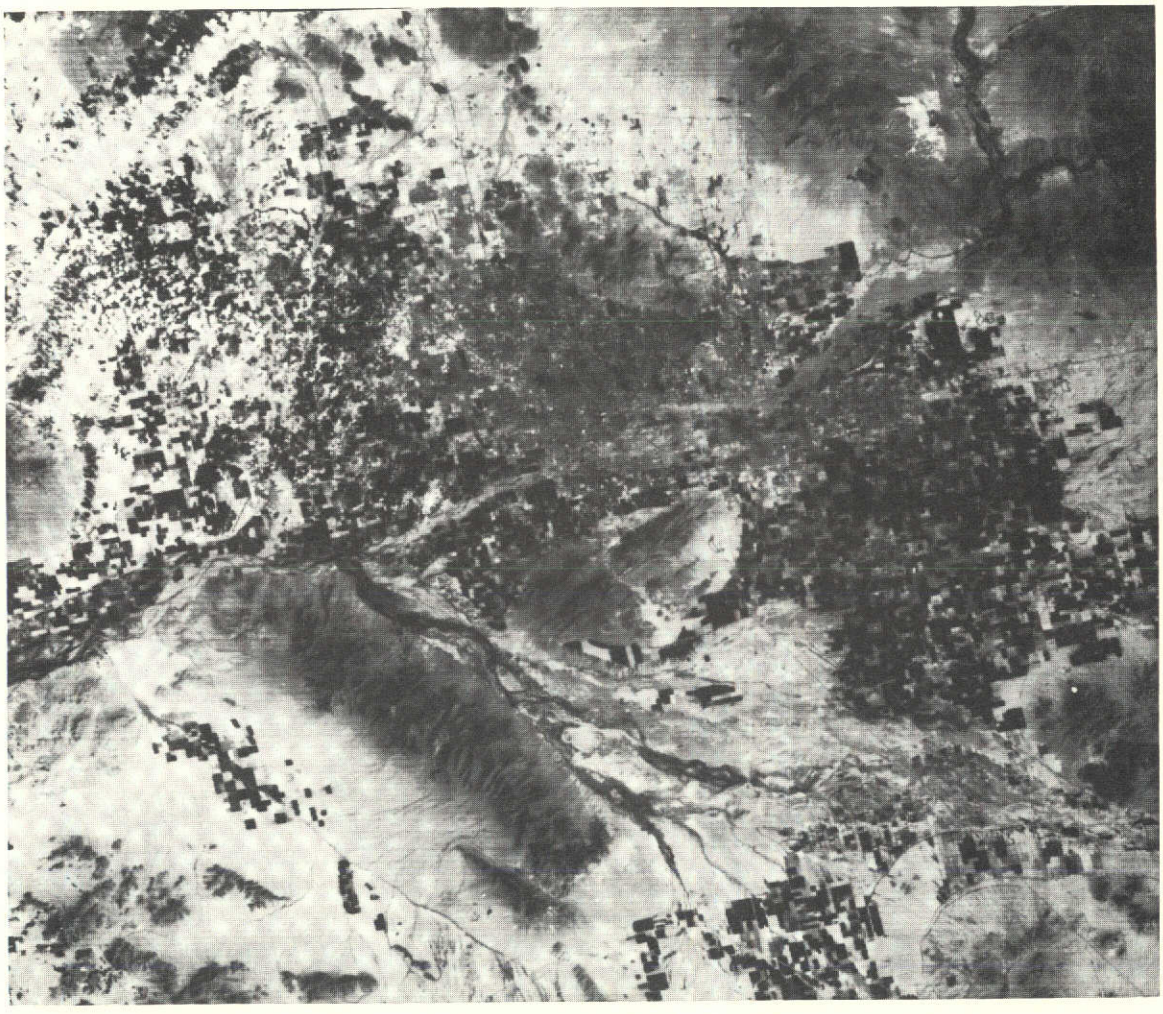


Fig. 2-2 — ERTS-1 image 1049-17324-5 recorded from digital data

The photointerpretation results are also compared to the digital interpretation results in order to determine the accuracy of the terrain recognition operations. (See Reference 1.)

4. TERRAIN SPATIAL SIGNATURES

In the ERTS-1 images, the basic unit which conveys information is the picture element. Each picture element corresponds to an area increment on the surface of the earth equal to the instantaneous field of view squared (about 80 meters square). Each area increment is characterized by a brightness value in each spectral band. These brightness values are the multispectral information conveyed by the picture element.

The first level of automatic interpretation is the classification of individual picture elements based on multispectral information. This level appears to be well developed as evidenced by the wide use of multispectral recognition software based primarily on the maximum likelihood criterion.

Most of the information in the ERTS images is contained in the distribution of the picture elements over the image area. This spatial information may be conveniently divided into the information carried by small localized areas defined as cells as well as the information carried by the distribution of the cells over the image.

The cell size that has been employed in this investigation is 32 x 32 picture elements which corresponds to about 2.5 x 2.5 kilometers on the surface of the earth. Each cell, therefore, is characterized by 1,024 picture elements.

Two questions, then, arise about each cell:

a. Which objects on the surface of the earth it might represent?

b. How could these objects be identified by manipulating these 1,024 numbers?

The first question is answered by the IFOV (80 meters) of the MSS scanner. The objects that are faithfully reproduced by this system are expected to be in the order of one kilometer. Examination of the ERTS photography shows that the objects that can be identified are various geographical features and terrain types, such as bays, lakes, rivers, mountains, urban areas, cultivated land, desert, forest, etc.

The second question is more difficult to anwer. It is desirable to develop signatures which uniquely characterize the terrain categories.

A signature is defined as a specific pattern of numbers.

It is also desirable that each terrain class be characterized by a signature containing less than ten numbers. Assume that such signatures were known. Then, to assign a cell to one of the terrain types, the 1,024 numbers describing the cell are combined to produce a vector with less than ten components which is known as a spatial feature vector. This vector is compared to the spatial signatures, and the cell is assigned to the terrain category whose signature most closely resembles the vector of the cell. Conceptually, the assignment of cells to terrain categories is simple, but there are some difficult practical problems in accomplishing it, such as:

a. The 1,024 numbers of the cell are quite redundant. Their replacement by a feature vector with less than ten components is a data reduction operation which may actually destroy information.

b. There is no theoretical treatment that allows one to select optimal terrain signatures.

c. Comparison of a feature vector to the spatial signatures requires the use of a criterion of "closeness". If the cell is to be assigned to the most probable terrain class, the statistics of each class must be known.

As far as is known, there have been two approaches for the determination of spatial signatures. One of these is to employ well known mathematical functions (as components of the feature vector), such as the variance of the cell numbers, and decide later by the results of the recognition whether these mathematical functions are reliable indicators of terrain type. The results of this approach have not been very satisfactory. The other approach consists of examining the Fourier transforms of cells and identifying patterns which are unique to each terrain type. This is the approach followed in this investigation.

4.1 Diffraction Pattern Analysis

4.1.1 Experimental Procedure

In the ERTS-1 images, the terrain types that one would like to recognize consist of:

Mountains	Clouds	Rivers
Desert or range	Urban areas	Bodies of water
Cultivated land	Hills	Transportation networks

Fourier transforms of images displaying one terrain type provide a means for isolating terrain signatures.

The Fraunhofer diffraction pattern of an image is related to its Fourier transform. It is well known that, if a transparency is introduced into the front focal plane of a lens and illuminated by a coherent plane parallel beam, the lens will form an image at its back focal plane whose amplitude distribution is the Fourier transform of the amplitude transmission of the transparency. Specifically,*

$$U_f(x_f,y_f) = \frac{A}{j\lambda f} \iint_{-\infty}^{+\infty} t_a(x_o,y_o)\exp[-j\frac{2\pi}{\lambda f}(x_o x_f + y_o y_f)]dx_o dy_o \qquad (4\text{-}1)$$

where $U_f(x_f,y_f)$ is the amplitude distribution at the back focal plane, x_f,y_f are the coordinates at the back focal plane, λ is the wavelength of light, f is the focal length of the lens, A is the amplitude of the illuminating beam, and $t_a(x_o,y_o)$ is the amplitude transmission of the transparency. The diffraction pattern $D(x_f,y_f)$ is the amplitude squared of $U_f(x_f,y_f)$:

$$D(x_f,y_f) = U_f(x_f,y_f)\cdot U^*(x_f,y_f) = |U(x_f,y_f)|^2 \qquad (4\text{-}2)$$

In addition:

$$t_a(x_o,y_o) = [T(x_o,y_o]^{1/2}e^{+j\Phi(x_o,y_o)} \qquad (4\text{-}3)$$

where $T(x_o,y_o)$ is the intensity transmission of the transparency for coherent light and $\Phi(x_o,y_o)$ is a phase function due to emulsion thickness that is unrelated to the image.

*J. W. Goodman, Introduction to Fourier Optics, McGraw-Hill Book Co., Inc., New York (1968), p. 86.

It is obvious from the above equations that the diffraction pattern is not equal to the Fourier transform of the image. However, the diffraction pattern is useful because it has the same general structure as the Fourier transform.

Noise or errors in the diffraction pattern associated with the phase function $\Phi(x_o, y_o)$ can be eliminated by inserting the transparency in a liquid gate filled with a refractive index matching fluid.

The optical bench employed is shown in Figure 4-1. The laser has a wavelength of 632.8 nanometers and its beam is expanded and filtered to produce a beam that is Gaussian in intensity. The primary lens forming the diffraction pattern has a focal length of 48 inches, to produce large diffraction patterns. The frequency scale of the diffraction pattern is proportional to the focal length of the lens and the wavelength of illumination.

To reduce the overall dimensions of the bench, the transparency is located near the lens rather than on its front focal plane. The result is that a quadratic phase factor now multiplies $U_f(x_f, y_f)$ but does not affect the diffraction patterns.

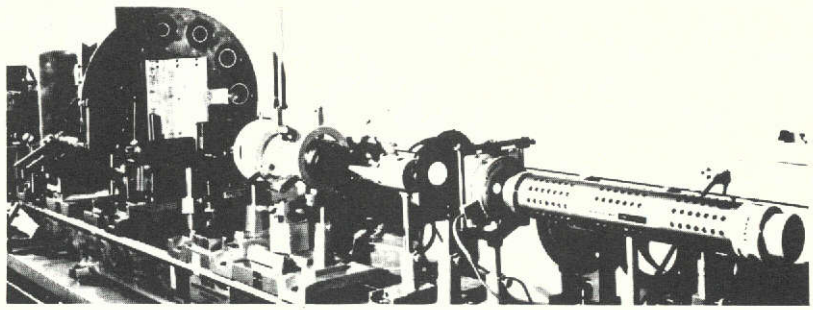
Liquid gates are not employed because they are inconvenient when diffraction patterns from a large number of images are to be obtained. The optical bench permits photographing of the diffraction patterns and the images simultaneously. It also allows photometric measurements of any part (rings or wedges) of the diffraction patterns. Such measurements of diffraction patterns have been obtained.

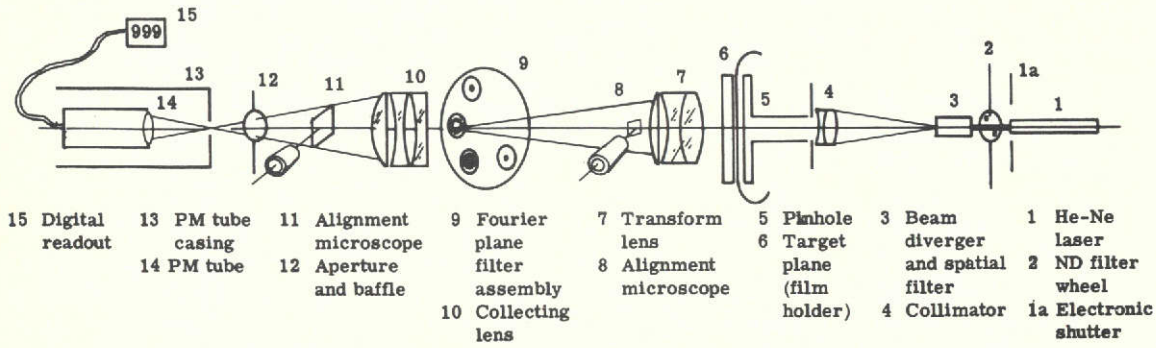
The ERTS images selected are described in Section 4.1.2. The images initially employed were positive transparencies on 9-1/2-inch-wide film. They produced small diffraction patterns. It was decided, therefore, to reduce the scale of the ERTS-1 images by a factor of three and develop the reduced images to a gamma of two to compensate for the square root factor in Equation 4-1. The maximum resolution in the reduced transparencies was estimated at about 25 line pairs per millimeter.

The diffraction patterns have some artifacts not related to the terrain images:

a. Rings due to the circular aperture employed to limit the image area being transformed; these rings are also present in a digitally computed Fourier transform.

b. Two frequency spots due to the line structure present in the ERTS images; the line structure is a characteristic of the multispectral scanner and electron beam recorder system.





15 Digital readout
13 PM tube casing
14 PM tube
11 Alignment microscope
12 Aperture and baffle
9 Fourier plane filter assembly
10 Collecting lens
7 Transform lens
8 Alignment microscope
5 Pinhole
6 Target plane (film holder)
3 Beam diverger and spatial filter
4 Collimator
1 He-Ne laser
2 ND filter wheel
1a Electronic shutter

Fig. 4-1 — Coherent optical bench

c. Artifacts due to the optical bench, such as lens aberrations, laser beam nonuniformity, phase modulation by the transparencies.

d. The central order is so bright that it usually saturates the film over a sizable area surrounding the central order and masks low frequency components.

To eliminate or suppress artifacts in the diffraction patterns, a mask is employed when photographing the diffraction patterns. The mask is itself the photograph of the diffraction pattern of an image area (from ERTS images) with no detail, such as areas of water from lakes or the ocean.

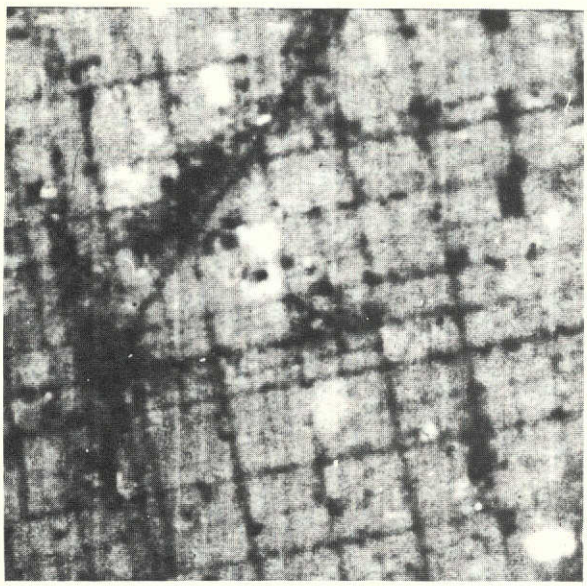
Figure 4-2 provides a comparison of a typical diffraction pattern with and without the mask. It is obvious that the mask enhances the low frequency structure of the diffraction pattern and suppresses artifacts. For this reason, these masks have been used consistently.

Rings in the diffraction patterns are associated with the aperture only, not with image detail within the aperture. The spacing (d) between adjacent dark bands in the Airy disk is $d = 1.22\lambda f/a$, where λ is the wavelength of light, f is the focal length of the lens, and a is the aperture diameter.

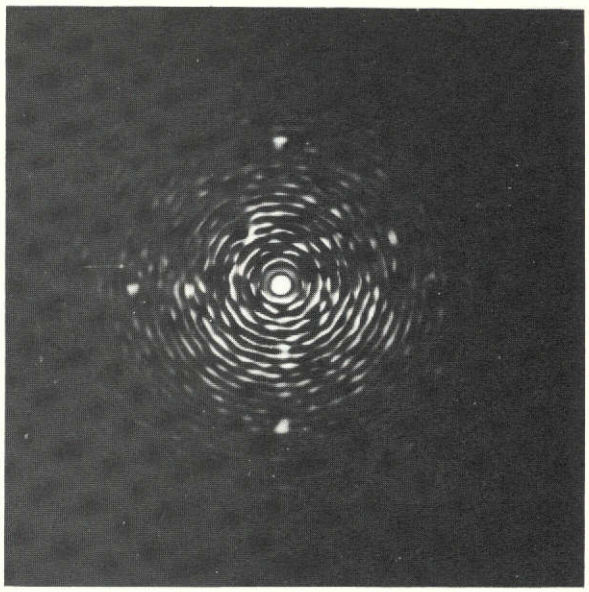
Increasing the aperture decreases the ring spacing and improves the resolution of the diffraction pattern. However, the central order becomes very intense and obscures low frequency components even though a Fourier plane mask is being employed. In addition, if the aperture is too large, the image will include more than one type of terrain.

