LARS Contract Report 071079

Computer-Aided Analysis of Landsat Data Taken over Spain

Ramon Bermudez de Castro

Laboratory for Applications of Remote Sensing
Purdue University West Lafayette, Indiana 47906 USA
1979

TABLE OF CONTENTS

Acknowledgments

This training/research work was done under an IBM fellowship.

I want to thank, in a very special way, Dr. Luis Bartolucci for his inestimable advice as well as for the help he gave me at every moment. Without him the performance of this work wouldn't be possible.

I thank also all the LARS staff for their assistance, and in a special way to Barrett Robinson for his advice in the field of radiation theory, and to Ron Boyd for his help in operating the software.

Finally I want to thank my fellow, Juan M. Gonzalez, from INIA (Spain) for lending me the material with which I worked and for his opinions.

I. INTRODUCTION

This analysis has been done as a result of a training program, and for this reason it has many things that wouldn't be done operationally.

In trying to deal with the largest number of classes, in order to make the analysis as complex as possible, and due to the relatively small number of classes that are available or present in a small area, the number of training areas chosen was very large (17), and several of these training areas were used to test the classification after it was performed. Another point to take into account is the reference data used. These data, infrared photography, not only because its poor quality but because the different date between it and the CCT, are of little use when trying to identify land cover other than forest, urban areas, or other cover types that do not change with time. For this reason it is impossible to determine quantitatively the percent of correct classification of the site, as well as to assign an exact name to each of the different spectral classes.

This job has been done in order to develop a methodology to deal with a large number of classes by using the classification of the data into spectral families, and later on making the decision between members within a family.

II. LANDSAT: A DESCRIPTION OF THE SYSTEM.

ERTS I (currently called Landsat I) was launched on 23 July, 1972, and placed into orbit by the Delta launch vehicle. Since that date two more satellites have been launched, Landsat-2 and Landsat-3. Both Landsat-1 and Landsat-2 have been launched into a circular sun synchronous orbit with a 9.30 A.M. descending mode. The complete coverage cycle consisting of 251 revolutions, takes 18 days and provides complete global coverage between 81 degrees north and 81 degrees south latitude. The satellite makes 14 revolutions of the earth per day.

The satellite orbit has also been designed so that the swaths viewed during one 18 day coverage cycle repeat or overlay the corresponding swaths viewed on all previously coverage cycles. This facilitates comparation of imagery of a given area collected during different coverage cycles.

With two satellites operating simultaneously, the relative phasing of their orbits is nominally established to offset the 18 days repeat pattern of one spacecraft by 9 days with respect to the other. In this way the spacecraft overlay a given area every 9 days. Landsat has two subsystems by which it acquires data: the M.S.S. and the R.B.V.

The R.B.V. (Return Beam Vidicon) camera subsystem is used to obtain high resolution T.V. pictures of the earth. The subsystem uses three cameras to take pictures of earth scenes simultaneously in three different spectral bands (475-575 μm, 580-680 μm, 690-830 μm). Landsat-3 RBV views 4 subscenes at higher resolution (½) in one spectral band.

The three cameras are aligned in the spacecraft to view the same nominal 185 km square ground scene. When the cameras are shuttered the images are stored on vidicon photosentsitive surfaces, then scanners to produce video outputs.

The MSS subsystem is a sensor system that collects data of the earth in various spectral bands.

Landsat 1 and 2 MSS respond to earth reflected sunlight in four spectral bands, Landsat C carries an additional band responding to thermal IR radiation.

The MSS continually scans the earth in a 185.2 km swath perpendicular to satellite track. Scanning is accomplished in the cross track direction by an oscillatory mirror. Satellite motion along the orbit provides the along track scan.

The spectral bands to which Landsat MSS responds are:

Band 4	0.5 - 0.6 μm
Band 5	0.6 - 0.7 μm
Band 6	0.8 - 1.1 μm

Each MSS band (4 through 7) utilizes six detectors. Band 8 in Landsat B uses 2. Bands 4, 5 and 6 utilize photomultipliers, Band 7 utilizes silicon detectors. Due to the number of detectors the scanner forms six scan lines during each mirror sweep in the active scan direction (west to east). The instantaneous field of view of each detector for bands (4 through 7) is 79 μm, the IFOV for band 8 detector is 237 m.

The analog signals produced by the MSS detectors are digitized and formatted into a 15m bit data stream for transmission to a receiving station or for on board recording.

During every mirror retrace period the radiance from the earth sceen is blanked out by a mechanical shutter. The individual sensors of bands 4 through 7 are exposed to a rotating variable density wedge optical filter illuminated by on board calibration lamps. The resulting calibration data are subsequently utilized to perform radiometric corrections to the MSS detector signals.

In the image processing facility data is transformed into framed imagery with a 10% of overlap between frames.

In order to increase the signal to noise ratio and facilitate data storage and retrieval, data compression techniques are used to compress the high radiance level signals. The compression is used for bands 4, 5, 6 only. Band 7 is transmitted in linear mode. All data are then encoded in 6 bit digital words representing 64 discrete levels. The data are recorded in 6 bits regardless of whether the data was linearly processed or compressed.

After the data are transmitted to earth, the inverse of the compression function is applied to both video and calibration wedge signals before calibration, by means of hook-up tables. The input byte values 0-63 are output as 0-127. Two decompression tables are employed, one serves bands 4 and 6 and another band 5. Band 7 does not require decompression because it has been transmitted in linear mode. The ground resolution element of Landsat is approximately half a hectare (79 x 56 m) and for this reason its efficiency is not good for fields of that or smaller size. Further than that, Landsat has four spectral bands, two in the visible and two in the near IR and for this reason only objects with different behaviour within these bands can be distinguished with the spectral response values. Multitemporal analysis and photointerpretation can help when trying to discriminate between objects. Figures 1 through 3 show some characteristics of Landsat.

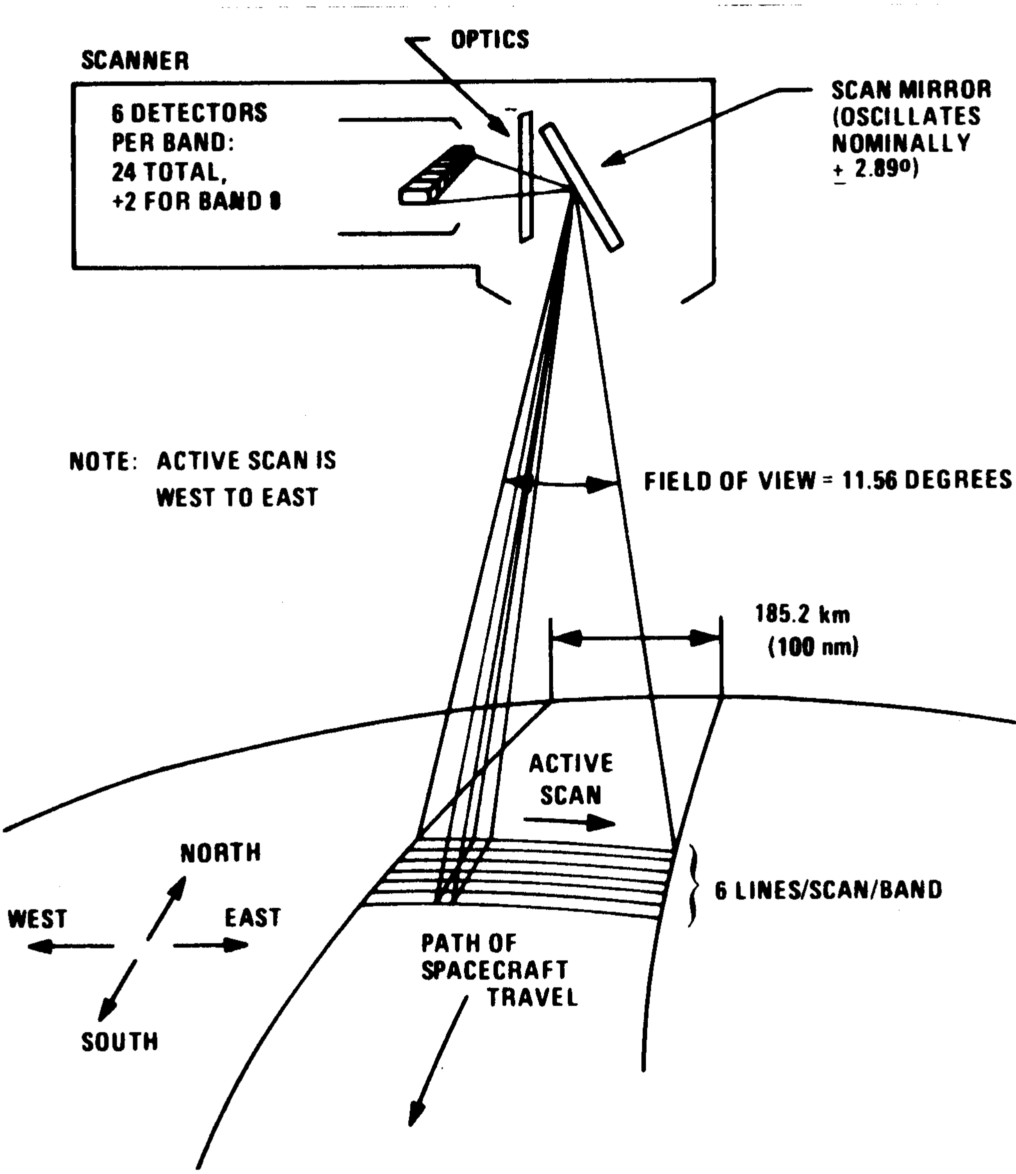


Figure 1. MSS Scanning Arrangement.