The aperture size was adjusted so that low frequencies are not obscured by the central order while the ring structure is not so coarse that frequency spots due to the image are lost by excessive broadening. A sinusoidal frequency ν present in the image will produce a spot located at a distance γ from the center of the diffraction pattern such that $\gamma = f\lambda\nu$, where f is the focal length of the lens and λ is the wavelength of light. The aperture was set at 2 millimeters and it produced ring spacing of 0.471 millimeter, which corresponds to a frequency of 0.61 cycle per millimeter in the image.

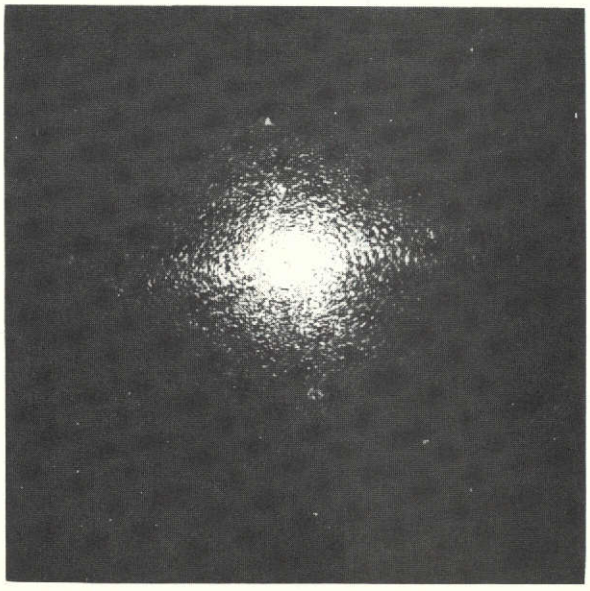
The important spectral content of the diffraction patterns exists for frequencies larger than 1 cycle per millimeter. This frequency corresponds to about 1/3 cycle per millimeter for the original image size (9-1/2-inch format) or 1 cycle per 3 kilometers on the ground.



Phoenix, Arizona. ERTS-217-1



Diffraction pattern with mask. 217-1



Diffraction pattern without mask. 217-1

Fig. 4-2 — Phoenix, Arizona—two diffraction patterns from same area: one with a spatial filter, one without

4.1.2 Description of ERTS-1 Diffraction Patterns

Table 4-1 lists the ERTS-1 images that were employed for developing spatial signatures for various types of terrain. These images are shown in Figures 4-3, 4-5, 4-7, 4-9, 4-11, 4-13, and 4-15. Diffraction patterns were obtained from the image areas identified by numbered circles. The corresponding diffraction patterns are shown in Figures 4-4, 4-6, 4-8, 4-10, 4-12, 4-14, and 4-16, respectively. The diffraction patterns were photographed by using the Fourier plane mask discussed in section 4.1.1.

Table 4-1. List of ERTS-1 Images

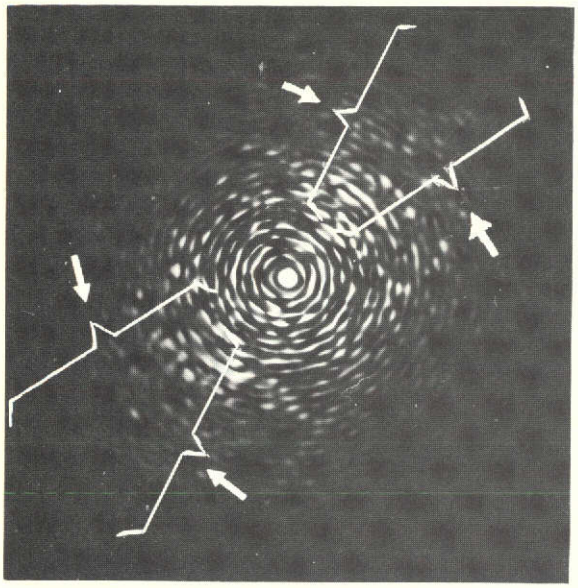
	Image Identification	Area Covered
4-3	030 NASA ERTS-E-1070-17495-5	Imperial Valley, California
4-5	084 NASA ERTS E-1070-16073-5	New Orleans Area
4-7	098 NASA ERTS E-1041-18253-5	Cascade Mountains, Washington
4-9	101 NASA ERTS E-1015-17415-7	Salt Lake City Area
4-11	116 NASA ERTS E-1040-18201-5	Cascade Mountains, Washington
4-13	181 NASA ERTS E-1031-17325-5	Phoenix Area
4-15	217 NASA ERTS E-1031-17325-7	Phoenix Area

4.1.3 Cultivated Land Signatures

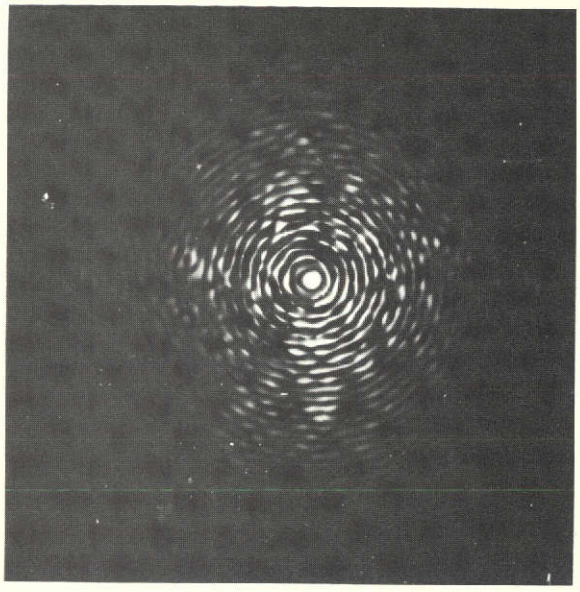
In Figure 4-4, diffraction pattern no. 030-3 from a portion of the Imperial Valley displays a unique signature for the cultivated land of this region. The signature consists of two orthogonal rows of frequency spots. One of the rows is slightly titled to the horizontal. A third row, which is horizontal, is barely visible and is a result of the line structure of the image. The two orthogonal rows are due to the square fields, whose size and dimensions are highly repeatable. The frequency spots are multiples of the fundamental frequency.

That so many high frequency spots are visible is due to the Fourier plane mask. In Figure 4-6, there are two diffraction patterns, nos. 084-1 and 084-3, from cultivated land along the Mississippi River. The fields are elongated rectangles, with their long dimension approximately normal to the river. In diffraction pattern no. 084-3, there are two orthogonal rows of spots but the spacing of the spots in the rows is different because

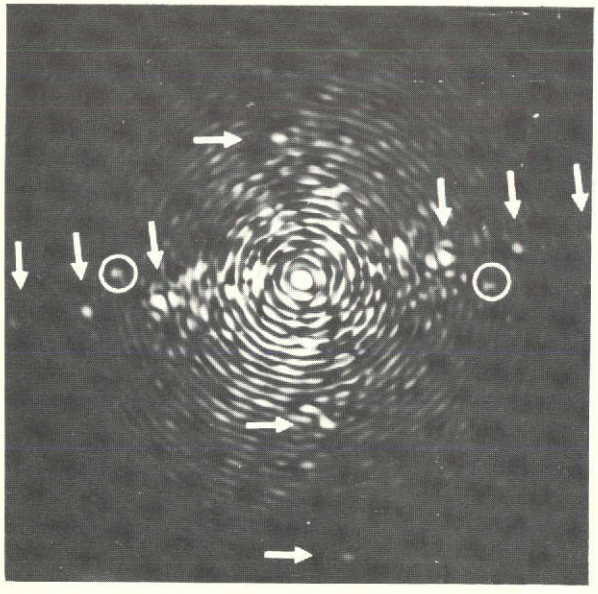
Fig. 4-3 — Salton Sea, California—circled areas show scenes from which diffraction patterns were produced



030-1 Mountainous



030-2 Urban



030-3 Cultivated

↑ Spots due to fields

◯ Spots due to scan lines

Fig. 4-4 — Salton Sea, California—diffraction patterns from circled areas of Fig. 4-3

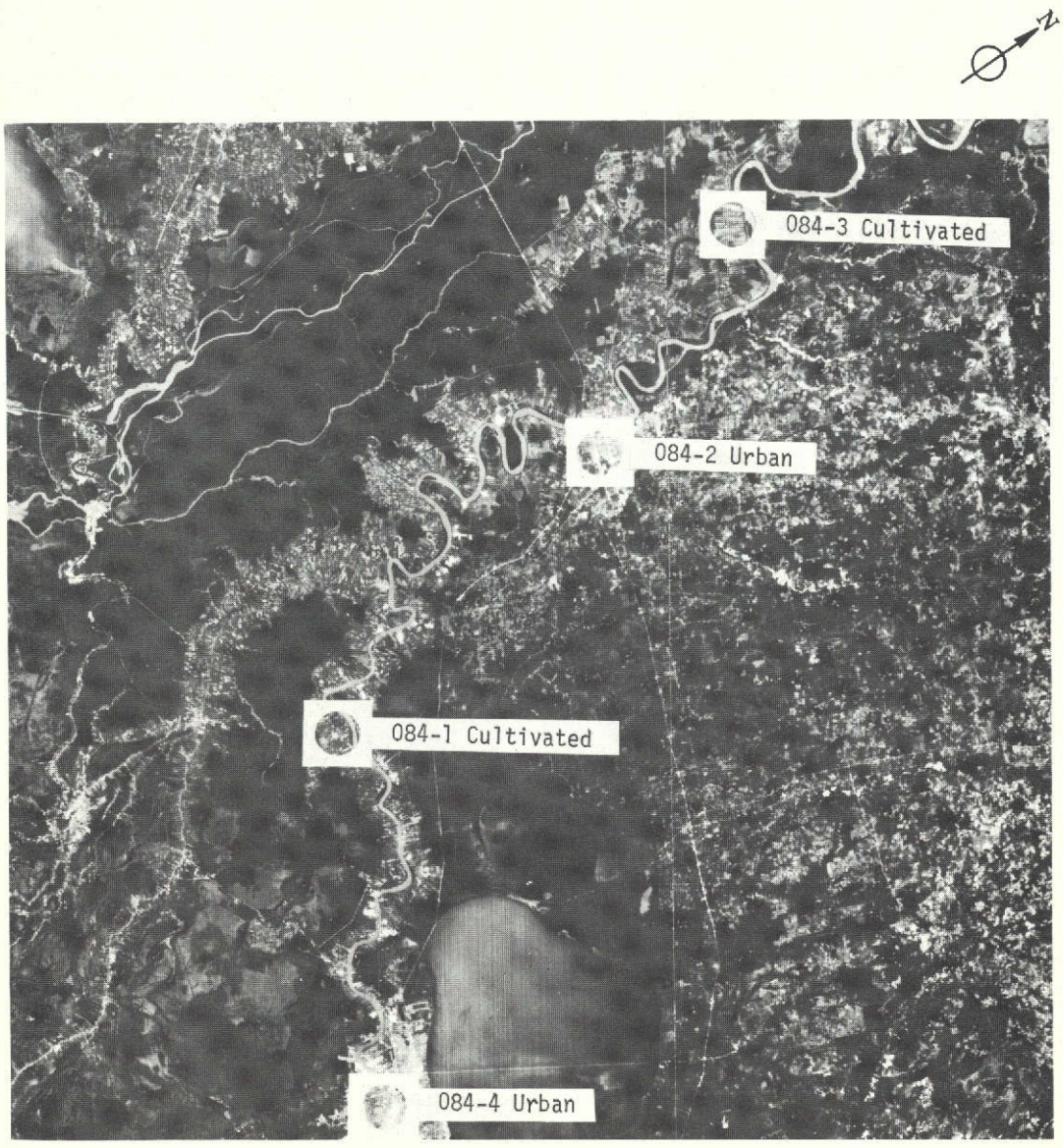
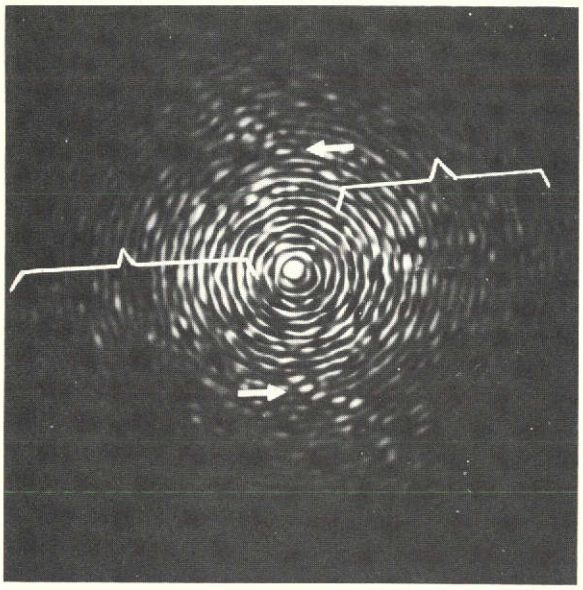
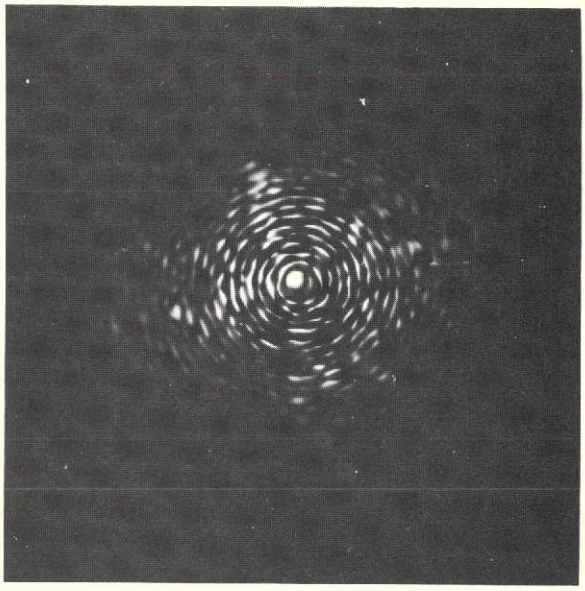


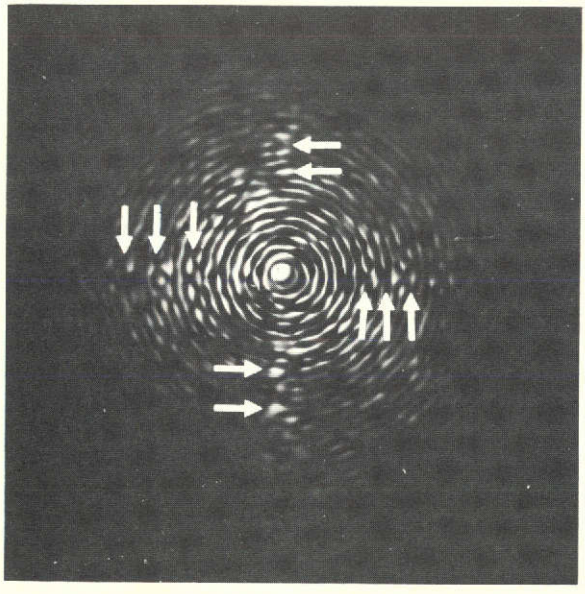
Fig. 4-5 — New Orleans, Louisiana—ERTS image 1070-16037-5



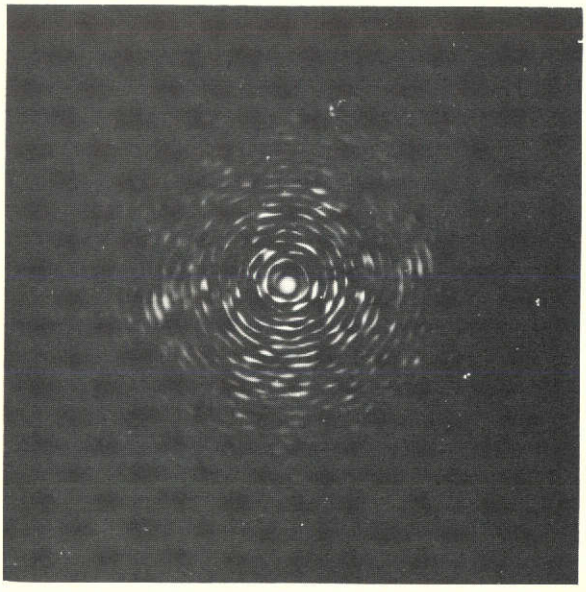
084-1 Cultivated



084-2 Urban



084-3 Cultivated



084-4 Urban

Fig. 4-6 — Diffraction patterns from circled areas of Fig. 4-5

of the elongated rectangular shape of the farms. In diffraction pattern no. 084-1, there is only one row of spots. Orthogonal to it, there is an almost continuous linear structure instead of another row of spots. In Figure 4-5, the farms appear to have the same shape as the farms that produced diffraction pattern no. 084-3. However, crops in adjacent farms appear to be similar (in terms of reflectance) and the result is that the farms appear to be wider and of irregular shape. This comparison brings out a crucial point: the diffraction pattern is affected not only by the field size and shape but also by the crop reflectances.

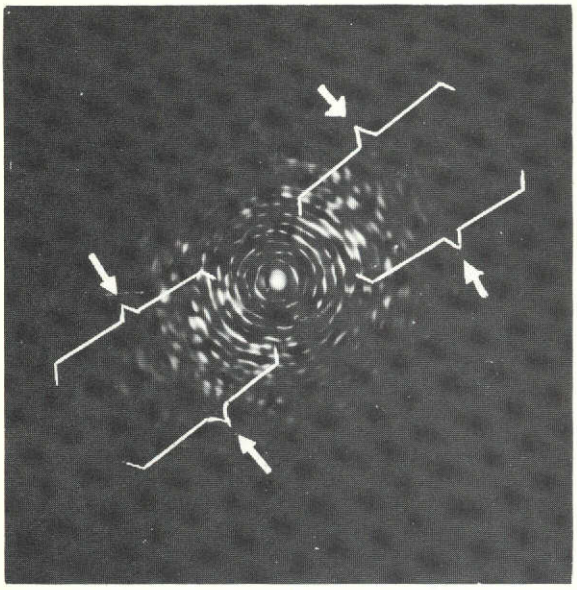
In Figure 4-11, there are two encircled areas that contain circular fields. In Figure 4-12, diffraction pattern no. 116-1 has a structure that corresponds to the farm pattern. There is a broken circular ring that resulted from the interference at the circular aperture with the numerous circular fields. Then, there are at least two frequency spots at about double the frequency of the circular ring. These appear to belong to two orthogonal and equally spaced rows of frequency spots, which are due to the uniform spacing of circular fields in rows and columns.

Diffraction pattern no. 116-2 displays a complex pattern. There appears to be a near vertical row of frequency spots and a much fainter horizontal row normal to the vertical one. Also, there are two orthogonal fans (light diffracted into broad lines) and a smaller fan located about 45 degrees to the other two fans. The orthogonal rows appear to be related to the regular spacing of the circular farms. Displacement of frequency spots in the row is slightly smaller than the displacement of the frequency spots for diffraction pattern no. 116-1. This agrees with the observation that the farms for diffraction pattern no. 116-2 appear larger than the farms for diffraction pattern no. 116-1. The origin of the fans is not obvious, but they may be related to the variation of reflectance in the farms (white, gray, and black).

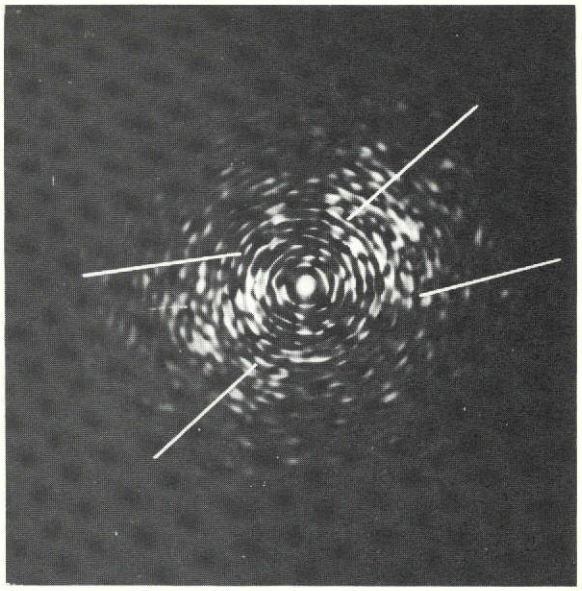
There are two cultivated areas in Figure 4-13, with diffraction patterns no. 181-2 and 181-3 shown in Figure 4-14. Both patterns have two orthogonal rows of frequency spots, which are, however, very weak. In addition, there is a lot of light diffracted in other directions. The farms in Figure 4-13 are characterized by many vertical and horizontal edges. However, due to the reflectance of the crops, the edges appear to be random in size and distribution. In turn, the diffraction patterns are characterized by many frequency spots. However, the principal spots marked in the figure do correspond to the predominant farm size in that region, which is a quarter-mile-square section.

In conclusion, it appears that cultivated land can be identified by a signature, in the diffraction patterns, that consists of two orthogonal rows of frequency spots. The spacing of the frequency spots may not

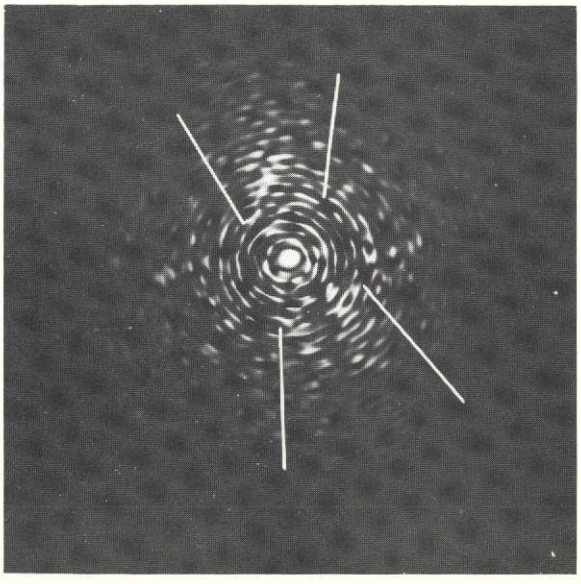
Fig. 4-7 — Cascade Mountains, Washington—ERTS image 1041-18253-5



098-1 Mountainous



098-2 Mountainous



098-3 Snow-covered mountains

Fig. 4-8 — Diffraction patterns from circled areas of Fig. 4-7

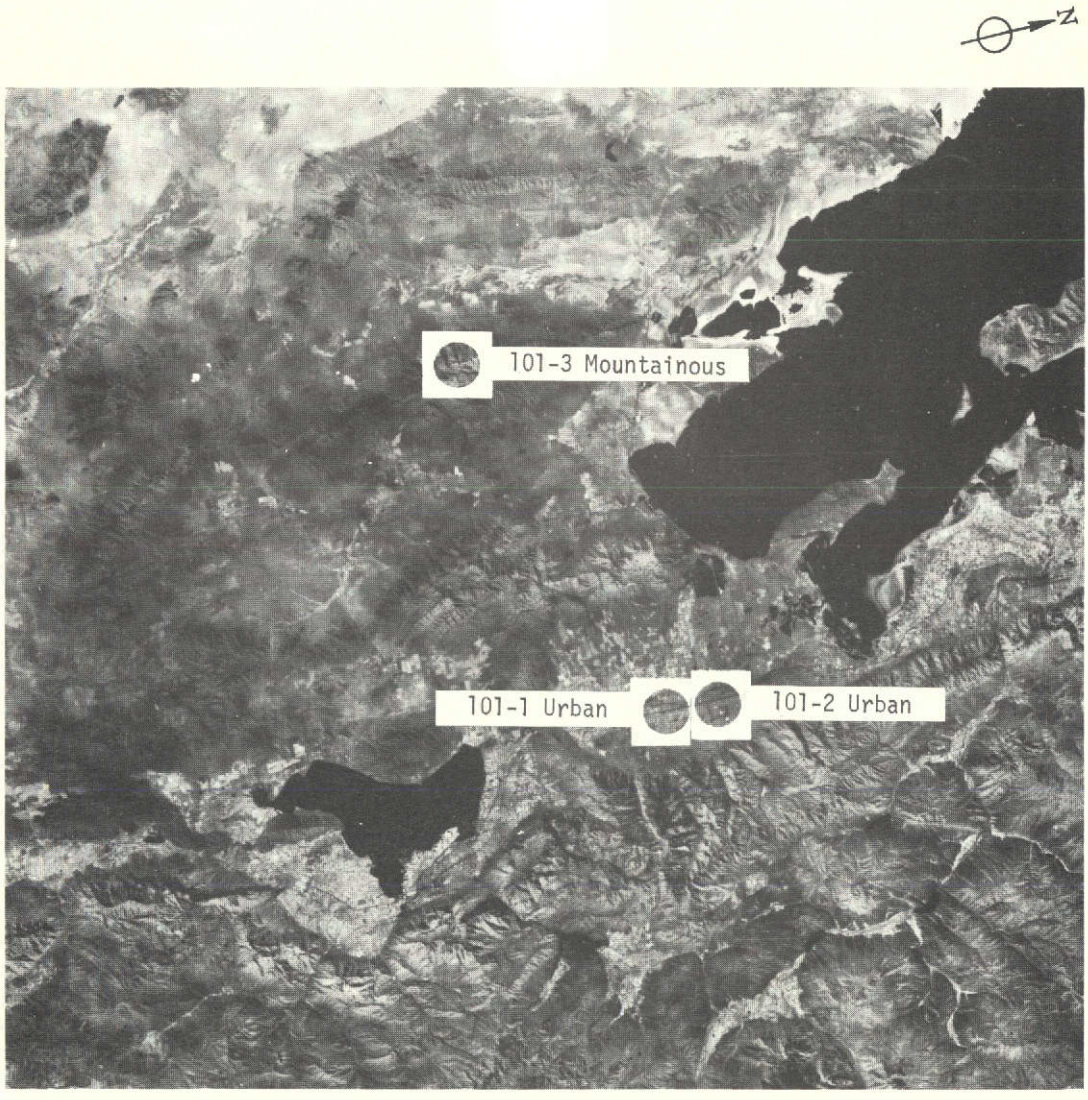
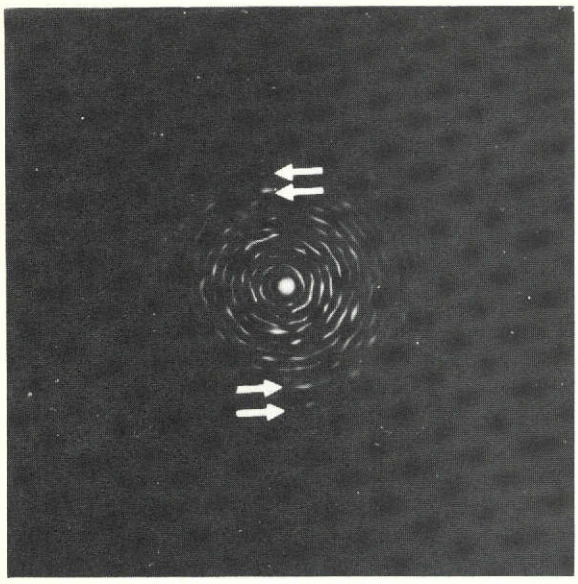
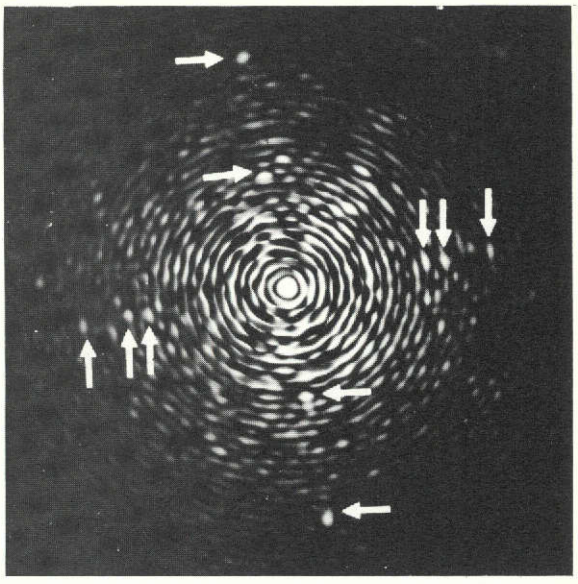


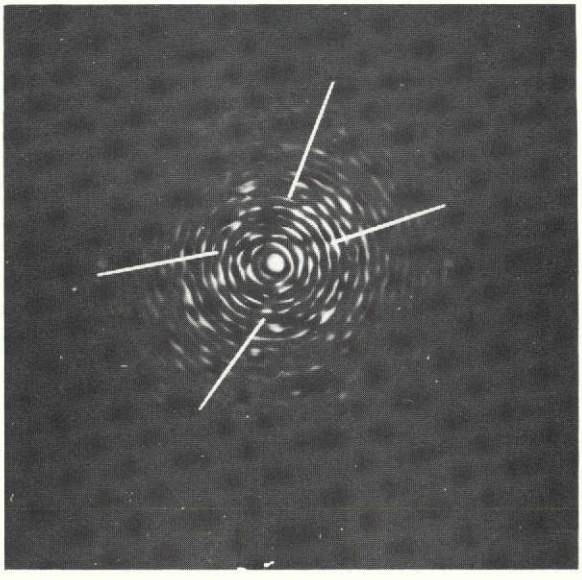
Fig. 4-9 — Great Salt Lake, Utah—ERTS image 1015-17415-7



101-1 Urban



101-2 Urban



101-3 Mountainous

Fig. 4-10 — Diffraction patterns from circled areas of Fig. 4-9

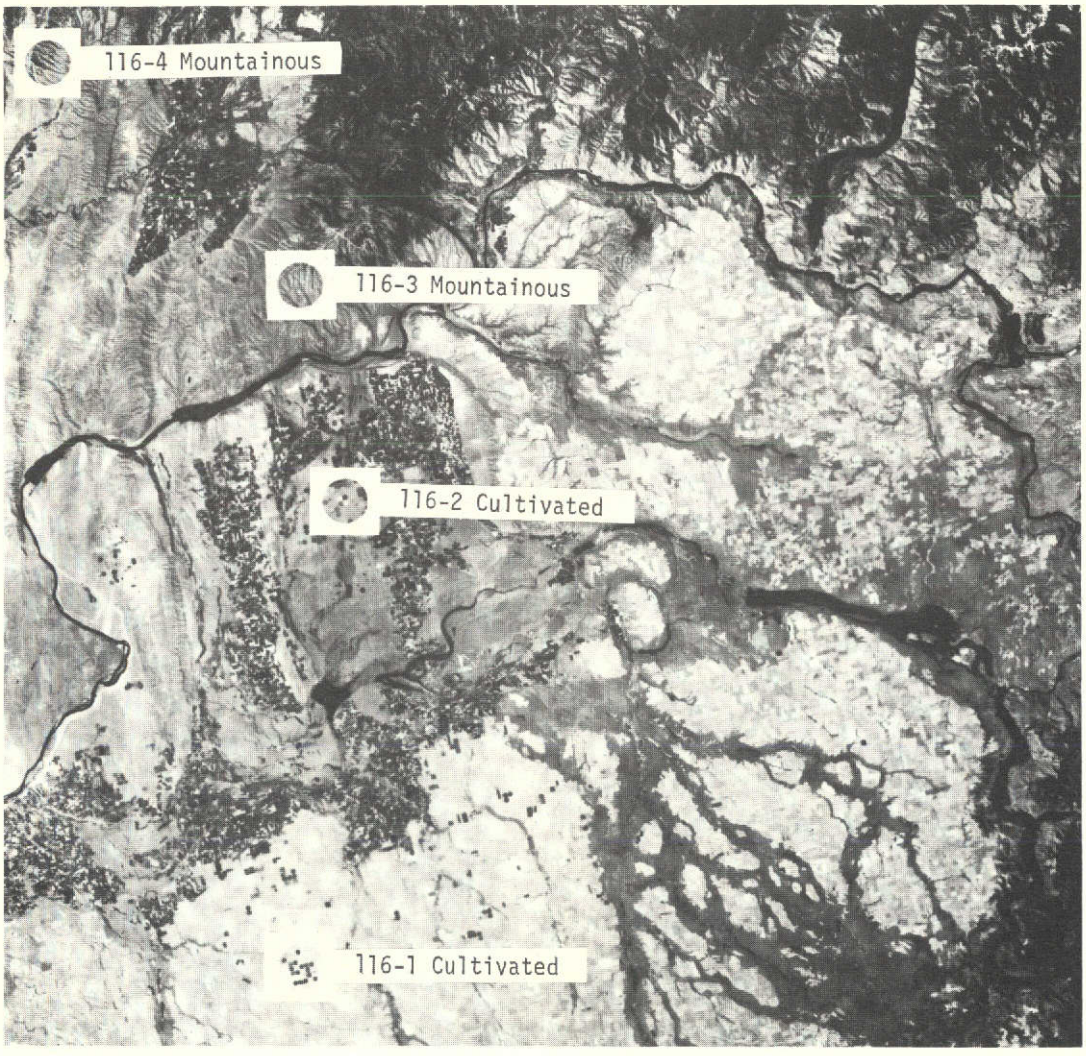
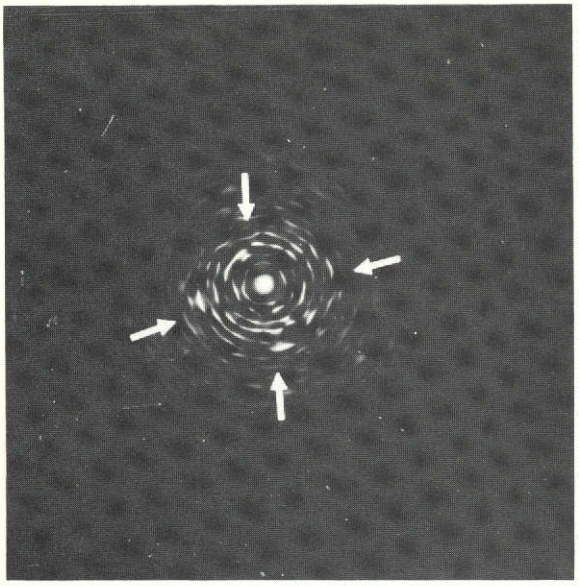
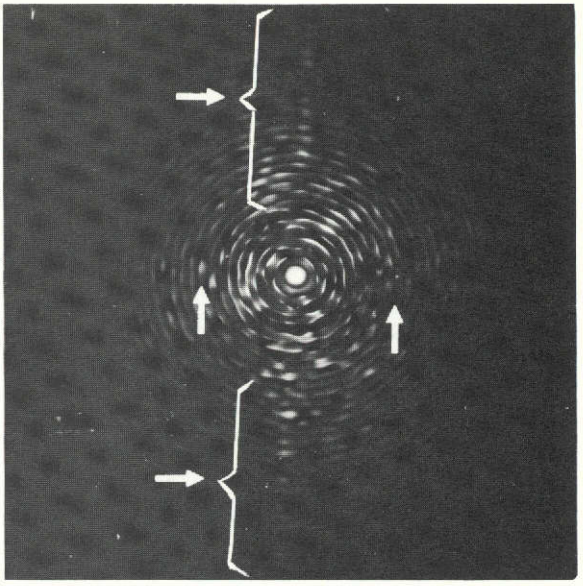