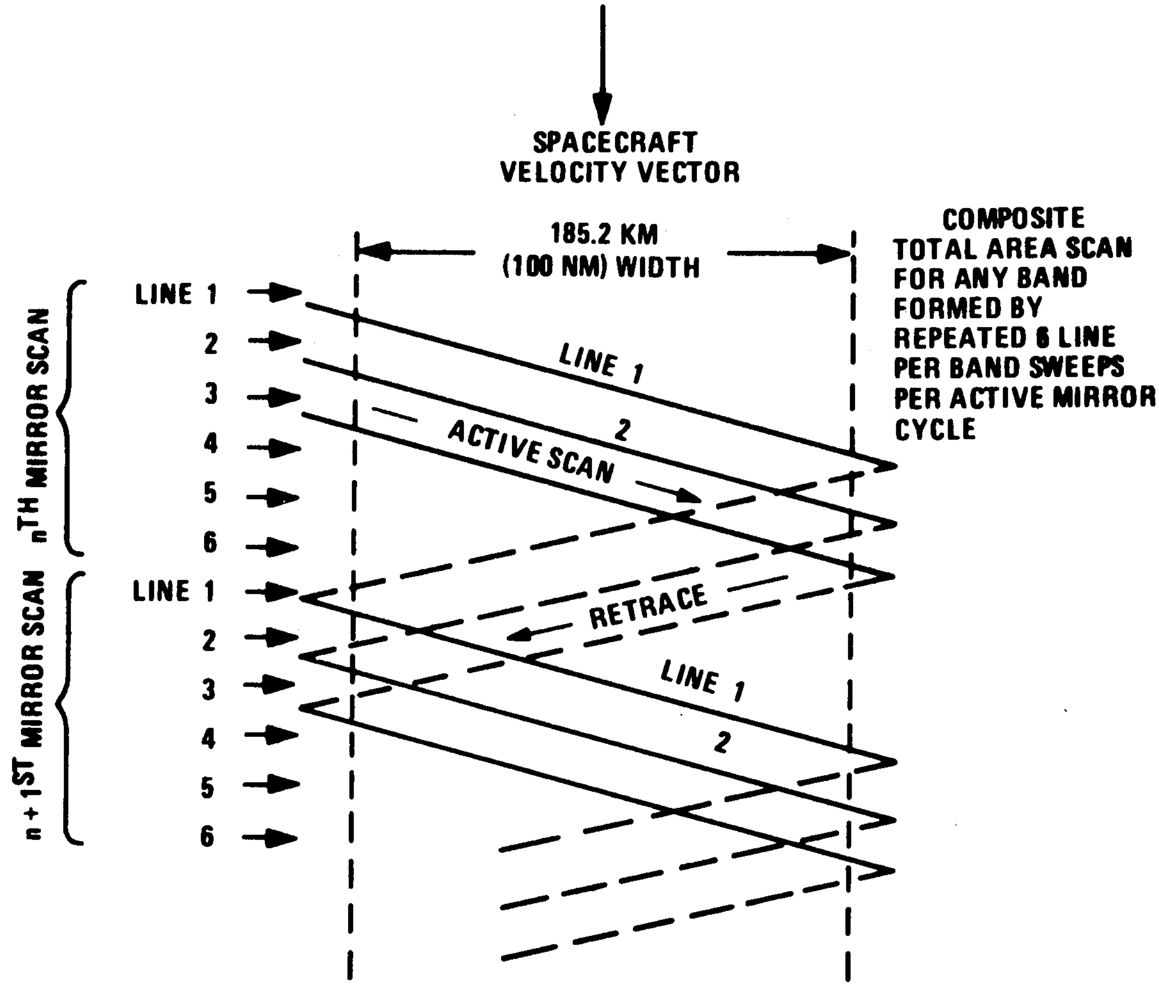


Figure 2. Ground Scann Pattern for a Single MSS Detector.

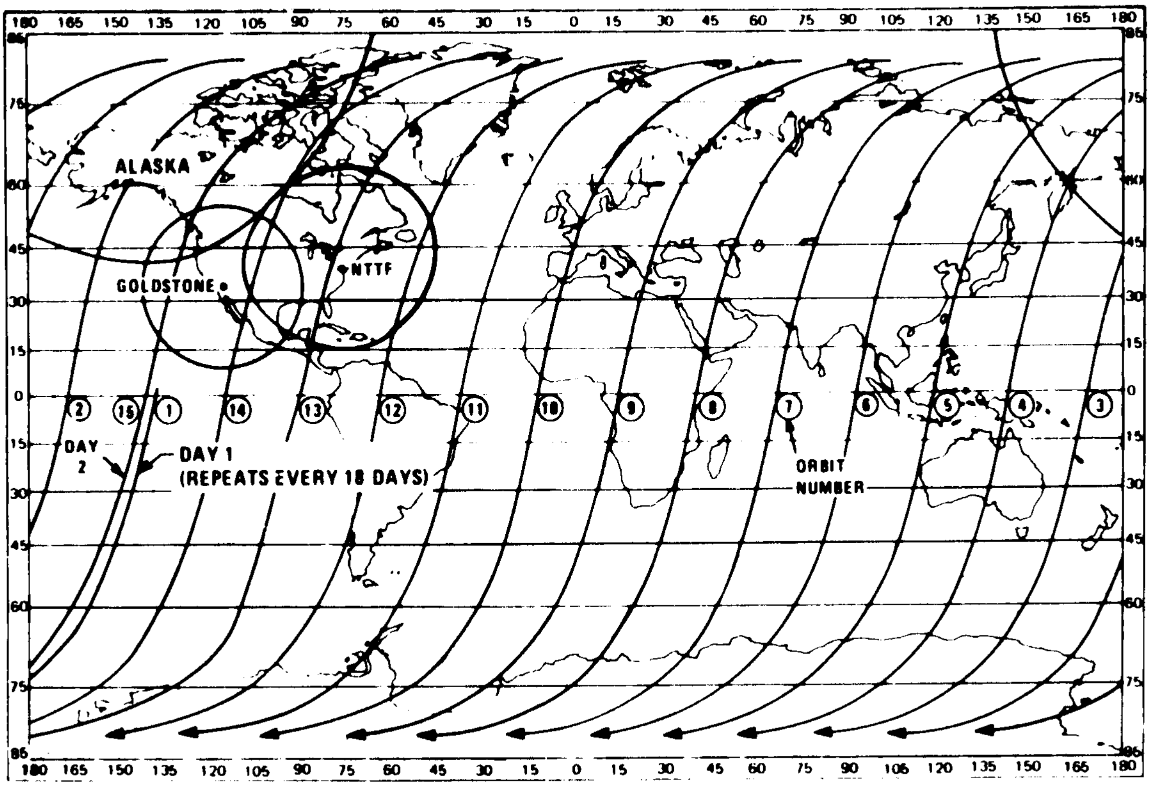


Figure 3. Landsat Ground Trace for One Day.

III. IMAGE PROCESSING.

Image processing techniques could be described as those oriented to extract the largest amount of information from a certain data set.

Two major types of treatment of the data are utilized: one is known as image enhancement, the other is multispectral classification. Image enhancement techniques are those which alter the image in order to improve its quality. The output is a transformed image with good visual qualities and can be analyzed and classified through conventional photointerpretation methods. Multispectral classification involves a group of computer aided analysis techniques with which we shall deal in later chapters.

Although both methods are related to each other, photointerpretation of Landsat images (either as individual band images or as color composites) behaves better when mapping broad categories of cover types because they can give a synoptic view of the situation of a given region enabling the understanding of certain significant ecological relationships. Photointerpretation is an inexpensive method of analysis and leads to good results when the level of detail needed is small as it is in the case of mapping land covers with small scales (1/250,000 or less). Nevertheless, when a higher level of mapping detail is needed, Landsat data in digital format, together with computer aided analysis techniques can provide much more information.

The number of gray levels that a normal human eye can distinguish is very low if compared with that of the CCT's which can be distinguished by a computer. Besides that, the analyst is enabled to work with spatial resolution elements of half a hectare and finally, the numerical analysis techniques offer the advantage of being able to work simultaneously with data from several spectral bands (multivariate space) which increases the capability of spectrally discriminating objects that in individual bands (multiple univariate space) would not be possible.

IV. DESCRIPTION OF THE ANALYSIS SITE.

The analysis is contained in the fram whose characteristics are shown in the next page. From this frame we have selected the region between lines 0 and 1100 and columns 0 - 2060. This region as we can see in one of the next figures, covers aproximately half of Borgos Province and contains the following sheets, scale 1:50,000, of the official map of Spain.

198 to 202

236 to 240

274 to 278

312 to 315

The digital data for this region was geometrically corrected. Color infrared photography of several zones were available. These photographs were taken 3 years after the date the Landsat data was taken.

V. ANALYSIS STEPS

A. DISPLAY OF RAW DATA.

The first step to be done in every analysis process, is the representation of the data in a pictorial format for visual inspection. Three ways are commonly used to display the data:

1 - Display of the data by the use of a cathode ray tube (image display.)

2 - Printing a map of gray levels, in which spectral ranges of the data input are represented by different gray levels.

3 - Printing an alphanumeric computer map in which spectral ranges are represented by different symbols.

Once the data are represented in a pictorial format we can see their quality as well as the distribution of clouds and furthermore choose the areas of interest for detailed study and choose the training areas needed to train the classifier. The last two methods were used in this work.