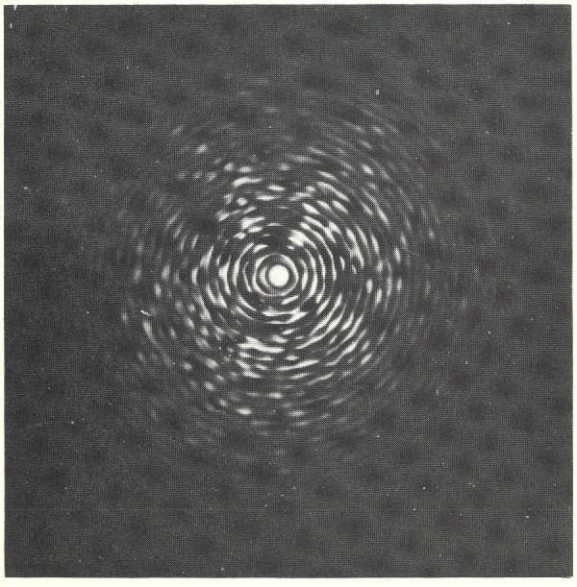
Fig. 4-11 — Cascade Mountains, Washington—ERTS image 1040-18201-5



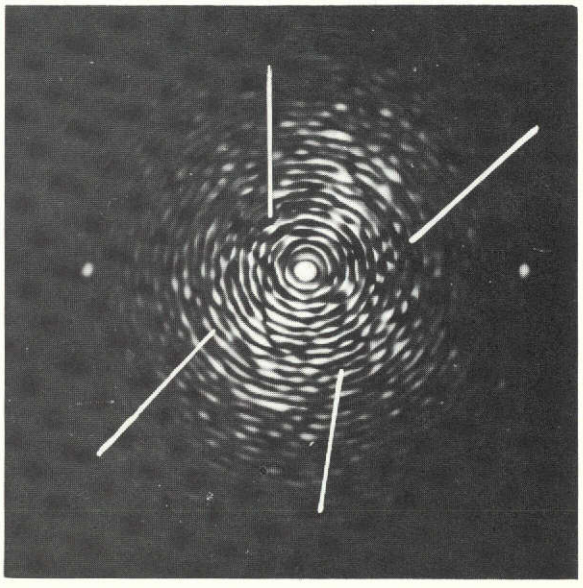
116-1 Cultivated



116-2 Cultivated



116-3 Mountainous



116-4 Mountainous

Fig. 4-12 — Diffraction patterns from circled areas of Fig. 4-11

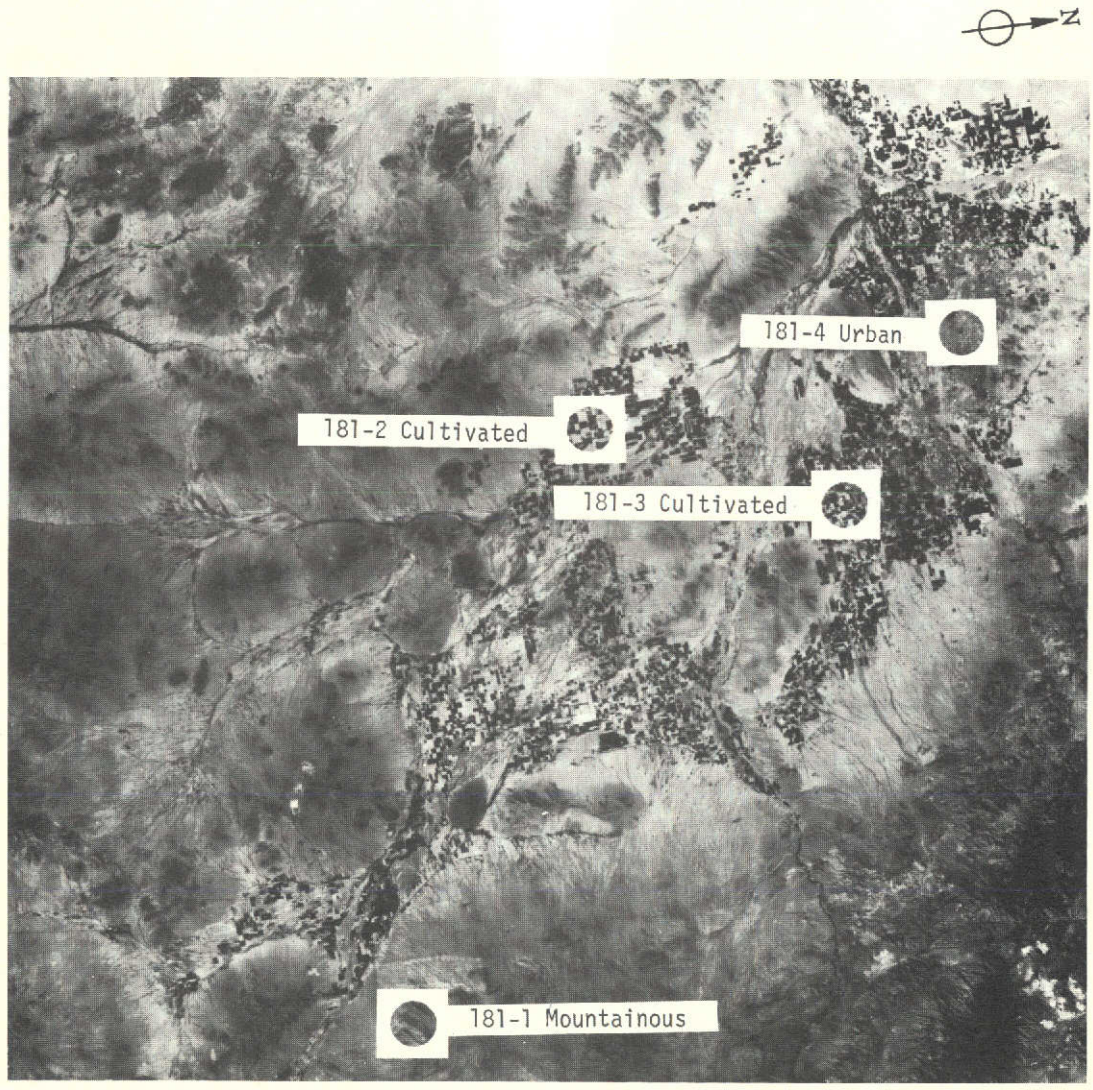
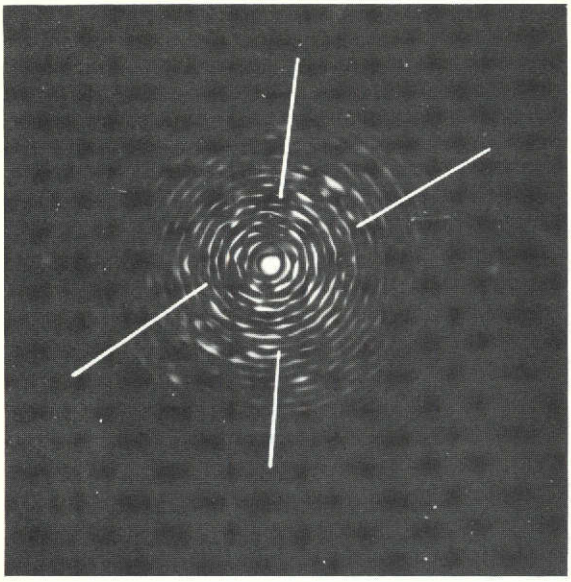
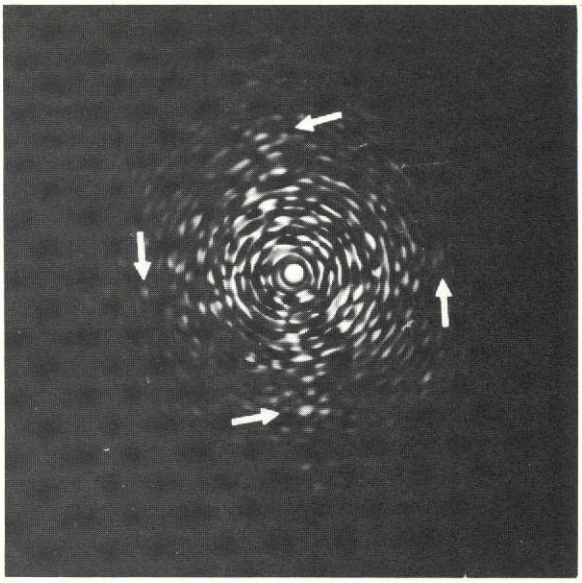


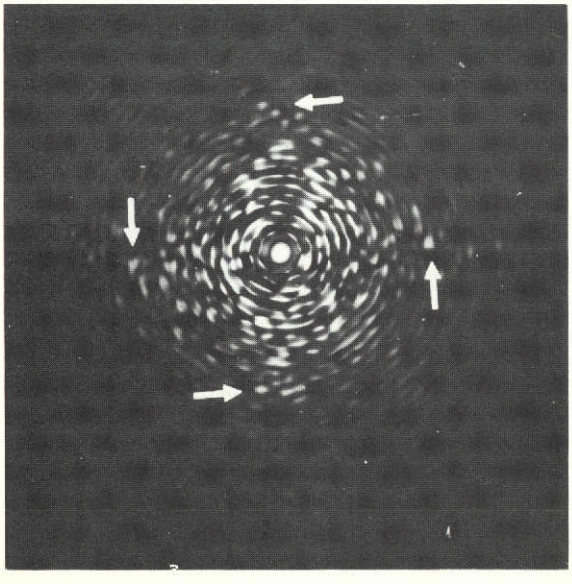
Fig. 4-13 — Phoenix, Arizona—ERTS image 1031-17325-5



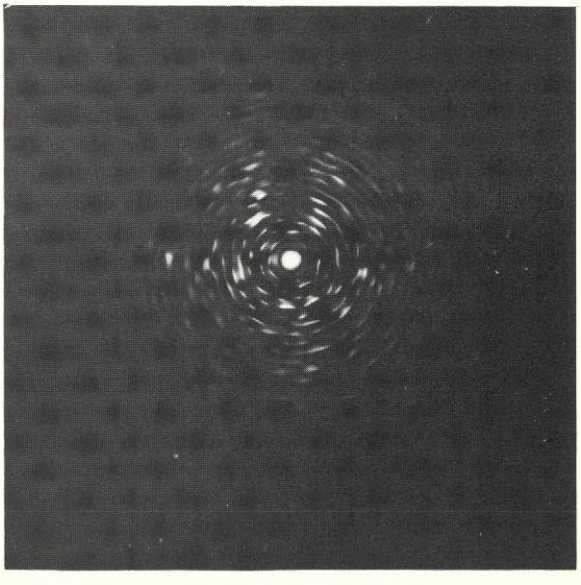
181-1 Mountainous



181-2 Cultivated



181-3 Cultivated



181-4 Urban

Fig. 4-14 — Diffraction patterns from circled areas of Fig. 4-13

be the same in both rows. For some images, the signature must be detected in a diffraction pattern that contains many other strong frequency components.

4.1.4 Mountainous Terrain Signatures

In Figure 4-4 diffraction pattern no. 030-1 shows diffraction of light (fans) in a direction normal to the mountain ridges (see Figure 4-3). The fans are rather broad and contain strong frequency components.

In Figure 4-12, there are three diffraction patterns from mountainous terrain in the Cascade Mountains. These diffraction patterns have characteristics similar to diffraction pattern no. 030-1.

Diffraction pattern no. 101-3 in Figure 4-10 was produced by the mountainous terrain of Figure 4-9. This pattern is also characterized by a broad fan in a direction normal to the mountain ridges. In Figure 4-12, diffraction patterns nos. 116-3 and 116-4 are from the mountainous terrain of Figure 4-11. Diffraction pattern no. 116-4 is again similar to the other patterns of mountainous terrain. Pattern no. 116-3 displays an extremely broadened fan, which can be attributed to variations in the orientation of the corresponding mountain ridges in Figure 4-11.

In Figure 4-14, diffraction pattern no. 181-1 was produced by mountainous terrain from Figure 4-13. This pattern also shows a broad fan normal to the mountain ridges. In conclusion, the diffraction patterns of mountainous terrain show diffraction of light along broad fans normal to mountain ridges. An obvious, unique signature for mountains has not been identified.

4.1.5 Urban Area Signatures

In Figure 4-4, diffraction pattern no. 030-2 was produced by a small town in the Imperial Valley (see Figure 4-3). There is nothing distinctive about this pattern, and diffraction of light along fans appear to be due to the edges of the cultivated region surrounding the town.

In Figure 4-10, diffraction patterns no. 101-1 and 101-2 are from the urban area of Salt Lake City. Several major arteries are seen in the encircled areas in Figure 4-11. The diffraction patterns show several rows of frequency spots. The horizontal row is due to the line structure of the image (a scanner/recorder artifact). Tilted to that row is a pair of orthogonal rows due to the repetitive pattern of orthogonal major arteries of Salt Lake City.

In Figure 4-14, diffraction pattern no. 181-4 is produced by an area of Phoenix from the image of Figure 4-13. This pattern shows only one

horizontal row of frequency spots owing to the line structure of the image.

In Figure 4-16, diffraction pattern no. 217-1 shows two orthogonal rows of frequency spots owing to the major street pattern of Phoenix (see Figure 4-15). In conclusion, urban areas can be identified by signatures resulting from patterns of parallel major arteries. This signature becomes evident only in the IR-2 (0.8- to 1.1-micrometer) band for the cities of Phoenix and Salt Lake. The difference in signature from the same area photographed in two different spectral bands is seen by comparing the relative strengths of the diffracted spots in Figures 4-14 (no. 181-4) and 4-16 (no. 217-1). In New Orleans, such a street pattern does not exist. Instead, New Orleans is characterized by intersections of major highways that are visible in the red band image. It appears that a signature for small towns does not exist since secondary street patterns are not resolved.

4.1.6 Other Terrain Signatures

Diffraction patterns were obtained also from other terrain features such as broken clouds, rivers, and transportation networks. These have not been presented because they are not expected to produce signatures.

4.1.7 Conclusions

The search for terrain signatures has produced the following results:

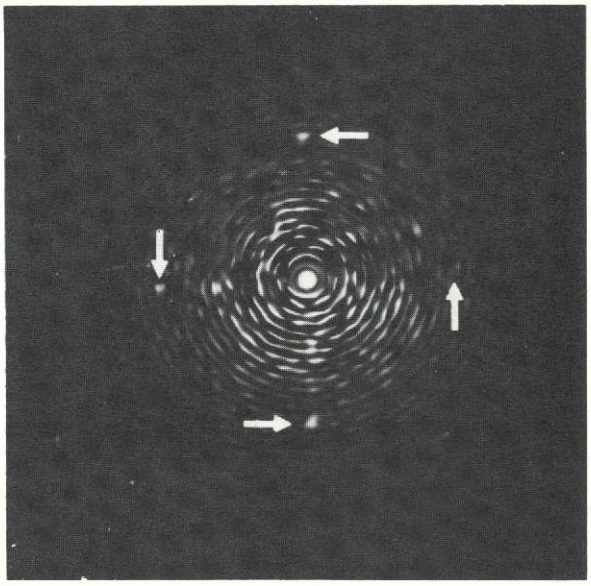
a. The diffraction patterns of mountains show clusters of intensity at low spatial frequencies and diffracted light spreading in wedge-shaped patterns (fans) away from the center.

b. For cultivated land, a signature consisting of orthogonal rows of frequency spots has been identified. For some images, the signature may be weak in relation to other components of the diffraction pattern.

c. For urban areas, a signature consisting of orthogonal rows of frequency spots has been identified. This signature becomes evident only in the IR-2 image and cannot be confused with the signature for cultivated land, which is detected in the red band (0.6 to 0.7 micrometer). In the red band, a signature for some urban areas can be obtained with a high degree of probability from intersecting major highways.

d. Bodies of water have diffraction patterns with very low energy at frequencies other than the central lobe. Water can be identified by low reflectances in nearly all spectral bands, and,

Fig. 4-15 — Phoenix, Arizona—ERTS image 1031-17325-7



217-1 Urban

Fig. 4-16 — Diffraction pattern from circled area of Fig. 4-15

therefore, the use of a spatial signature is usually redundant.

e. Deserts also have diffraction patterns with low energy levels of light diffracted at most frequencies other than the central lobe and very low frequencies.

f. Rivers and transportation networds are linear features that can be recognized by using algorithms operating directly on the image and testing for topographic characteristics. Although linear features have Fourier transforms that display high spatial frequency content in a direction perpendicular to the length of the feature (i.e., normal to the river bands), there is no indication that the Fourier transform provides better signature discrimination that the image itself.

In summation, the terrain features have been divided into three categories:

a. Features for which spatial signatures are not appropriate - deserts and bodies of water.

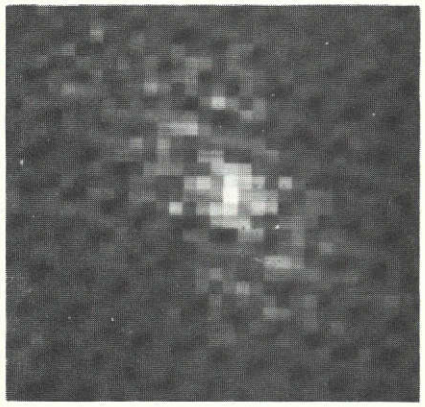
b. Features that can best be identified by spatial signatures developed directly from an image - broken clouds, rivers, and transportation networks.

c. Features that can be identified by spatial signatures isolated in the diffraction patterns or the Fourier transforms - cultivated land, some urban areas, mountains, and hills.

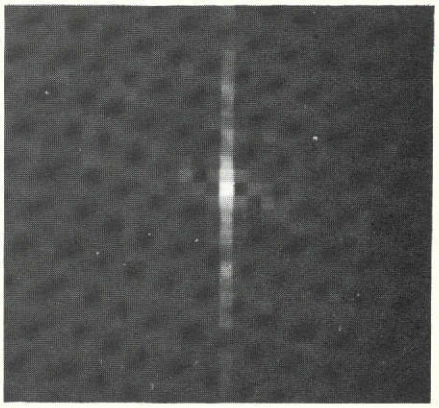
4.2 Digital Signatures

The diffraction pattern annalysis of the previous section suggests that signatures for the various terrain types can be found in the digital Fourier transforms of cells.

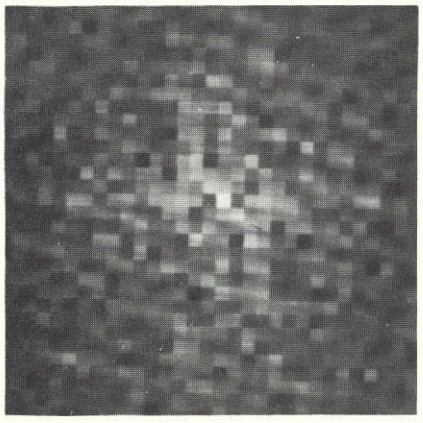
Figure 4-17 shows the amplitude of the Fourier transforms of four cells. The horizontal direction coincides with the scanning direction of the multispectral scanner. The frequencies represented by the squares vary from 0.39 to 5.9 cycles per kilometer. Due to the ground resolved distance of the MSS data, the street pattern of Phoenix is not resolved in the red band image. Also, the boundaries between adjacent farm plots are not resolved. The farm pattern within a cell seems to be established by crops of equal brightness. If such crops exist in adjacent plots, they have the appearance of larger farms of irregular shape. Also, the farm pattern is affected by clouds and highways. Comparison of the digital Fourier transforms to the diffraction patterns of the previous section shows many similarities as expected.



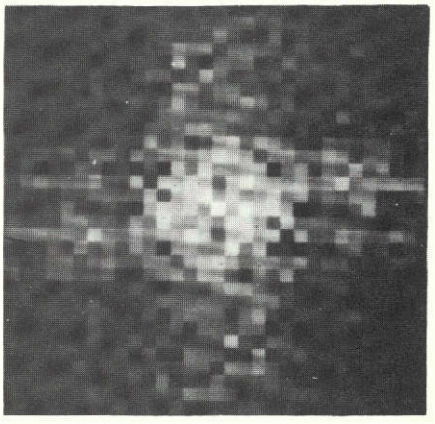
Mountain



Desert



City



Farms

Fig. 4-17 — Digital Fourier transforms of cells

The Fourier transforms were extensively analyzed to determine what portions of the transforms contain information that can be used to recognize terrain types. It is desirable that areas of the transforms not contributing to the recognition of terrain be excluded from the spatial feature measurements. Furthermore, a minimum number of spatial features should be extracted.

The analysis of the Fourier transforms has produced the following results:

a. It appears that frequencies larger than 3.5 cycles/km contain the information needed to discriminate between the terrain types. Frequencies less than 3.5 cycles/km carry significantly less information.

b. Regardless of terrain type, there is significant energy along the vertical frequency column ($f_x = 0, f_y$) where x is the horizontal direction and y the vertical. In addition, this column has local peaks at the frequencies f_y = 2.1, 4.2 and 6.3 cycles/km. These appear to be multiples of 1/6 the scanning rate (12.6 lines/km) of the Multispectral Scanner. Since the scanner has six detectors per spectral band, it is conceivable that these frequencies are artifacts due possibly to small errors in detector response remaining after calibration. The calibration errors may not be large enough to produce line structure in an MSS image. In any case, the frequency column mentioned contains frequency components of the same order of magnitude as the components which have been related to farms. The discrimination results between the various terrain types were improved when this column was replaced by zeros.

Measurements of four spatial features are made as follows:

a. All frequencies less than 3.5 cycles/km and larger than 5.9 cycles/km are eliminated.

b. In the remaining Fourier transform, the largest peak (or maximum) is determined.

c. The energy in a sector which is $\pi/8$ radians wide and centered on the largest peak is determined. The energy within this sector (S_1) is one of the features.

d. The energies in similar size sectors which are displaced from the first one by $\pi/4$, $\pi/2$ and $3\pi/4$ radians in a clockwise direction are also determined and constitute the features S_2, S_3 and S_4, respectively. (See Figure 4-18)

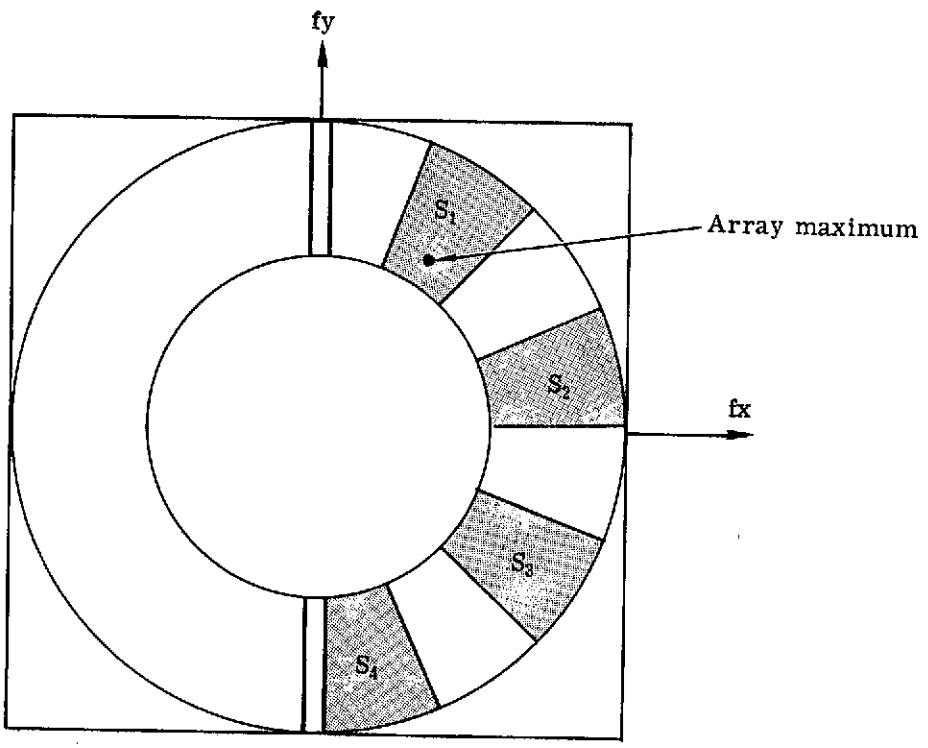


Fig. 4-18 — Fourier transform integration regions

Other spatial features that have been employed are: a number representing a count of high derivative values within the cell and two numbers obtained from the histogram of the cell pixels.

Upon completion of the spatial feature selection, the classification of the cells can then be carried out using only the feature vectors. The classification involves assignment of the cells to the terrain types by partitioning the feature space. For this investigation, the division of the feature space was done with a heuristic algorithm, in order to analyze the usefulness of the features that were selected. This heuristic algorithm consists of logic statements containing many thresholds. (See section 5.1.)

4.3 Spatial Feature Measurements

The generation of the feature vectors is a time consuming data reduction operation. On a cell basis, the 1,024 pixels are reduced to seven numbers. The reduction in data is about 439 to 1 when the spectral data (MSS 4, 5 and 7) is taken into account. For each cell, the input data consists of 3 x 1,024 = 3,072 numbers. The time required to compute the seven feature vector components is about 1.3 seconds in the IBM 370/158 computer. The time required to compute the vectors for an entire ERTS image is over two hours. It is desirable that the feature measurements be executed at a substantially faster rate.

Also, it is desirable to reduce the size of the cell to less than 2.5 km square in order to reduce the probability that a cell will contain more than one type of terrain. Reducing the cell size digitally to less than 32 x 32 pixels makes the Fourier transform grainy and introduces considerable frequency quantization noise to the spatial features S_1, S_2, S_3, S_4. To reduce this quantization noise, it will then be necessary to interpolate and resample the Fourier transforms. For example, if the digital cell size is reduced to 16 x 16 pixels, the Fourier transforms will also contain 16 x 16 samples. These would have to be interpolated and resampled so that they contain at least 32 x 32 samples. The interpolation and resampling operations will consume more computer time.

The spatial features can be measured more rapidly and efficiently by measuring the energy in specified portions of the diffraction patterns of cells. In other words, the spatial features of cells can be measured electro-optically with photodetectors after the diffraction patterns of the cells have been formed optically as described in section 4.1. The detector outputs can be digitized and then introduced into a computer as feature vectors.

An added advantage of the electro-optic measurement of features is that the cell size can be reduced without introducing frequency quantization noise as occurs with the digital Fourier transforms.

However, the main advantage of the electro-optic measurement of the features is the speed with which the feature vectors can be measured. If one assumes that the feature vector for a cell is measured conservatively at the rate of one vector per millisecond, the speed advantage is 1,300:1 in comparison to the computer. If the cell size is reduced to an equivalent of 16 x 16 pixels, the feature vectors of an entire ERTS image can be measured at about 21 seconds.

5. TERRAIN TYPE RECOGNITION

5.1 Heuristic Recognition Algorithm

This algorithm was developed by examining the feature vectors of cells of the same class. Then, the vector space was divided by thresholds forming hyperplanes. The algorithm was then employed to classify the cells from a portion of ERTS-1 image no. 1049-17324-5 containing Phoenix, Arizona and the surrounding area in southern Arizona (about 8,000 square km). The thresholds in the algorithm were then adjusted until the recognition results were optimized≠.

In order to determine the accuracy of the heuristic algorithm, two assignment matrices of cells were developed by photointerpretation. One matrix represents assignments using only the ERTS-1 red band image 1049-17324-5. The other matrix represents assignments using all available information (maps, aerial photography, and a land use map*). Comparing these two matrices, one finds that the first one has about 2.5% errors, which indicates the error rate of a photointerpreter if he were to use only the red band ERTS-1 image. Most of the errors are due to isolated urban areas near cultivated land, mountains, or desert.

The computer classification matrix was compared to the most accurate photointerpreter matrix. The results are tabulated in Table 5-1.

≠Gramenopoulos, N., "Terrain Type Recognition using ERTS-1 MSS Images", NASA SP-327, Symposium on Significant Results Obtained from the Earth Resources Technology Satellite-1, March 5 - 9, 1973.

*Poulton, C. E., Schrumpf, B. J., Johnson, J. R., "Ecological Resource Analysis from High-Flight Photography for Land Use Planning", Applied Remote Sensing of Earth Resources in Arizona, Proceedings 2nd ARETS Symposium, University of Arizona, November 2 - 4, 1971.

TABLE 5.1 COMPARISON OF CLASSIFICATION RESULTS*

COMPUTER ASSIGNMENT CLASS	PHOTOINTERPRETER'S ASSIGNMENT 1	2	3	4	5	6	7	8	TOTAL
0	0	0	3	15	22	6	14	0	60
1	61	0	0	0	0	0	0	1	62
2	0	2	0	1	0	0	0	1	4
3	1	0	417	6	7	8	5	0	444
4	0	0	3	248	1	9	2	1	264
5	0	0	5	3	131	2	4	0	145
6	0	0	1	5	3	71	8	0	88
7	0	0	1	0	0	0	4	0	5
8	0	0	0	0	0	0	0	13	13
TOTAL	62	2	430	278	164	96	37	16	1085

Class Identification

0 = Unidentified
1 = Clouds
2 = Water
3 = Desert
4 = Farms
5 = Mountains
6 = Urban
7 = Riverbed
8 = Cloud Shadows

*Heuristic algorithm using spatial features only.

Many cells contain two terrain types, and the computer assignment of a cell is deemed correct if one of the terrain types is correctly recognized. Clouds and cloud shadows can confuse the classification of underlying terrain. Therefore, cells that contain even a small cloud or a cloud shadow are not assigned to a terrain type. The underlying terrain can be classified from different ERTS-1 images when it is cloud free.

Whenever cells cannot be classified with reasonable confidence, they are assigned to an unidentified group. This can be justified by the fact that all the information available (the other three spectral images) is not being utilized yet, and it is desirable to defer making decisions on questionable cells until the spectral information is introduced. Due to scale and resolution, only three types of natural terrain can be recognized: desert, mountains, and riverbeds or flood plains.

Due to the population density of the area and man's activities, the signatures of the various terrain types vary considerably, thus increasing the probability of error. The following examples have been noted:

a. Highways and canals cutting through desert can be confused with flood plains.

b. Urban development has extended into the hills north of Phoenix and new developments farther north into the desert have been initiated. These areas are now urban, but retain some of their former desert or mountainous characteristics, thus producing some classification errors.

(See Figures 5-1 and 5-2.)

c. The urban areas are spreading into farmland near Glendale and farms have been noted surrounded by urban areas or in the process of being coverted to new developments. Such areas have produced classification errors or unidentified cells. Sun City, a circular urban development near Glendale, is quite different in appearance than downtown Phoenix.

d. The Salt River bed runs through Phoenix. The waters of this river have been diverted for irrigation, and the river bed is now dry. The riverbed is used for urban or industrial purposes, but it is thinly populated in comparison to Phoenix. Classification errors and unidentified cells were found along the river bed.