DISPLAY OF DATA

Both methods operate in the same way; assigning either a gray tone or an alphanumeric symbol to a range of data values based on equal likelihood of appearance. The ranges can be specified manually or automatically determined. In the manually operated form, the user chooses the gray tone, or alphanumeric symbol with which he wants to represent a given range of spectral values from a CCT. The automatic procedure is as follows:

First, the machine computes a histogram of the data for each channel that is to be displayed. Second, it integrates the histogram to locate a user specified number of data ranges which have an equal probability of occurring. Finally, it assigns a gray tone or alphanumeric symbol to each of the several ranges in order of increasing magnitude of data brightness. Both methods of data display give a large amount of information, but since the analyst is used to dealing with photographs rather than with alphanumeric symbols, it is more useful when dealing with the problem of locating specific areas, to choose the gray tone map as data display. Human eyes, used to dealing with photographic representations, can distinguish in the gray tone map certain spatial characteristics that would be confused in the alphanumeric one. However, it is difficult to locate individual pixels in such images. In this work in which the LARSYS system was used as the analysis tool, the GDATA processor was chosen to display the data in gray tone format. Pictureprint, which is the processor that gives the alphanumeric map was also used in specific cases. The input to both processors is the multispectral storage tape. Data from any wavelength available may be individually displayed. To operate either processor function we must specify:

the area to be displayed

the channels to be displayed

the number of gray levels desired

In this study the area located between lines (0 - 1100) and columns (0 - 2060) was chosen. The data was geometrically corrected and calibration values added. Channels 1, 2, 3, and 4 were displayed.

B. LOCATION OF THE AREAS WHERE GROUND TRUTH IS AVAILABLE.

Once we have the display of data, the next thing to do is to locate the areas where we have ground truth. This is accomplished by comparing the grayscale output with photographs or maps where these areas are represented.

In order to get the exact location of the areas in the computer map, several steps were followed:

First, the areas were located in maps with scale 1/50,000 and drawn in there.

Second, due to the difference in scales between maps (1/50,000) and data display (1/125,000) and to the lack of a pantograph, a reduction of the scale of the 1:50,000 maps dividing by a 2.5 factor was done. Next the location of the areas in those sheets was accomplished using a grid based on reference points such as rivers, towns, etc. to place the areas in the right place. Once the different areas were located in the corresponding sheets, these sheets were joined together forming a map that was positioned on the data display by superimposing recognizable features such as rivers.

Taking into account that there is a bending of the meridians, the areas were located approximately in the right place as were subsequently realized. Once the location of the areas was done, the coordinates (rows and columns) that define the areas were noted.

Another possible method besides using the photograph would be:

A recognizable feature (town, convergence of two rivers, a specific known area, etc.) located in both computer output and reference data, can be used to locate every other thing by measuring its polar coordinates versus the reference point in the reference map and applying those after transformation of scale. The angle will not vary. Of course an experienced analyst doesn't need to do all these things because he is used to dealing with these computer maps and can distinguish features that can't be distinguished by a non-experienced one. These two methods of locating certain areas in the gray level map are just two ideas about how to solve these problems. Both of them were used in this work and satisfactory results were obtained.

The next page shows a table with the coordinates within the image of the areas where ground truth is available.

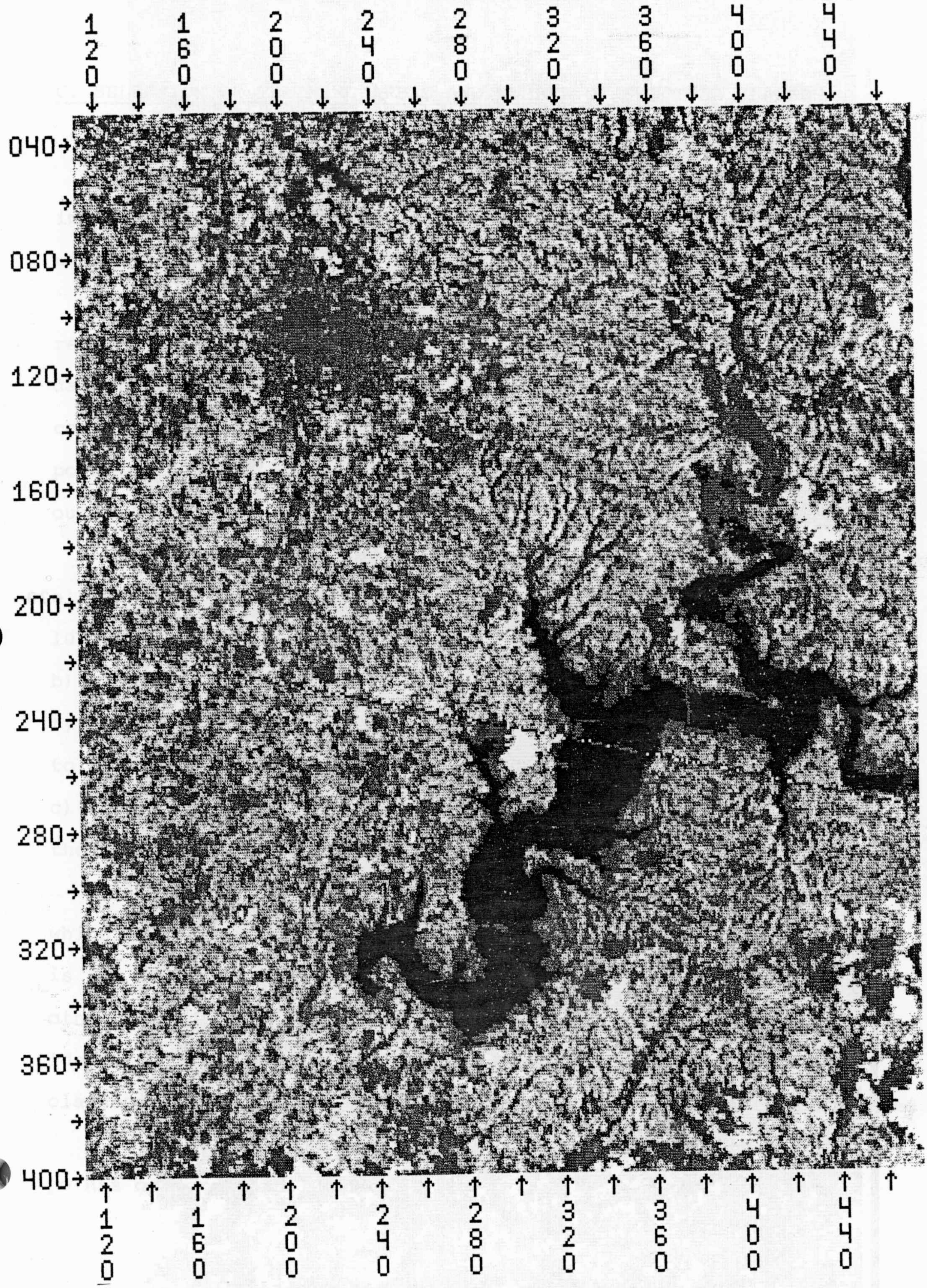
Description of the Training Areas

Training Areas	Lines	Columns	No. of Pixels
3	160-230	590-640	3500
6	500-570	700-800	7000
8	600-680	800-860	4800
9	690-770	1040-1140	8000
11	80-180	1670-1740	7000
12	280-360	610-710	4800
13	480-580	640-700	6000
17	490-570	980-1100	9600
18	610-680	1440-1540	7000
19	520-570	1580-1680	5000
20	640-700	1710-1830	7200
24	590-690	420-520	10,000
26	360-440	540-640	8000
27	660-720	880-990	6600
28	750-830	750-810	4800
32	620-690	1060-1140	5600
40	240-300	860-930	4200

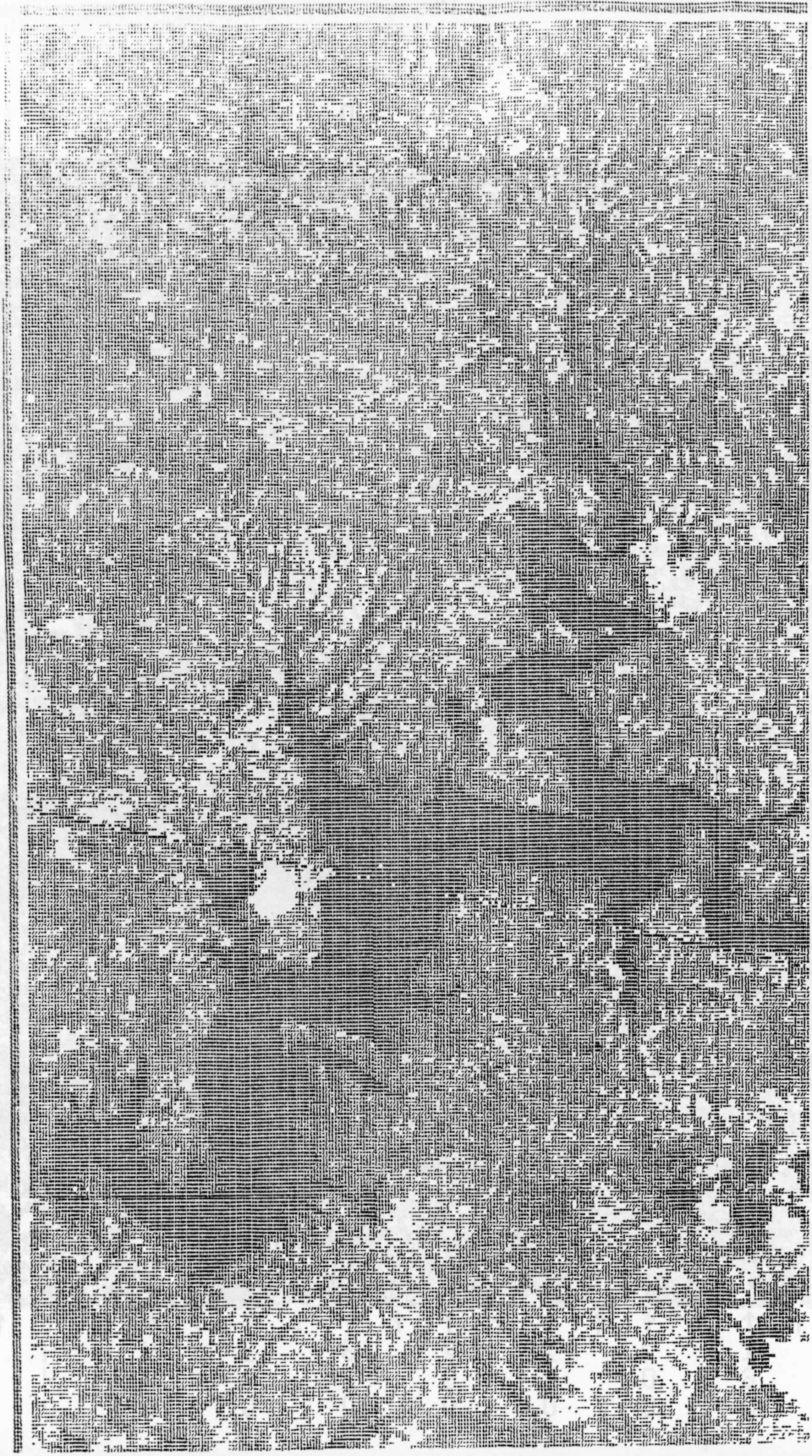
The next two following pictures show an output with gray levels and another with alphanumeric symbols, both of them with channel three.