It was also noted that farms are a dominant terrain type. Most cells containing two terrain types with farms being one of them were assigned to the farm category.

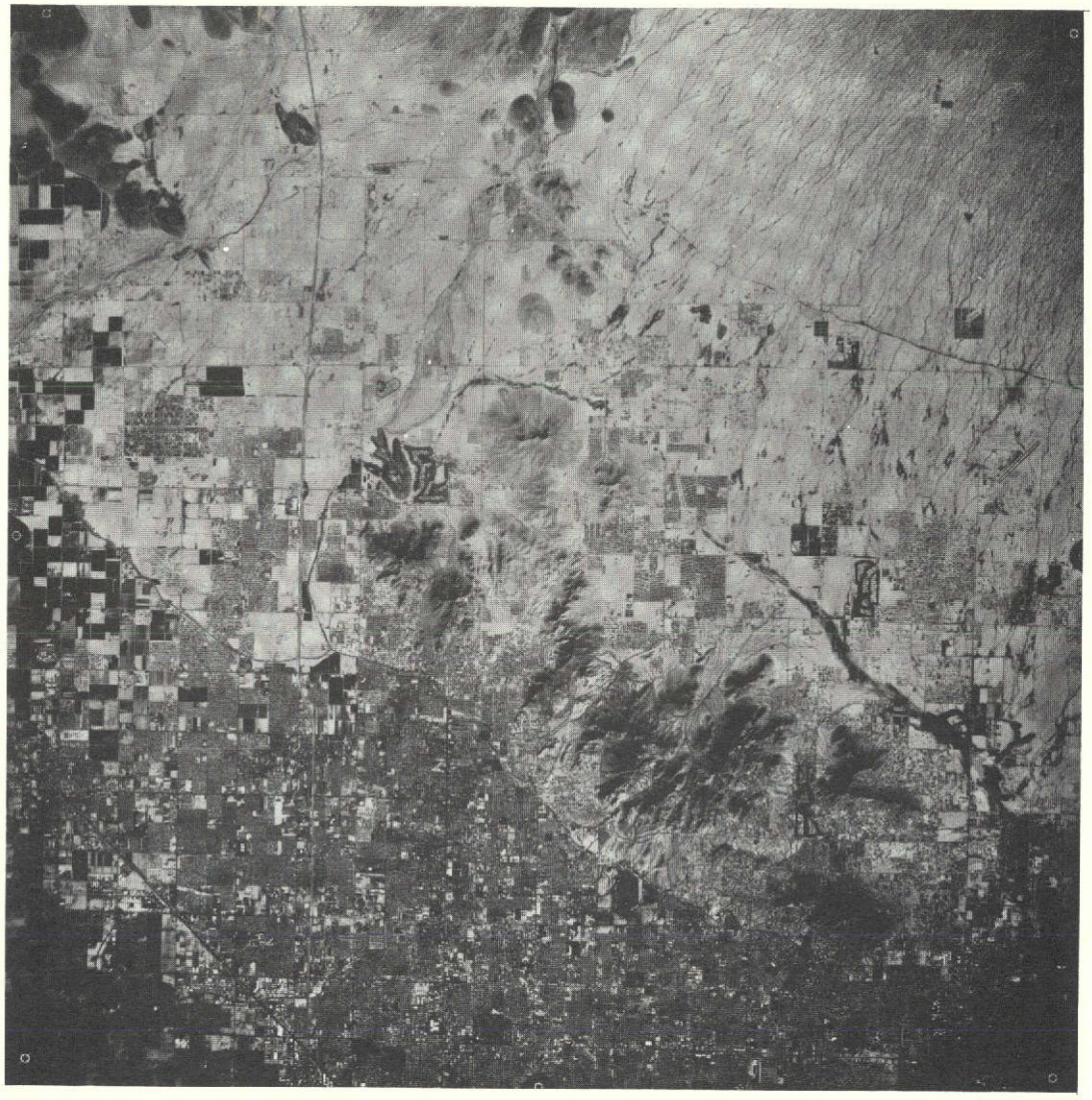


Fig. 5-1 — U-2 photograph of northern suburbs of Phoenix, Arizona

Fig. 5-2 — Low altitude photograph of Phoenix, Arizona suburbs—taken by NASA earth resources aircraft

The classification results are shown in the annotated photograph of Figure 5-3. The letters were superimposed on the MSS-5 image and are centered within each cell. The interpretation of the symbols is:

D	=	Desert	R	=	Riverbeds
U	=	Urban	C	=	Clouds
F	=	Farms	S	=	Cloud Shadows
M	=	Mountains	Blank	=	Unidentified

The recognition results for most terrain types were very good: 97% for desert, 89% for farms, 80% for mountains, 74% for urban areas, 98% for clouds, 100% for water, 81% for cloud shadows. Only river flood plains, which are peculiar geographic features of southern Arizona, were recognized poorly (11%).

Due to the resolution of the system, urban areas cannot be recognized with a very high degree of confidence until the information in the other spectral bands has been exploited. Most of the errors in the urban category seem to occur in thinly populated areas associated with recent urban expansion into desert, farms, riverbeds, and hills.

From the recognition results, it appears that the features selected are sufficient for the purposes of classifying terrain types. The four most important features are derived from the Fourier transforms of the cells. (See section 4.2.)

5.2 Maximum Likelihood Criterion

5.2.1 Integration of Multispectral and Spatial Information

Integration of spectral and spatial features is required in order to exploit the information available in the ERTS-1 imagery. It is also desirable to employ a general recognition algorithm which is independent of the data characteristics. The maximum likelihood criterion* which has produced good results with multispectral data has been selected. This criterion requires training data and is optimum only when the terrain classes form multivariate Gaussian distributions in feature space. In addition, no class is allowed to have a feature vector component which is always zero.

*Marill, T., Green, D. M., "Statistical Recognition Functions and the Design of Pattern Recognizers", IRE Transactions of Electronic Computers, December 1960, p. 472.

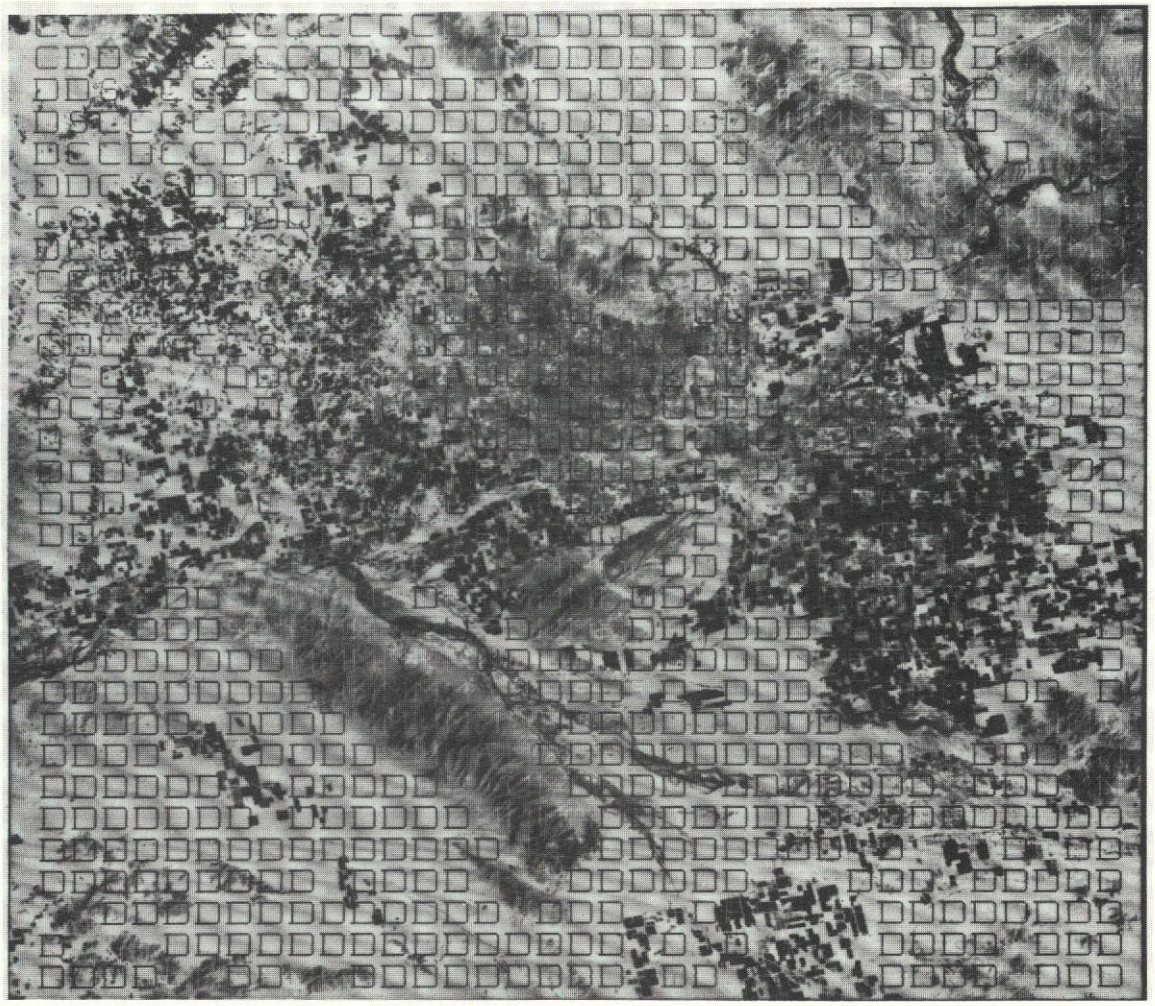


Fig. 5-3 — ERTS-1 image 1049-17324-5 and terrain classification results from heuristic algorithm

The four spatial features described above were combined with three spectral features to form a seven-dimensional vector describing each cell. Using image 1049-17324 as training data, the statistics (covariance matrices) for desert, farm, mountains and urban areas were computed. Then, using the statistics, the same input data and the maximum likelihood criterion, the cells were reclassified in one of the four categories. The recognition accuracy was 54% for desert, 92% for farms , 97% for mountains, and 92% for cities. Desert was poorly recognized (many desert cells were assigned to the mountain category), and it was suspected that the statistics for the classes were not Gaussian.

5.2.2 Gaussian Statistics of Class Vectors

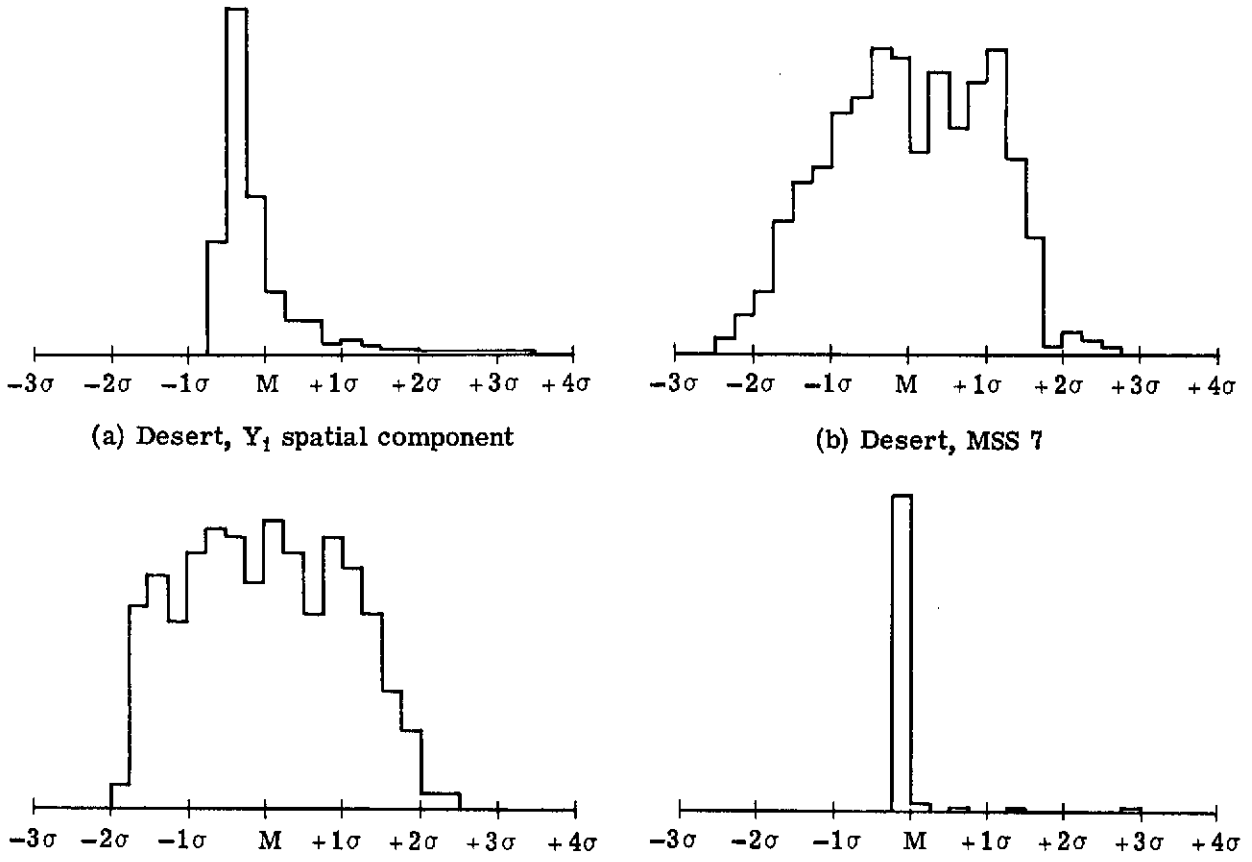
Analysis of histograms of the classes showed that no component of any class was Gaussian. In particular, the distribution for the spatial features appeared like the Rayleigh rather than the Gaussian. All components were positive with small means and large standard deviations. (See Figures 5-4 and 5-5.)

5.2.3 Non-Linear Transformations of Class Vectors

Non-linear transformations applied to the feature space change the statistics of the classes. A different transformation is required for each feature component. The transformations must be such that large component values are reduced while small values remain unchanged. (See Figure 5-6.) Known transformations of this type are the logarithms on various bases such as the natural logarithm or powers less than unity such as the square or third roots. If a logarithmic transformation is selected, the base of the logarithm is adjusted for each feature component so that the class distributions in this component become symmetrical about their means. Similarly, if a power transformation is employed, the power is adjusted for each component so that the class distributions in this component become symmetrical about their means. For example, if the power is 0.5 for all components, then each component of every feature vector is replaced by its square root.

Both logarithmic and power transformations were extensively analyzed. The logarithmic transformations gave good results, but the resultant distributions were sensitive to amplitude variations in the data. In other words, if the logarithm on a certain base is used for one component (for example, the MSS 7 band) the distribution for each class may or may not approximate the Gaussian depending on solar illumination which changes the range of values obtained in this band.