The area is one of Bloomington (Indiana) and is courtesy of LARS.

A bad scan line can be seen in the gray level output.



120
160
200
240
280
320
360
400
440
040
080
120
160
200
240
280
320
360
400



C. SELECTION OF TRAINING AREAS AND NUMBER OF SPECTRAL CLASSES FOR EACH OF THEM

The selection of training areas is perhaps one of the most important steps in the MSS data analysis. Training areas selected with a good criterion will lead to a good classification result.

The use of training areas come from the need of the pattern recognition algorithms to be provided with samples of typical data from each class of interest. In other words, training the computer to recognize certain classes in order to compare every point with these classes when the classification is being carried out.

The training areas must be selected using the following criteria

a) each area must have no less than 1600 pixels and no more than 10,000

b) each area includes more than one cover type and every cover type must be included in at least one area although it is preferable to be included in two or more.

c) the training areas should be distributed uniformly throughout the area to be classified.

All these recommendations work fine in a supervised analysis which includes good reference data. However, if the reference data is not so good or if it is from another date or if for any other reason, as in this report, special objectives were not identified and the goal was to get as many spectrally separable classes as possible, in this case the following steps are recommended.

a) as before, choose training areas with no less than 1600 points or more than 10,000

b) instead of cover types think in spectral classes, in other words look at the gray levels output, select areas that contain more than one cover type (represented here by a set of gray levels) and every cover type (gray level) contained in more than one area.

c) training areas distributed uniformly throughout the analysis area. Although these two groups of recommendations appear the same, in the first group we are dealing with analysis objectives and we are defining the areas based upon ground truth where we can choose different cover types in a very exact way. As mentioned in the Introduction the treatment of the problem here would not be good in an operational way. One reason is that 17 training areas were chosen in a relatively small analysis site. The normal number of areas for this site should be a number between four and eight. In this case training areas were the same as those where ground truth were available.

Once the training areas have been selected, the next step is to decide how many spectral classes should be requested from the cluster algorithm. This is a difficult step in which there are no distinct rules. The best help is ixperience. However several ideas used in this work follow that may be helpful.

a) Using photographs or any other kind of reference data, try to see how many informational classes are separable. Take into account that the things that are different for a person may not be for the Landsat scanner. Specific cover types should not be confused with spectral classes because there are a lot of mixtures which are different spectrally but aren't representative of any specific cover type. Especially in this work where the agricultural crop fields are so small, many mixtures exist that are different spectrally.

One can also use a color composite or the gray levels output to see how many different classes one can expect in an area, and then add several more as security factor.

b) be generous when choosing the number of classes. A large number of classes leads to a more difficult analysis by increasing the amount of data one must deal with, but a very low number may lead to reclustering the area which is anti-economic.

Experience in this study shows that typically a good number of classes would be between 14 and 20 although some areas will require a smaller number and others a larger one. Once the number of classes were chosen for each area, the clustering was performed. The list of the number of classes chosen in this work can be seen on the following page.

SPECTRAL CLASSES PER TRAINING AREA

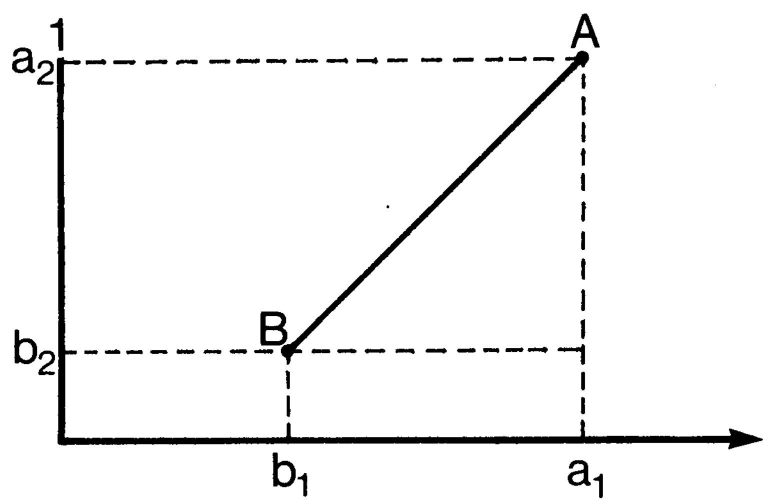
Training Area	Number of Spectral Classes
3	14
6	14
8	14
9	14
11	20
12	25
23	14
17	14
18	14
19	14
20	20
24	17
26	14
27	14
28	12
32	14
40	16

D. CLUSTER

Once the training areas and the number of classes per area have been chosen, everything is ready to cluster the different spectral responses (pixels) of each area into the number of groups or classes described above.

The purpose of clustering is, as mentioned above, to group the different spectral responses that comprise the different pixels into the prespecified number of classes, each class representing a group of pixels with similar spectral characterisitcs. A mean and a standard deviation is then used to characterize the distribution of data values that belong to the points belonging to the class.

The clustering algorithm is described below. The first thing to consider is the distance between points in the feature space. There are several ways to measure the distance between two points but perhaps the most familiar one is the Euclidean distance which in the two-dimensional space would be the calculation of the hypotenuse of a triangle using the Pythagorean theorem as shown in the figure.



For the bi-dimensional case the Euclidean distance between A and B, would be

$$D_{AB} = \sqrt{(\alpha_2-b_2)^2 + (\alpha_1-b_1)^2}$$

and generalizing for an n-dimensional space

$$D_{AB} = \left[\sum_{i=1}^{n} (\alpha_i-b_i)^2\right]^{-\frac{1}{2}}$$

Another distance to take into account is the L_1 distance. The sum of the distance components

$$D_{AB} = \sum_{i=1}^{n} \left|\alpha_1-b_1\right|$$

which on some computers is easier to calculate. Other distances such as Mahalanobis (etc. which apply different weights to the various components,) are available as interpoint distances, but the most commonly used in clustering algorithms are the Euclidean and L_1 distances. The second thing to consider is the distance between groups of points. Conceptually, the simplest ways to obtain the distance between two collections of points A and B should be to compute the average distance between all pairs of points for which one element of each pair is from A and the other from B. The most commonly used are the statistical distance measures which have the advantage of accounting for within group variability in the process of computing between group distance. An explanation of some of them will be given later in this chapter. The last thing to consider is the clustering criteria. A clustering criterion associates a measure of quality with each assignment of data points to clusters. Although there are many criteria to perform clustering, the most commonly used is to assign pixels to a cluster, in such a way as to obtain minimum distance between points within a cluster and maximum distance between clusters. At this point, a clustering algorithm will be described. The algorithm proceeds as follows.

1) Once every vector has been read the algorithm selects n vectors to serve as initial cluster centers (n being the number of classes asked for this area). The selection is arbitrary, but no two of the initial cluster centers must be identical.