Finally, the non-linear transformations selected are all powers less than 1. For image 1049-17324, the powers for the various



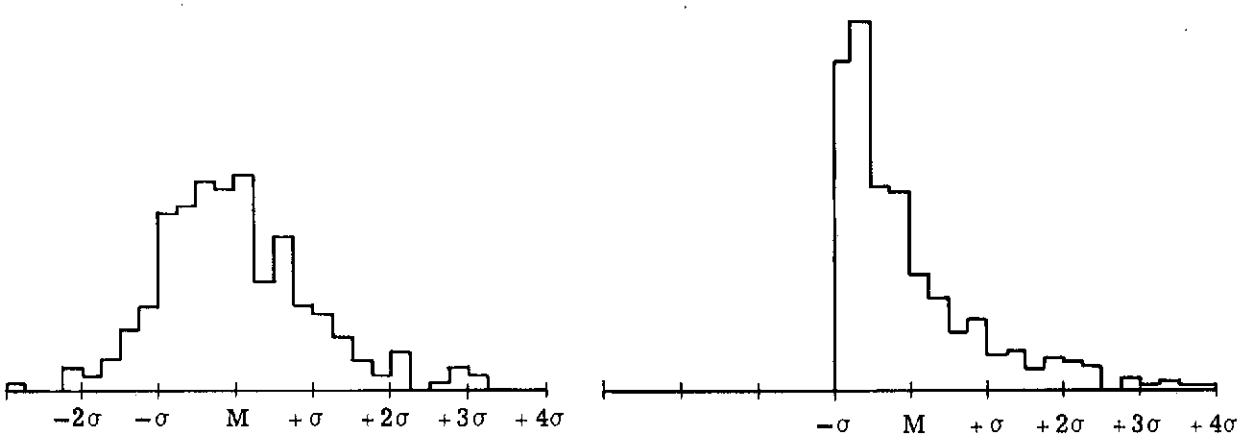
(a) Desert, Y_1 spatial component

(b) Desert, MSS 7

(c) Desert, MSS 5

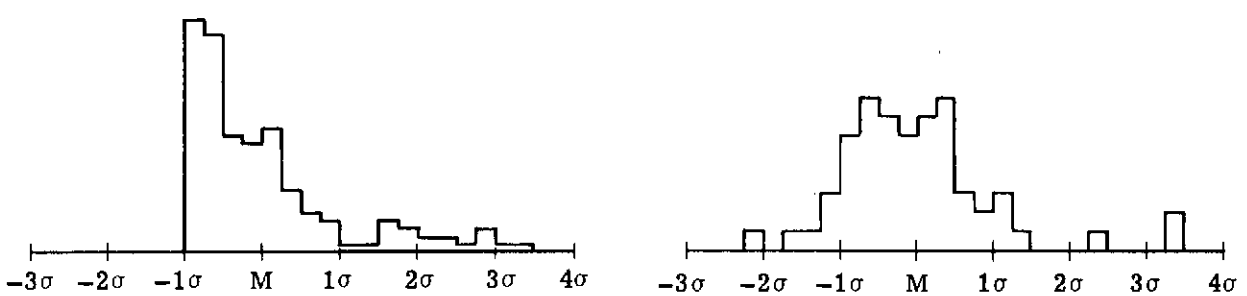
(d) Desert, Y_2 spatial component

Fig. 5-4 — Histograms of desert distributions in four components (in units of standard deviations about the means)



(a) Farms, derivative count

(b) Desert, derivative count



(c) Mountains, derivative count

(d) Cities, derivative count

Fig. 5-5 — Histograms of terrain classes in one component (in units of standard deviations about the means)

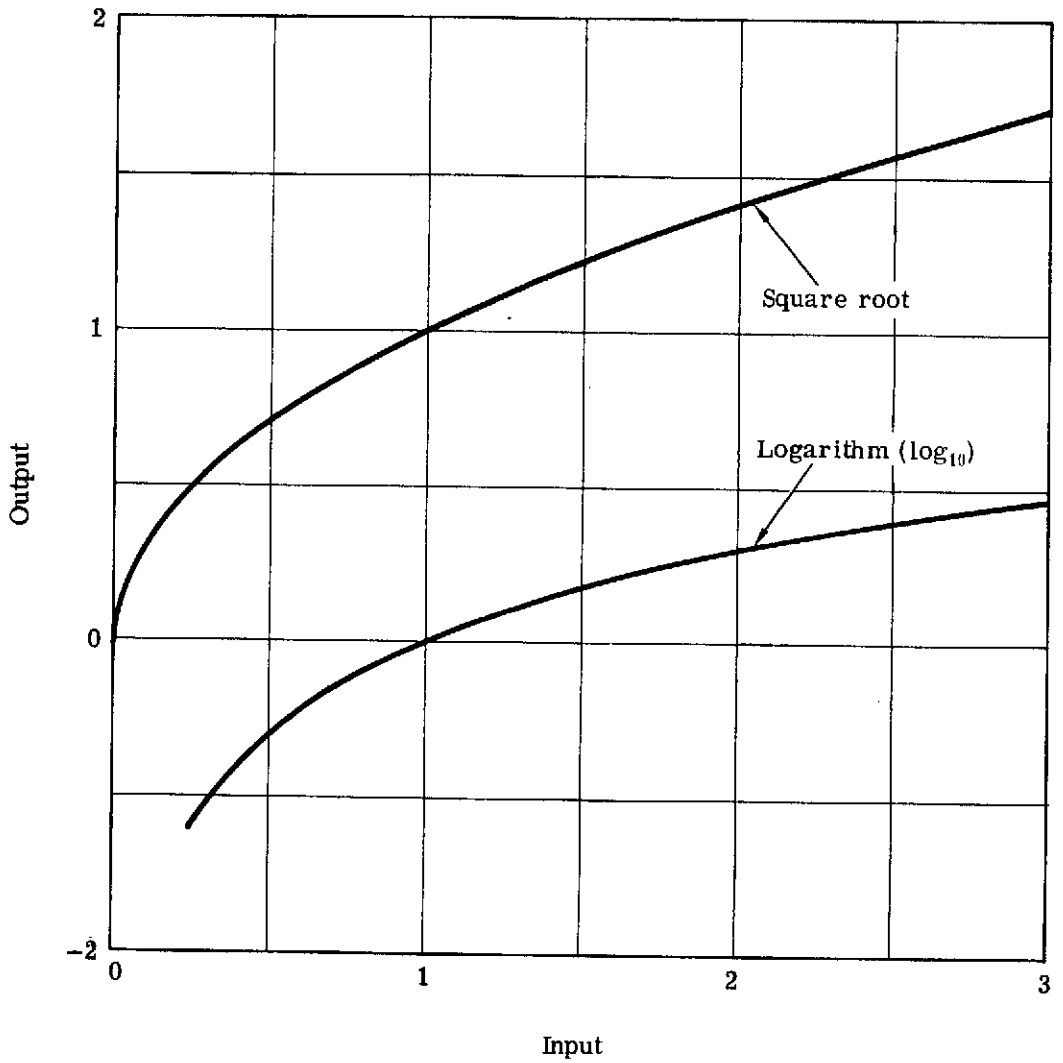


Fig. 5-6 — Typical nonlinear transformations

components range from 0.8 to 0.025 and have been optimized to produce excellent recognition results in all four classes: desert 89%, farms 97%, mountains, 96%, and city 95%. In arriving at an optimum transformation for each component, Gaussian distributions cannot be achieved for all classes. In other words, as one adjusts the transformation of say the MSS 5 band, the desert distribution may become more Gaussian while the city distribution may become less Gaussian. There is no reason to believe that all distributions can always be made Gaussian in all components by appropriate transformations. To reduce the possibility of such problems and maintain high recognition rates, the number of different classes should be kept small. (See Figures 5-7 and 5-8.)

Also, the importance of each component in the recognition of each class is not the same. A component may be essential for recognizing one class and relatively unimportant for recognizing other classes. Knowledge of this can be taken advantage of when selecting a transformation for a component. The transformation is adjusted so that a symmetrical distribution is obtained for the class that the component is most important. The effectiveness of each transformation can be judged from the recognition rates achieved.

5.2.4 Recognition Results

An important question that arises is the variability of the terrain class statistics with time of year and geographic location. If the statistics vary substantially, one must have statistics for each image to be processed. On the other hand, it would be desirable to have one set of statistics stored that would be applicable to a large geographic area for imagery acquired at any time of the year.

To explore this possibility, data from ERTS-1 images 1031-17325 and 1103-17332 were subjected to terrain classification using the statistics developed from training data of ERTS-1 image 1049-17324. The data from all three images cover the same general area around Phoenix, Arizona. The dates of acquisition are August 23, 1972, November 3, 1972, and September 10, 1972, respectively.

The data of image 1031-17325 was first subjected to the non-linear transformations described in the previous section and was then classified using the statistics from image 1049-17324. The recognition results are shown in Table 5-2. The detection rates for the four classes were: 93% for desert, 93% for mountains, 100% for farms, and 85% for urban areas. The accuracy of recognition of urban areas was not above 90% because many urban cells containing parks or golf courses were assigned to the farm category.

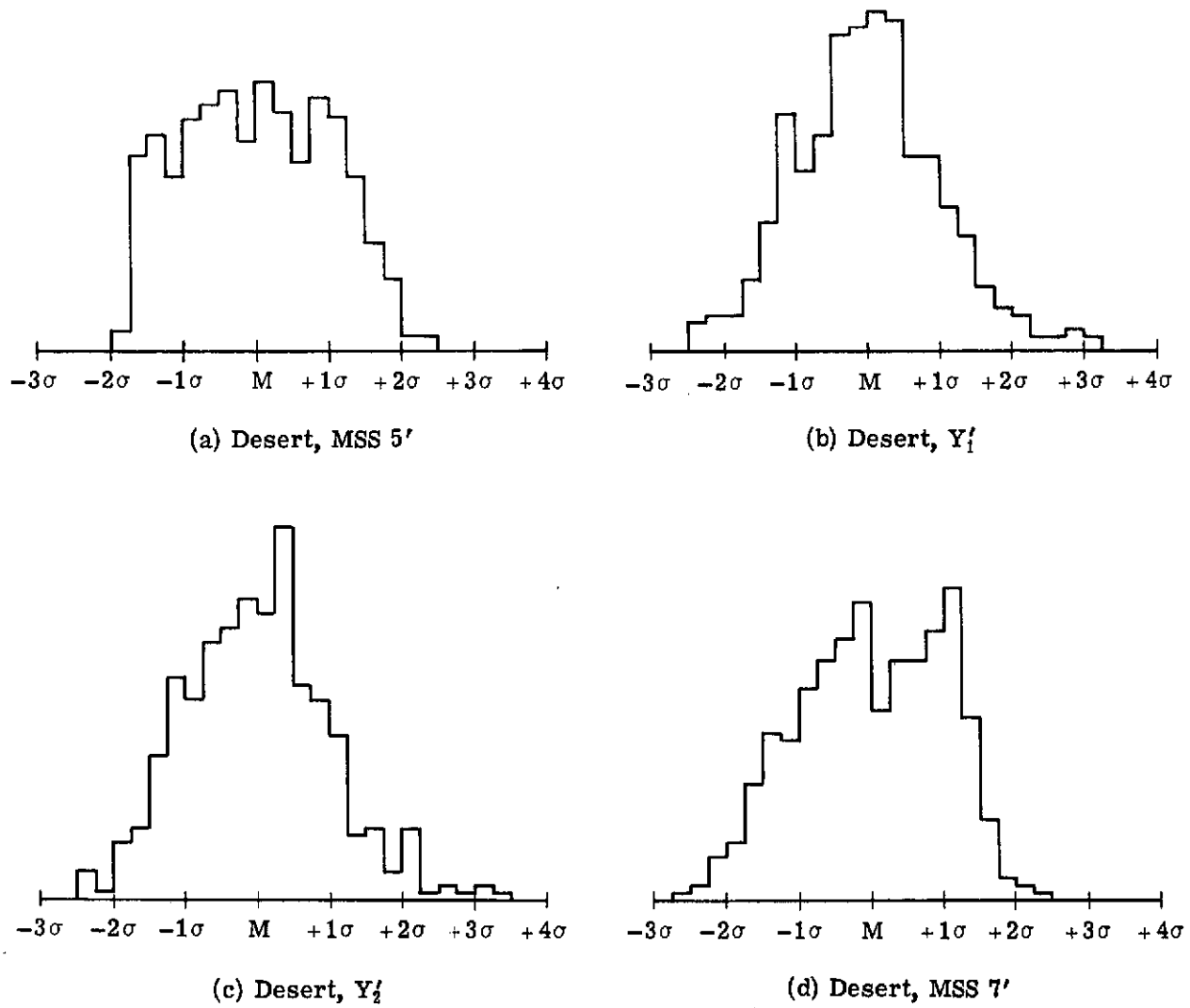
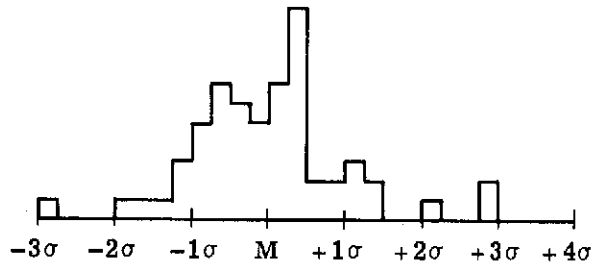
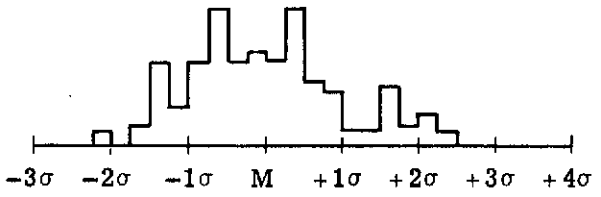


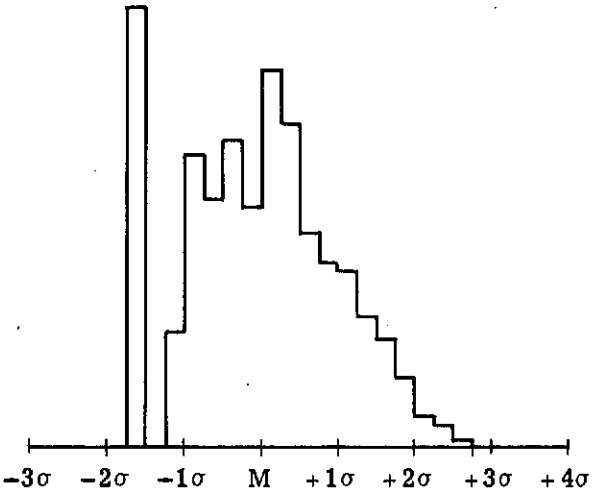
Fig. 5-7 — Histograms of desert distributions after nonlinear transformations, compare to Fig. 5-4



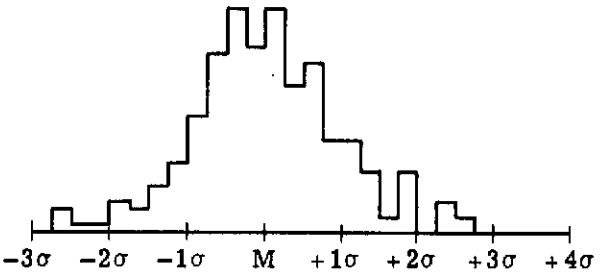
(a) Cities, derivative count'



(b) Mountains, derivative count'



(c) Desert, derivative count'



(d) Farms, derivative count'

Fig. 5-8 — Histograms of class distributions after nonlinear transformations—compare to Fig. 5-5

Table 5-2. Comparison of Classification Results - ERTS-1 Image 1031-17325

		Photointerpreter				
	CLASS	1	2	3	4	5
Computer	1	429	0	6	0	6
	2	1	407	0	10	7
	3	31	0	114	0	1
	4	1	1	3	57	11
	TOTAL	462	408	123	67	25

Table 5-3. Comparison of Classification Result - ERTS-1 Image 1103-17332

		Photointerpreter				
	CLASS	1	2	3	4	5
Computer	1	448	1	14	13	2
	2	17	443	7	6	31
	3	6	0	74	2	1
	4	0	0	0	20	0
	TOTAL	471	444	95	41	34

1 = Desert
2 = Farms
3 = Mountains
4 = Urban
5 = Riverbeds

For image 1103-17332, it was noted that the average brightness in all bands was substantially reduced due to lower solar elevations during the month of November. It was felt that the data should be compensated for the change in solar elevation. The average feature space vectors were determined for images 1049-17324 and 1103-17332. Then, the vectors of image 1103-17332 were linearly scaled so that the average vectors for both images would be the same. This compensation assumes that the change in statistics between September and November is simply a change in scale of the feature space. The compensation applied was only an approximation. Then, the compensated data was subjected to the non-linear transformations described in the previous section and was afterwards classified using the statistics from image 1049-17324. The recognition results are shown in Table 5-3. The detection rates were as follows: 95% for desert, 100% for farms, 78% for mountains, amd 49% for urban areas. It is obvious that the statistics employed were not optimum for image 1103-17332 despite the simple compensation applied to the feature space. Some mountain and urban areas were assigned to the desert and farm categories. Figure 5-9 shows the ERTS image number 1103-17332 with the recognition results superimposed as annotation.

Table 5-4 summarizes the recognition results for the maximum likelihood criterion and the heuristic algorithm. This table provides a comparison of detection rates for various sets of data and features being employed:

a. Data from ERTS-1 image 1049-17324 was used. The feature vectors were three-dimensional, and the features were the average cell brightness in the MSS 4, 5 and 7 spectral bands. The same vectors were used as training data to compute the class statistics and were then classified by the maximum likelihood criterion. The recognition results are what one could expect when using the maximum likelihood criterion in multispectral recognition.

b. Data from ERTS-1 image 1049-17324 was used. The feature vectors were four-dimensional with spatial features only. All other conditions were as in case a, above.

c. The spatial and spectral features of case b and a, respectively, were combined to form seven-dimensional vectors. All other conditions were as in case a, above.

d. The vectors of case c were non-linearly transformed as described in section 5.2.3. Then, the class statistics were computed from them, and they were subsequently classified by the maximum likelihood criterion.