2) Assign each vector in the data set to the nearest cluster c center

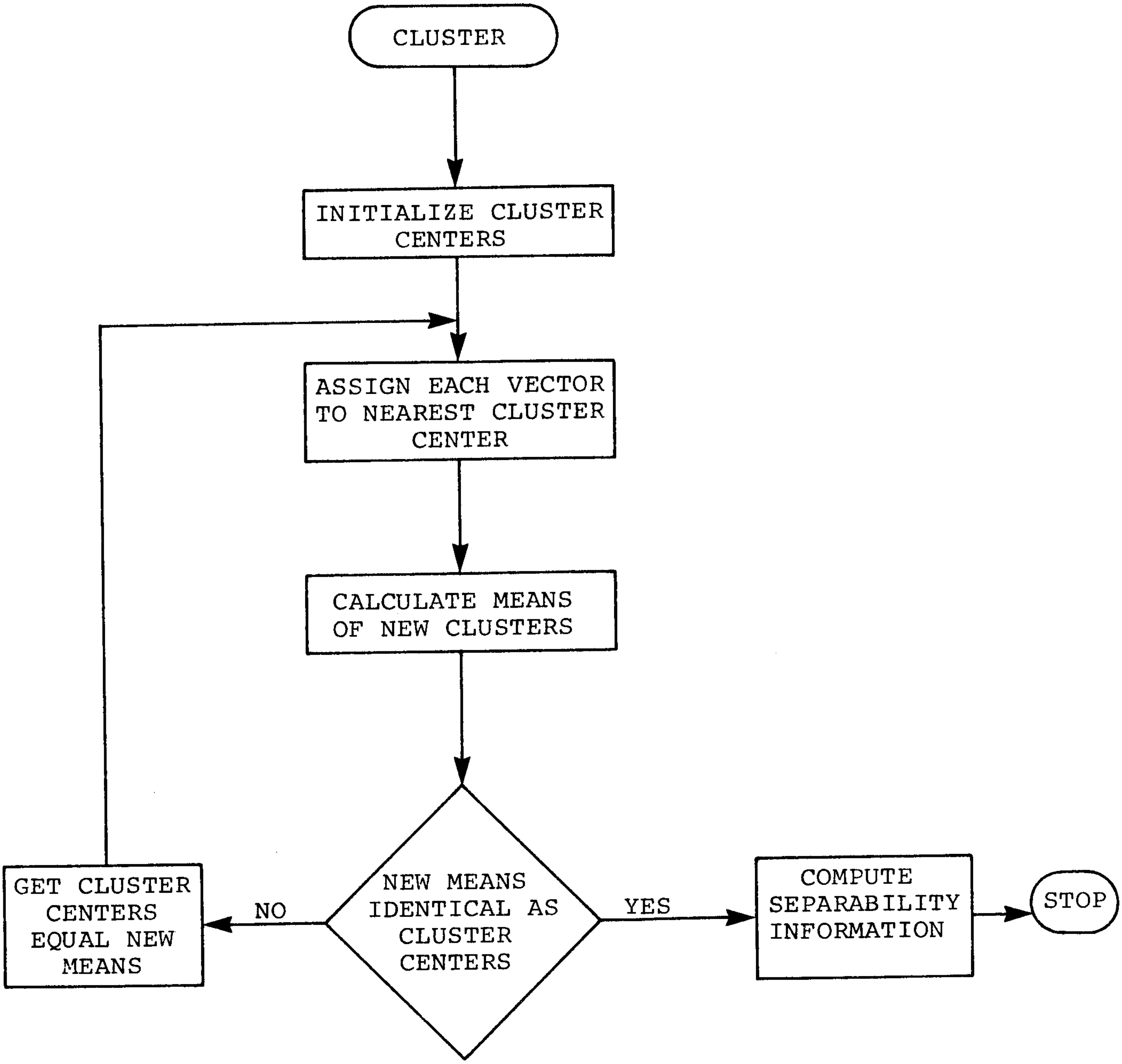
3) Compute the mean vectors for the data assigned to each cluster

4) If the new cluster means are identical with the cluster centers, then the cluster is completed. If not go to the point number 2 and repeat the process.

A scheme of the clustering algorithm could be as shown on the next page.

This method described is known as unsupervised classification because the analyst has no control over the decision regions.

Using the clustering algorithm of LARSYS over seventeen training areas, 254 classes were obtained with these 254 classes the analysis was done. The procedure to analyze these classes is described in the next steps of this chapter.


CLUSTER
INITIALIZE CLUSTER CENTERS
ASSIGN EACH VECTOR TO NEAREST CLUSTER CENTER
CALCULATE MEANS OF NEW CLUSTERS
NEW MEANS IDENTICAL AS CLUSTER CENTERS
NO
GET CLUSTER CENTERS EQUAL NEW MEANS
YES
COMPUTE SEPARABILITY INFORMATION
STOP

E. CALIBRATION OF THE STATISTICS FROM CLUSTERS AND PLOT OF THE DATA

The first step done, after the cluster, was the calibration of the means that represent the different classes. In order to clarify what is meant by calibration here, several concepts about it will be reviewed.

The first thing to remark is that this is not a calibration in the real sense of the word. Numbers obtained here don't have physical units which express values of radiance, reflectance, etc. But the combination of the four values belonging to the four bands, can represent a curve very similar to those obtained with spectrometers for different ground cover types. Without considering any of the preprocessing steps, let us consider the data as they are in the CCT.

Data from the Landsat CCT are digital counts, and these counts represent levels of energy. There must be a reference in order to associate these data in counts with any physical magnitude such as radiance, voltage, etc. Sensors in the satellite, transform the radiance input into a voltage output. There is a lamp inside the satellite to calibrate the data. The calibration is made by means of the response of the lamp in the four spectral bands. There is a masimum value of input radiance in every band, which has an associated maximum output voltage. This value, known as R_{max} is the value of the radiance emitted by the lamp in every spectral band. The values of R_{max} for each band are:

Band 1	2.48
Band 2	2.00
Band 3	1.76
Band 4	4.60

all of them in mw/cm^2 sr.

The output voltage in every band is then digitized and transmitted to ground stations. In order to reduce the large transmission band width needed by the huge volume of data, and at the same time to facilitate data storage and retrieval, and computer analysis, data compression techniques are used. These data compression techniques must be information preserving in order not to destroy the information available in the original data. The compression in Landsat data is made for the first three channels. The fourth channel is transmitted in what is called linear mode. (Figure 4)

Data compression is made by means of two quasi-logrithmic amplifiers - one for bands 4 and 6, and one for band 5. After the compression is done, data from the four bands are sent to ground in six bit words resulting in 64 possible digital counts. On the ground data is decompressed and radiometrically corrected. The first three channels then have 128 possible digital counts and the fourth channel remains with 64.

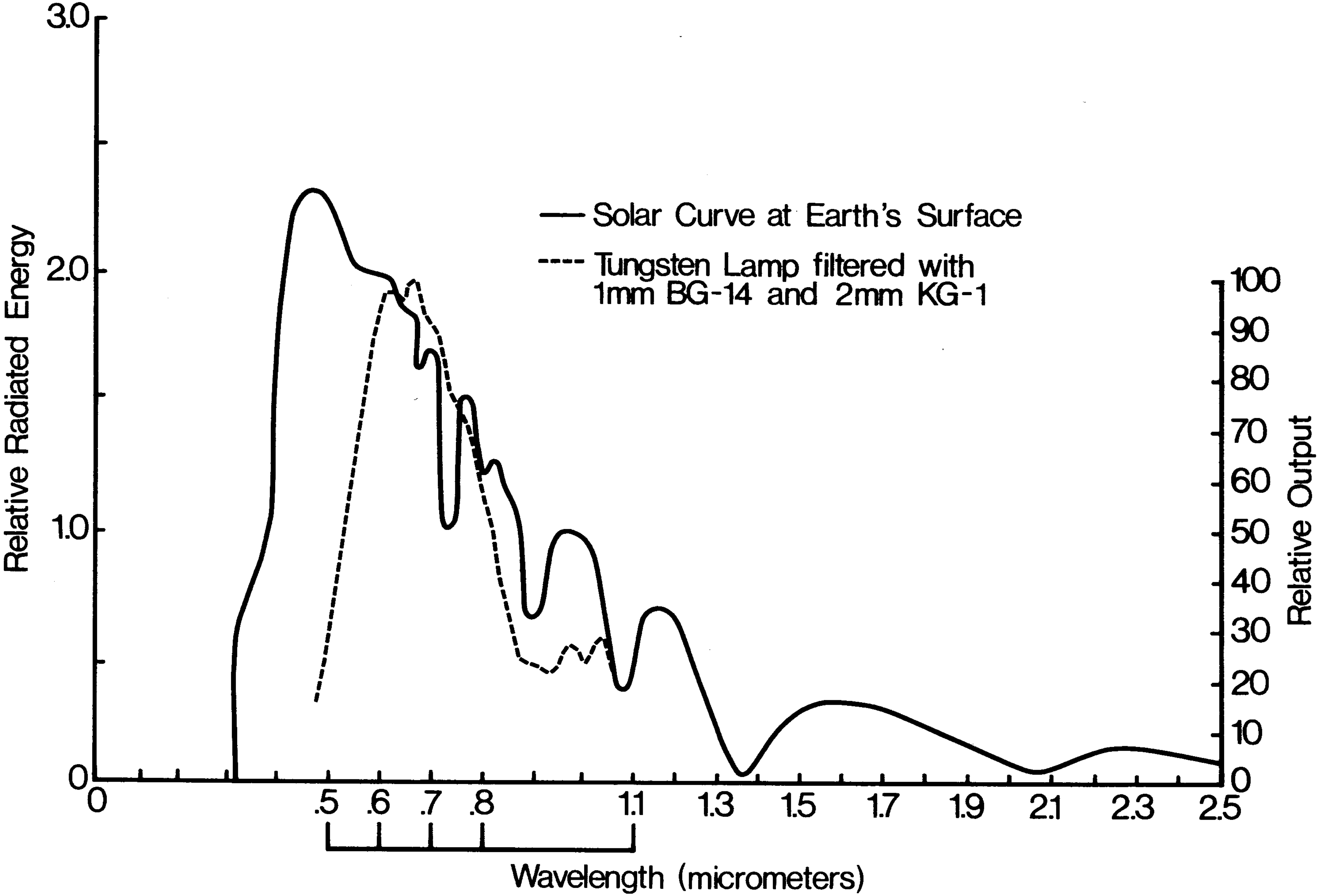
The value of radiance associated with a given digital count will be:

$$R = V_C \frac{R_{MAX}}{k}$$

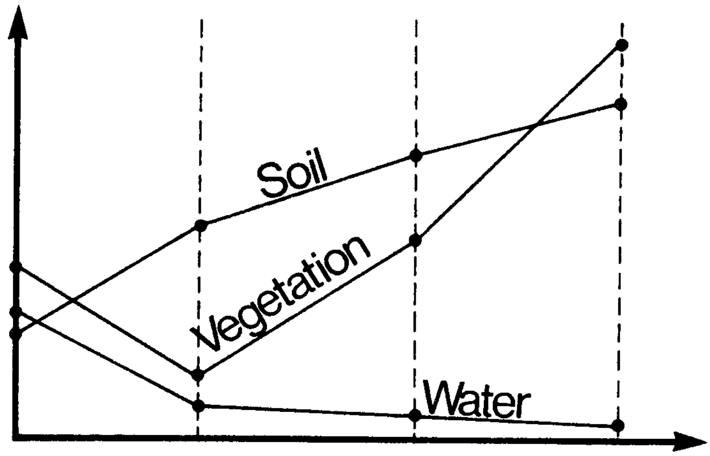
where k = 128 for bands 4, 5, 6

k = 64 " " 7

This factor k was not taken into account throughout this work and this is the reason why one can not talk about calibration of data. The number of counts has been multiplied by R_{MAX} for every channel, and the numbers obtained from this classification have been used to draw a curve that resembles a reflectance curve. The reason for the similarity between these curves and the reflectance


Solar Curve at Earth's Surface
Tungsten Lamp filtered with 1mm BG-14 and 2mm KG-1
Relative Radiated Energy
3.0
2.0
1.0
0
Relative Output
100
90
80
70
60
50
40
30
20
10
0
0
.5
.6
.7
.8
1.1
1.3
1.5
1.7
1.9
2.1
2.3
2.5
Wavelength (micrometers)

curves is the response of the lamp. The response of the lamp is similar in shape to that of the sun at the earth level because of the low modulation of the atmosphere on the spectral response of the sun. Once these curves are plotted one can decide to which of the following major ground cover types these curves (classes) belong.



Unfortunately, not all the spectral responses can be associated with pure types as the ones in the figure. Most of them are mixtures of several ground cover types. Due to this reason a classification in families of these spectral responses are necessary.

A family is a group of curves with similar characteristics (similar shape). All the curves within a family represent similar cover type or mixture of cover types and the differences in level of response among them could be explained as modifications caused by variations in topography, moisture content, soil composition, etc., over a given pixel. Curves within a family may be separable or not, but they belong to the same covertype modified of course by several factors.

Let's remark again that these curves can be used only as a tool for the analyst, and the values obtained are useful only in the way they can help the analyst to decide what condition of the cover type is associated to the class he is dealing with. Research

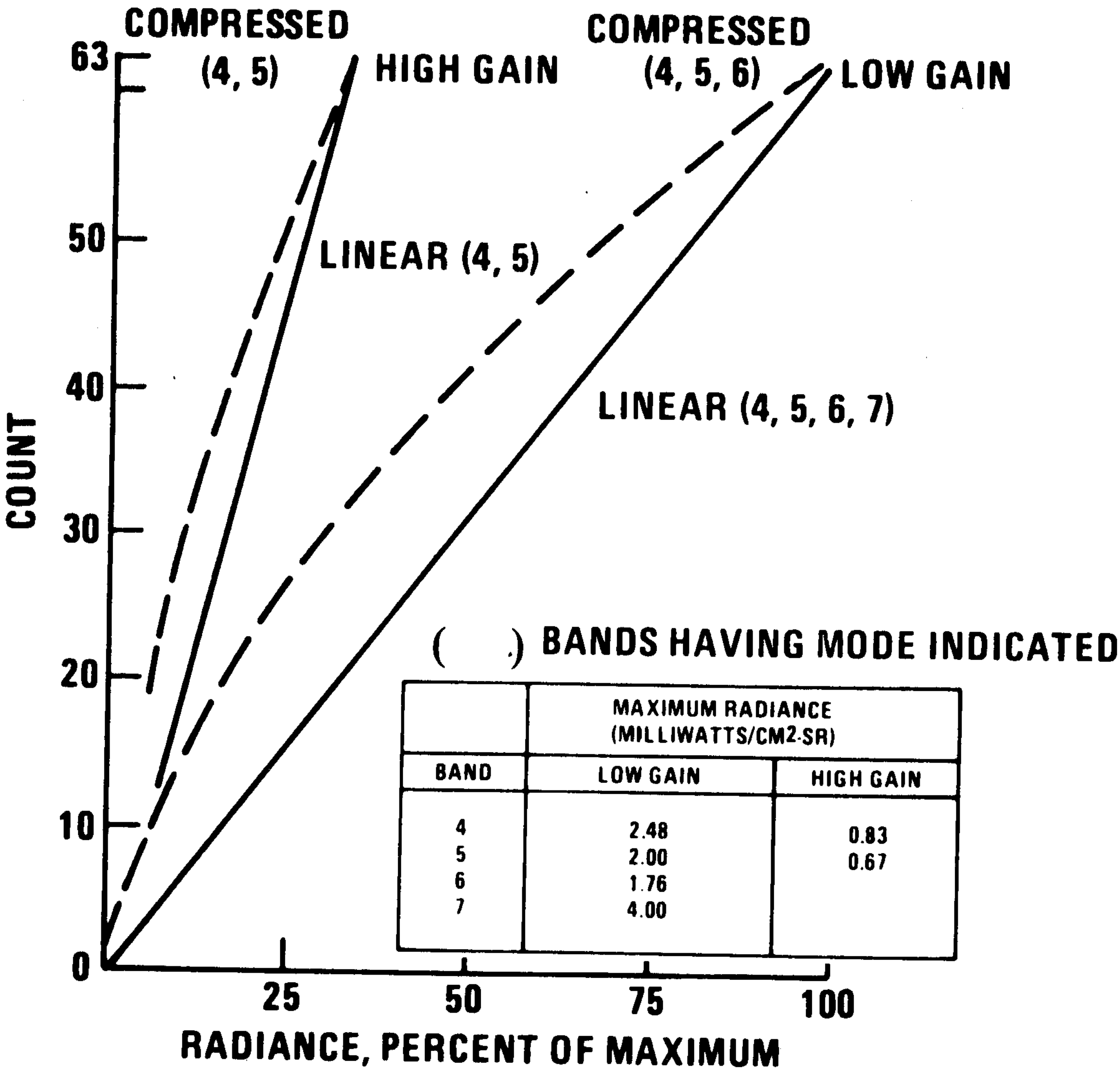


Figure 4. MSS Output vs Radiance, Compressed and Linear Modes.

is being carried out in order to see the influence of calibration on clustering.

After the calibration of the classes, and once the calibrated spectral curves were plotted, the analysis of them was performed. This analysis, based upon spectral families, is summarized in the following steps.

1. - Classification of the spectral curves into several spectral families according to certain parameters.

2. - Analysis, poolings or deletions within families. The decision making rule will be related to the shape of the curves, separability between them, covariance matrix and number of points.

3. - The objective throughout the analysis will be to get the maximum number of classes spectrally separable without taking into account the representativeness of the ground. The last step of the analysis will be to perform a classification of several well known text areas into the classes obtained, and to compare this classification with the ground truth available in order to associate spectral classes with cover types. (information classes)

F. CLASSIFICATION OF THE SPECTRAL CURVES INTO SPECTRAL FAMILIES

Once the spectral curves have been plotted the first thing done was classification of them into spectral families. Although it can be done by means of a computer program (a cluster program) specifying several parameters in order to classify the curves with highest correlation into the same family, in this work it was done manually. Admittedly doing it in this way is much harder, but as it was pointed out at the beginning of the report the goal of this work was to learn a methodology of analysis dealing with a large number of classes and by doing it manually, one could study every family as a particular case.

The decision as to the required number of families and as to which classes belonged to a family was made by studying the shape of the curves. There are three fundamental things in nature according to the Landsat bands: vegetation, soil, and water. Everything can be considered as a particular type of, or as a mixture of the three.

Based upon this, and following our objective which is to get the maximum number of classes spectrally separable, the 254 classes were divided at the very beginning into these three families. (Nevertheless, the problem is the large number of classes that can't be assigned to any of those families. Those classes are mixtures, and is hard to decide which family they belong to).

In order to group every class into a certain family and to avoid very large families, one criterion based upon the slopes between channels was used.

According to that criterion the following families were considered:

1 occurs between channels 1 and 2; slope 2 between channels 2 and 3; slope 3 between channels 3 and 4.

Family 1. -

Slope 1 less than 0, slopes 2 and 3 positives.

Family 2. -

Slope 1 positive, slopes 2 and 3 also positives but slope 2 bigger than slope 1.

Family 3. -

All the slopes are negatives.

Family 4. -

All the slopes are positives, but slope 2 bigger than slope 1.

Family 5. -

Slopes 1 and 3 positives, and slope 2 negative.

The families were chosen by examining the 254 curves plotted as described above. Once every class was assigned to a family a graph where every class in a family was represented by the mean of the visible channels versus the mean of the IR channels (bi-spectral plot) was requested for every family, in order to see if the classes were assigned to the right family.

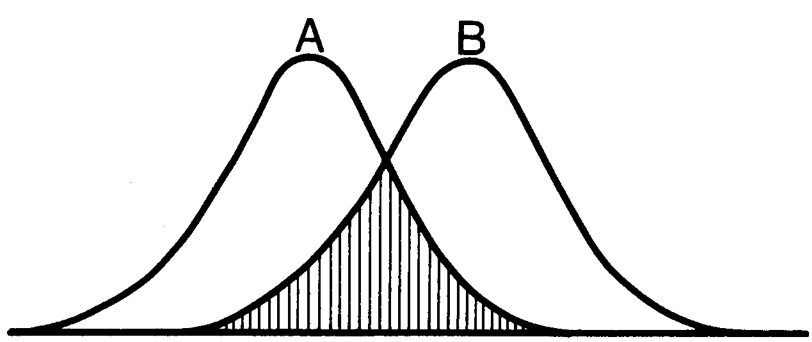
This bi-spectral plot was done due to the fact that the families were chosen manually and many mistakes could be made. By using both aids the criteria of the analyst is stronger. An analytical study of this bi-spectral plot will be done in the next chapter.

G. Analysis of data with the goal of getting maximum number of classes spectrally separable

The next step to be done was to compute all distances among classes within a family, in order to know which of them are separable and which of them are not.

Several methods are used in pattern recognition, in order to compute the distance between pairs of classes. Some of them will be discussed here.

Some mention was made of the distance between groups of points when talking about cluster, let's concentrate now on several ways to measure these distances. Given a pair of classes A and B represented by their normal probability density functions as shwon below



it appears evident that the greater the overlap between classes, the greater the probability of considering a point belonging to A as a point of B, and therefore the smaller the distance between classes. As we can see in the figure, if the standard deviation of the classes is held constant as the distance between means increases, the overlap decreases, and therefore the separability increases. On the other hand, when the distance between means is constant but the standard deviation increases, the overlap increases and the separability decreases.

Based on those concepts normalized distance between the means was defined as

$$\text{dnorm} = \frac{\mu_A - \mu_B}{\sigma_1 + \sigma_2}$$

This parameter gives an idea of how separable two classes are.

But looking at the dnorm expression, we can assert that this distance does not behave in the way we would like, due to the following reasons:

1 - The dnorm will be 0 when $\mu_1 = \mu_2$ without taking into account the standard deviation.

2 - The dnorm can be used only in a 1 dimensional space, and we are working in multidimensional ones -

For these reasons two distances, J-M distance and divergence will be introduced.

DIVERGENCE

Digergence is a measurement of statistical separability used in pattern recognition. It is related to the likelihood ratio Lij for two classes i and j

$$L_{ij} = \frac{P(x/w_i)}{P(x/w_i)}$$

For a multidimensional space it can be expressed as follows:

$$\text{Dij} = 1/2 \ \text{Tr}\left[(\overline{Z}_i - \overline{Z}_j) \ (\overline{Z}_j^{-1} - \overline{Z}_i^{-1})\right]$$
$$+ \ 1/2 \ \text{Tr}\left[(\overline{Z}_i^{-1} + \overline{Z}_j^{-1}) \ (U_i - U_j) \ (U_i - U_j)^{-1}\right]$$

where Tr (Δ) denotes the trace of matrix A and U; $\overline{Z}$; are the mean vector and covariance matrix respectively.

The expression consists of two terms, the first one represents a contribution resulting only from differences in covariance matrices. The second term is a normalized distance between means.

As it can be seen Dij won't be "zero" unless both the mean vectors and covariance matrices are identical. This divergence is a pairwise distance measure and for this reason it is used to compute the separability between two classes. If we want to know if a set of classes has good separability among them, we would use the average divergence

$$\text{Dev.} = \sum_{i=1}^{M} \sum_{j=1}^{M} p(w_i)\, p(w_j)\, D_{ij}$$

J-M DISTANCE

The Jeffries-Matusita distance, like the divergence, is a measure of statistical separability. It is defined as follows:

$$J_{ij} = \left\{ \int_x \left[\sqrt{p(x/w_i)} - \sqrt{p(x/w_j)} \right]^2 dx \right\}^{1/2}$$

This distance is a measure of the average divergence between two class density functions. When the classes have normal density functions, the J-M distance is

$$J_{ij} = \left[2(1-e^{-\alpha}) \right]^{1/2}$$

where

$$\alpha = 1/8(v_i - v_j)^{-1} \frac{(\bar{Z}_i + \bar{Z}_j)^{-1}}{2} (v_i - v_j)$$

$$+ 1/2 \log \left[\frac{|(\bar{Z}_i + \bar{Z}_j)/2|}{\bar{Z}_i \cdot | \cdot |\bar{Z}_j|)^{1/2}} \right]$$

As it can be seen observing the expression of the J-M distance, it has a saturating behavior that has been verified experimentally. Also, through experimentation it has been shown that a saturating transformation of divergence can provide similar results and is somewhat easier to calculate than J-M distance. This saturating form of the divergence is called transformed divergence and is given by

$$Dij^T = 2 \left[1 - \exp(- Dij/8)\right]$$

where Dij is the divergence described above. LARSYS, which is the software used in thie analysis procedure, calculates the transformed divergence. The maximum possible value given by the software is 2000, representing complete saturation of the measure. The transformed divergence varies from this value (2000) that represents complete separation between classes to representing identical classes and its relationship to classification accuracy is shown in Figure (5).

The way in which the analysis was performed, was applying the transformed divergence to measure separability between classes within a family. Two classes within the same family with a very low separability, represent the same thing, and in that way the decision to pool them or delete one of them can be verified.

The way in which decisions were made in this work could be explained as follows. Once the classes were grouped in families, and plotted onto a bi-spectral plot, the distance between classes which were not on the border of another family was measured by means of the separability function. The classes which were on the border of another family, were not pooled or deleted in this first step, in order to use them as a link between families.

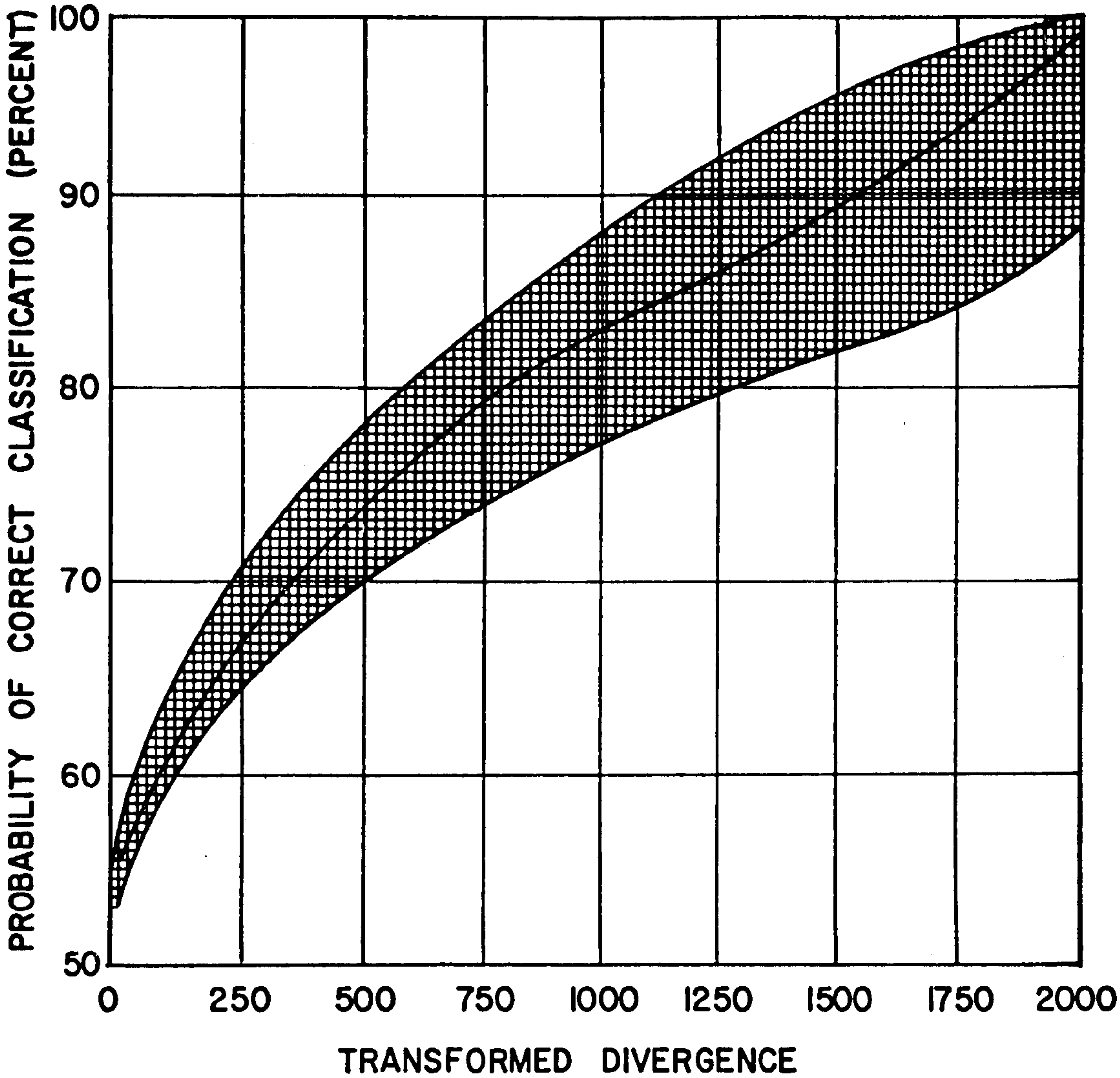


Figure 5. Empirical relationship between transformed divergence and percent correct recognition (from Swain, 1973).