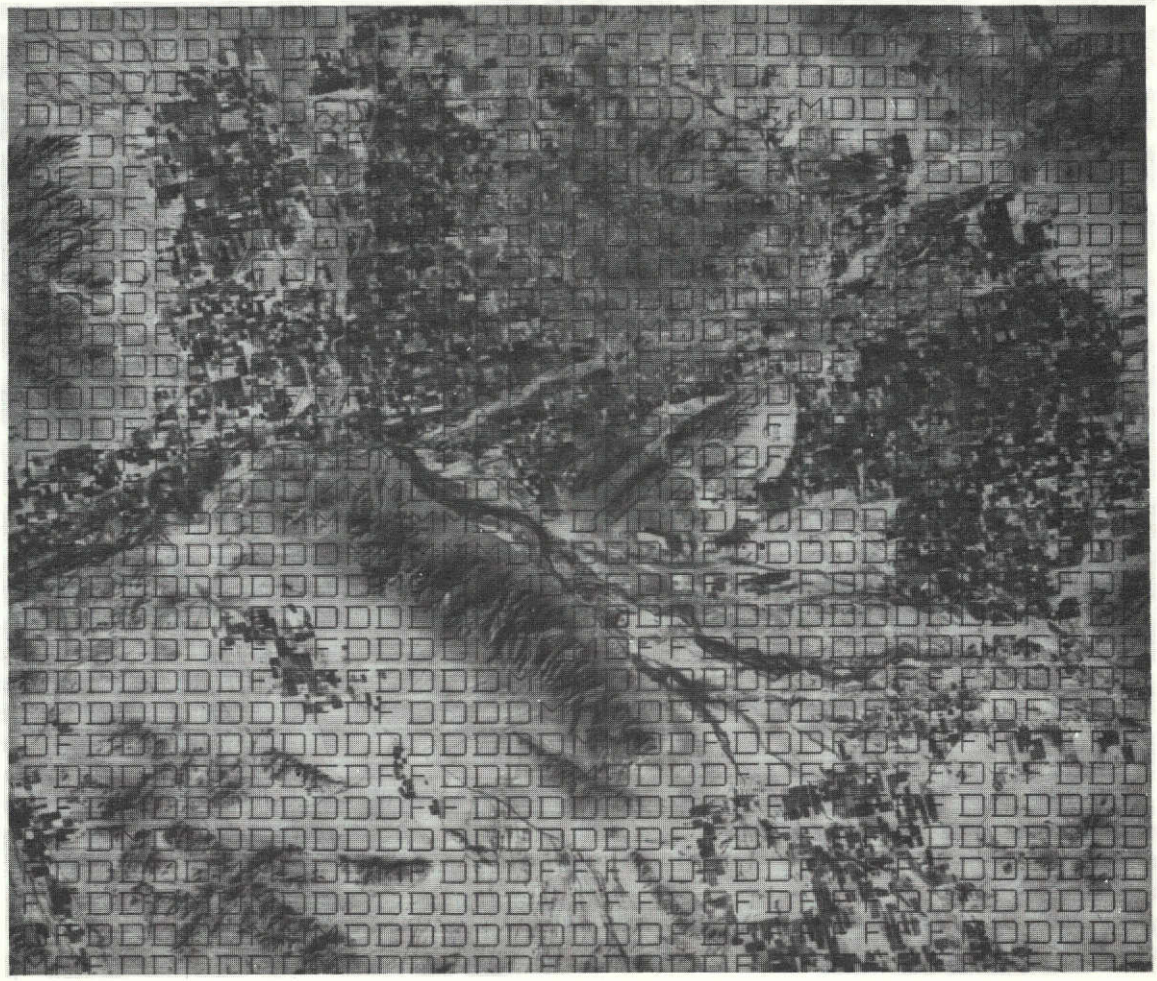


Fig. 5-9 — ERTS-1 image 1103-17332 and terrain classification results (maximum likelihood criterion, combined spectral and spatial features)

Table 5-4. Comparison of Detection Rates (%) with the Maximum Likelihood Criterion

Processing Conditions	Desert	Farms	Mountains	Urban Areas
1049-17324: 3 spectral features only	87	93	96	86
1049-17324: 4 spatial features only	8	83	97	70
1049-17324: Combined 7-dimensional feature vectors	54	92	97	92
1049-17324: Transformed 7-dimensional feature vectors	89	97	96	95
1031-17325: Transformed, 1049 statistics	93	100	93	85
1103-17332: Compensated to solar elevation and transformed 1049 statistics	95	100	78	49
1049-17324: Spatial features only, heuristic algorithm	97	89	80	74

e. The data was seven-dimensional vectors from ERTS-1 image 1031-17325. The vectors had been non-linearly transformed as in case d. The class statistics employed were from case d. (ERTS-1 image 1049-17324)

f. The data was seven-dimensional vectors from ERTS-1 image 1103-17332. The vectors were scaled so that the average vectors for images 1049-17324 and 1103-17332 were equal. This was an approximate scheme for compensating the lower solar elevation for image 1103. All other processing operations were the same as in case e.

g. The data was from ERTS-1 image 1049-17324 processed using only spatial features and the heuristic algorithm. (See section 5.1.)

Case b shows that the maximum likelihood criterion gives poor results with spatial features for the desert class. This has been attributed to the wide deviation of the spatial features from Gaussian statistics. (See Figure 5-4.) This conclusion is supported by case g which shows very good results when the spatial features are used with the heuristic algorithm rather than the maximum likelihood criterion.

Case a shows good results using the spectral features with the maximum likelihood criterion. The reason is that the class distributions for these features, though not Gaussian, do not deviate as sharply as the class distributions for the spatial features. This is fortuitous because there has been very extensive use of the maximum likelihood criterion with multispectral data, and we are not aware of any other investigations into the nature of the statistics (Gaussian or not).

Cases c and d show that by combining the spectral with the spatial features, the recognition rates for urban areas have improved substantially. Comparing case d to cases b and c shows the importance of the non-linear transformations in achieving high recognition rates. Also, comparing case d to cases a and g further indicates that combining the spectral and spatial features improves the recognition rate of the urban areas significantly.

For cases e and f, the recognition rates for desert and farms has improved over the corresponding recognition rates of case d. However, for mountains and urban areas, the recognition rates have progressively deteriorated. It appears, therefore, that the statistics of some classes in a given geographic area vary substantially with the seasons.

From the results of cases a through g, it appears that the spatial features have not increased the number of classes that are recognized

over what can be achieved with multispectral data only. However, the addition of the spatial features has increased the accuracy of recognition significantly for the urban classes.

There seems to be redundancy between the spectral and spatial features, particularly for the desert and farm classes. This, of course, is a characteristic of the geographic area and the ground resolved distance of the ERTS-1 data. For example, the spectral features are actually identifying live, dense vegetation rather than farms. In this part of southern Arizona, such vegetation is associated with irrigated farmland. In another part of the country, it may be associated with deciduous forests, grassland, etc. On the other hand, the spatial features are identifying farms, whether the farms have live vegetation (growing crops) or are fallow.

Furthermore, the ground resolved distance of the ERTS-1 data is large and only large geographic features (bays, lakes, mountains, farmland, desert, forests, large urban areas, etc.) can be identified. These tend to be highly correlated spectrally and spatially.

5.3 Clustering Algorithm

Various clustering algorithms have been investigated over the years. One of the established algorithms was developed at LARS, Purdue University and is part of the LARSYS software system. The algorithm has been described by Wacker[5] and Landgrebe and by Swain[6]

Conceptually, the clustering algorithms consist of the following parts:

a. Determination of potential clusters and their statistics.

b. Assignment of all feature vectors to the potential clusters.

c. A criterion of separability is applied to the potential clusters to determine which clusters are unique.

It turns out that part a is the most crucial and has the greatest impact on the success of the algorithm. There seem to be at least two approaches in determining potential clusters:

a. The first approach is to search the vector space for local concentrations of vectors.

b. The second approach is to make some initial assumptions about the number of clusters and their distribution in the feature

space. Then, each vector is assigned to a cluster, and the statistics of each cluster are computed. Repeating this operation several times, hopefully leads to convergence to the true clusters present in the vector space. This is the approach taken in the LARS clustering algorithm.

The first approach is the direct method of determining the cluster centers. However, the digital operation of searching the feature space for vector concentrations is very time consuming. When a large number of vectors are being clustered, the computer time required can become prohibitive.

The second approach has been employed to cluster large numbers of vectors because it produces simple algorithms and reasonable computation times. Unfortunately, there is no guarantee that these algorithms converge on the correct clusters every time. It might be necessary to perform the clustering two or more times with different initial conditions. If the same clusters are developed for all initial conditions, then a degree of confidence is achieved on the correctness of the results. However, if the clusters are dependent on the initial conditions, then ambiguities may still exist as to which group of clusters is the correct one.

In this investigation, a fast algorithm has been developed which determines the centers of clusters without any initial conditions. The centers are developed by computing the Euclidean distances between the vectors and selecting groups of vectors whose distances between each other are small. Instead of searching the seven-dimensional vector space for local concentrations of vectors, the algorithm determines which of the data vectors are surrounded by a number of closely spaced vectors. The computer time consummed by this algorithm is a few minutes because in addition to the algorithm characteristics, the number of vectors being clustered is 1,116.

The number of classes present in the data and their relative populations depend on the geographic area from which the vectors have been derived. All clustering algorithms work better when the classes represented in the data are equally populated. This, of course, does not occur naturally. (See Table 5-1.) However, if one assumes that at least 100 vectors per class are required to specify the statistics of a class with an acceptable accuracy level and that no more than 20 different classes exist, then it should not be necessary to cluster more than 2,000 vectors.

This condition can be satisfied since one has control of the number of vectors being introduced into the clustering algorithm as well as the geographic area they represent. For example, in the Phoenix area, one

could limit the number of vectors from the cultivated and desert areas while all the vectors from the urban and mountainous regions could be utilized. The conclusion is that a general knowledge of the geographic area by the analyst can be exploited to reduce the number of vectors processed by the clustering algorithm and improve the results of the clustering operation.

After the potential cluster centers have been determined, it is then necessary to employ a criterion of separability in order to determine which clusters are unique. Some of the potential cluster centers determined by the first operation are expected to be parts of larger clusters. After the criterion has been applied, some clusters are combined and others remain intact.

The criterion of separability that is being employed in this investigation is the divergence.*

Swain[6] has described another criterion which he terms the "Swain-Fu distance". This criterion is employed in the LARSYS program and is computed from the mean distance between two clusters and their "ellipsoids of concentration".

In this investigation, the divergence criterion has been examined extensively. Using vectors whose classification was already known (from ERTS-1 image 1049-17324) the class statistics were first computed, and then the divergence between any two classes was computed using the corresponding class statistics.

If the divergence is greater than 10, then the probability that a vector belonging to one of two classes can be correctly identified is greater than 80%. The divergence is also dependent on Gaussian statistics. Before the non-linear transformations of the vectors (when the classes were highly non-Gaussian), the divergence values varied between 400 and 6×10^8 and seemed to bear no relationship to the recognition rates achieved. After the non-linear transformations were completed, rendering the classes approximately Gaussian, the divergence values took on a more reasonable range (20 - 100) and generally speaking the higher divergence values were associated with pairs of classes between which very few errors were made.

The conclusion drawn from this analysis is that the divergence is a useful criterion of separability of clusters whenever the clusters are approximately Gaussian.

---

*Marill, T. and Green, D., On the Effectiveness of Receptors in Recognition Systems, IEEE Transactions of Information Theory, Jan. 1963, p. 11.

The number of clusters and the locations of their centers are determined by the two parts of the algorithm already discussed above: the determination of the potential cluster centers and the recombination of cluster centers which fail the divergence criterion.

The third part of the clustering algorithm recomputes the statistics of the combined cluster centers and assigns any remaining vectors to the nearest cluster on the basis of the maximum likelihood criterion. When all vectors have been assigned, the cluster statistics are recomputed.

During the reporting period the development of the clustering software has progressed sufficiently to test the characteristics of the algorithm described above. The test data are the vectors from the ERTS-1 image 1103-17332. The results of the tests were very good, but the software development has not been completed, so that vectors from other ERTS images could be processed.

6. CONCLUSIONS

Through the developed algorithms, it is possible to automatically recognize terrain types. By using the maximum likelihood criterion, it is possible to achieve high accuracy in the recogniton process and some degree of automation since heuristic algorithms need not be developed for each geographical area.

When the correct statistics are employed, the machine recognition appears to be more accurate than a human photointerpreter who has been constrained to use ERTS-1 color composites only. In this investigation the accuracy of the machine recognition was determined by a photointerpreter who employed aircraft photography (U-2 and low altitude photography). Stated another way, machine recognition appears to be more sensitive and can operate much closer to the resolution limit of the ERTS-1 imagery than the human photointerpreter.

The experimental results show that the accuracy of terrain recognition is affected by the statistics employed. The accuracy is best when the statistics are approximately Gaussian, and it is necessary to apply non-linear transformations to the feature vectors to achieve approximately Gaussian distributions. The class statistics vary appreciably between seasons, such as summer and fall so that for each geographic area, it will probably be necessary to determine the seasonal variation in statistics. It seems though that once known, the statistics can be used repetitively for a number of years until substantial land use changes have occurred in the geographic area. The experimental results show that the spatial features selected contain the spatial information necessary to recognize terrain types.

The experimental results also showed that combining the spectral data with the spatial features increased the accuracy of the machine recognition.

7. PROGRAM FOR THE COMPLETION OF THE INVESTIGATION

The clustering software development will be completed in September. Data from six ERTS-1 images will be processed through the clustering algorithm. The results will be described in the Final Report which will complete the investigation.

8. RECOMMENDATIONS

The results of this investigation are applicable to the production of land-use maps. An example of a land-use map developed by Poulton et al by photointerpretation of U-2 photography can be found in reference 8. Anderson et al[10] have described three levels of land-use classification. The terrain classes recognized in this investigation fall in the first classification level. The second level resembles Poulton's map and requires the use of much higher resolution imagery than ERTS-1.

Visual examination of U-2 photography of the Phoenix area discloses that a useful land-use map would require the division of the urban class into subclasses: residential, industrial, commercial, airports, etc. Also, the desert should probably be subdivided into several subclasses: flat desert, rolling desert, hills, flood plains, dissected land, etc. Unfortunately, the ground resolved distance of ERTS-1 is such that this level (the second level defined by Anderson) of classification is not possible either my machine or human interpretation. Nevertheless, the machine recognition results of this investigation are very encouraging and suggest that the interpretation techniques that have been developed can be directly applied to higher resolution data which will permit a subdivision of the major terrain classes and the generation of second level land-use maps.

One very important aspect of machine interpretation that requires further development is the generation of boundaries between the recognized classes.

9. REFERENCES

1. Gramenopoulos, N. and Alpaugh, H, Automated Thematic Mapping and Change Detection of ERTS-1 Images, Photointerpretation Results of ERTS-1 Images from the Brownsville, Texas Area, Technical Report of ERTS-1 Investigation. December 1972, Prepared for Goddard Space Flight Center, Greenbelt, Maryland 20771.

2. Goodman, Joseph W., Introduction to Fourier Optics, McGraw-Hill Book Co., Inc., New York (1968), p. 86.

3. Gramenopoulos, N. and Corbett, F. J., Automated Thematic Mapping and Change Detection of ERTS-1 Images, Diffraction Pattern Analysis of ERTS-1 Images, Technical Report of ERTS-1 Investigation, March 1973, Prepared for Goddard Space Flight Center, Greenbelt, Maryland 20771.

4. Gramenopoulos, N., Terrain Type Recognition Using ERTS-1 MSS Images, NASA SP-327, Symposium on Significant Results Obtained from the Earth Resources Technology Satellite-1, March 5 - 9, 1973.

5. Wacker, A. G., and Landgrebe, 1970, Boundaries in Multispectral Imagery by Clustering, Proceedings of 9th IEEE Symposium on Adaptive Processes, XI pp. 4.1 - 4.8.

6. Swain, P. H., Pattern Recognition: A Basis for Remote Sensing Data Analysis", LARS, Purdue University, Lafayette, Indiana NASA-CR-130757, LARS-111572 N73-17184.

7. Marill, T. and Green, D, Statistical Recognition Functions and the Design of Pattern Recognizers, IRE Transaction on Electronic Computers, December 1960, p. 472.

8. Poulton, E. C., Schrumpf, B. J., and Johnson, J. R., Ecological Resource Analysis from High-Flight Photography for Land Use Planning, Applied Remote Sensing of Earth Resources in Arizona, Proceedings 2nd ARETS Symposium, University of Arizona, November 2 - 4, 1971.

9. Marill, T. and Green, D., On the Effectiveness of Receptors in Recognition Systems, IEEE Transactions on Information Theory, January 1963, p. 11.

10. Anderson, J. R., Hardy, E. E. and Roach, J. T., A Land-Use Classification System for Use with Remote-Sensor Data, Geological Survey Circular 671, 1972, Washington.