To make decisions between classes we observed the following parameters:

1. - (transformed divergence)

2. - Standard deviation in every channel

3. - Number of points in every class

Given two classes with a very low separability between them, belonging to the same family, and one with larger standard deviations and lower number of pixels than the other class with largest standard deviation and lowest number of points, which will be properly represented by the other was deleted. A pooling will be made when the separability is not so small, or when the classes have similar standard deviations and numbers of pixels or when having low separability buth the shapes of the spectral curve were not exactly the same. Having done this, 26 spectrally different classes were obtained.

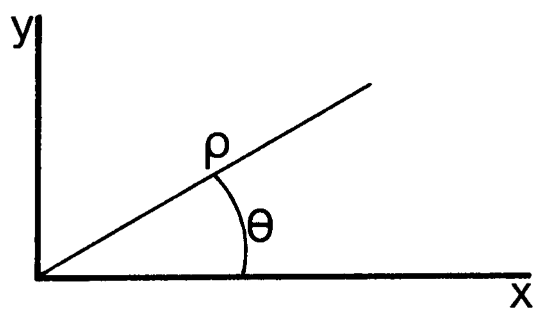
To compare the spectral curves among them a program that calibrates their mean values and print a plot of the response versus channel was used. This program prints beside the means, the standard deviations in a plot. An example of this printout is shown in one of the figures.

H. DESCRIPTION OF THE BISPECTRAL PLOT.

A very helpful tool when analyzing data is to obtain a scatter diagram of the spectral characteristics of the cluster classes by plotting the average of the means of the visible bands versus the average of the means of the IR bands. The kind of information that we can obtain from this diagram is based upon the following ideas: 1) observing nature one can assert that the earth surface is composed of three major components - vegetation, soil and water. These components are found in various combinations which correspond with the different cover types that we can find in a scene. 2) looking at the reflectance characteristics of vegetation, soil and water, one can observe that the response resulting from a given incoming radiation from vegetation will be smaller than the response from soil in the visible portion of the spectrum but higher than soil in the IR. Response for water decreases monotinically with increasing wavelength, in the visible and near IR portions of the spectrum.

Considering the four Landsat bands, two in the visible and two in the near IR, the response in each channel is the amount of energy reflected by a given target in that band.

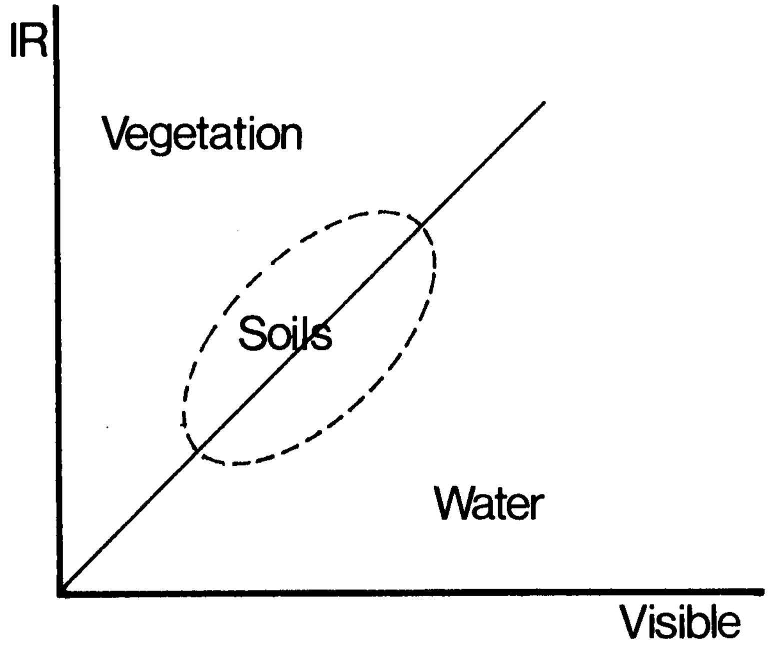
A transformation by which the four dimensional space of Landsat data can be converted into a two dimensional space involves computing the average response in the visible channels (y-axis) and the average response in the IR channels (x-axis)



Every point in the measurement space will be defined by a vector (ρ, θ) where $\rho = (x^2+y^2)^{\frac{1}{2}}$ and $T_g\theta = y/x$. For vegetation, we observe that the greener the vegetation is, the lower its response in the second channel, (absorption of chlorophyll) and the lower its coordinate y. On the other hand, the healthier the vegetation is, and the greater its percentage ground cover, the larger its response in the IR and its coordinate x.

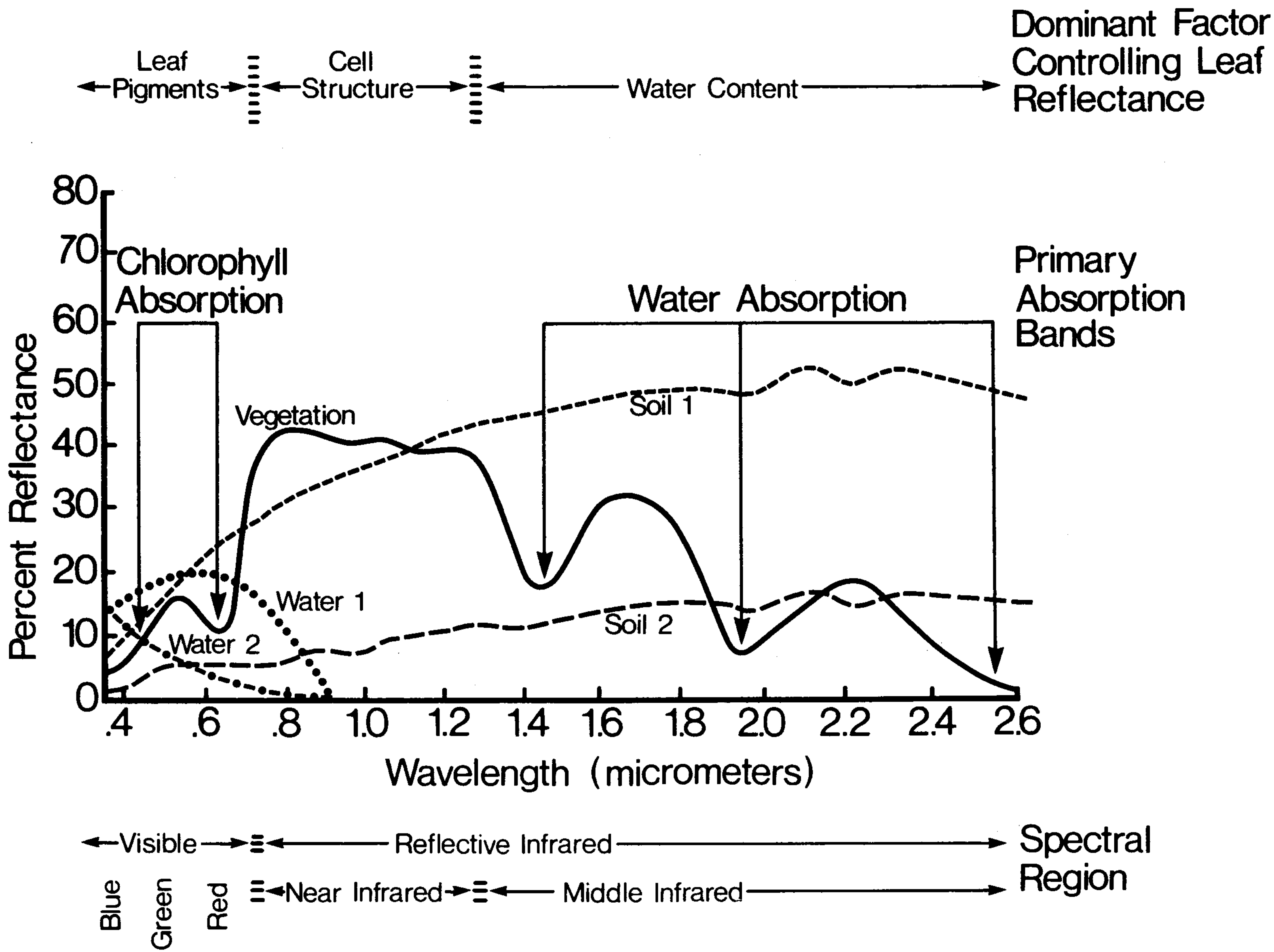
For these reasons the higher the quotient between y and x (or $T_g\theta$) the greener and more exhuberant is the vegetation. When the percent of ground cover of the vegetation decreases, and parallel to this more soil appears, the response in the visible increases and the response in the IR decreases, thus $T_g\theta$ increases.

ρ is the module of vector (ρ,θ), and for a given θ, is an increment of x and y, responses of visible and IR, which means an increment of the amount of energy in the four bands or an increment of brightness in the scene. We cannot say, of course, where the boundary between vegetation and soil is, because everything in nature varies gradually. According to the tasselled cup theory the line of soils, that has the associated concept of brightness and shade, is parallel to the diagonal of the four-dimensional space defined by the four bands of Landsat. We can see that this corresponds approximately to the diagonal of the two-dimensional space formerly described. For this reason in the two-dimensional space described above, the major cover types would be distributed as follows:



The bispectral plot of the final classes can be found later.

Reflectance Characteristics of Green Leaves



J. SELECTION OF TEST AREAS AND CLASSIFICATION

The next step in the analysis is the selection of test areas and the performance of a classification of each area into the classes chosen as described above. In order to do this several areas where the ground truth was available were chosen. After this they were classified using a minimum distance to the means classifier.

By this minimum distance criterion, every pixel in the area is assigned to the class to which it is closest. This criterion is different from the maximum likelihood decision rule by which every pixel is classified as the class which it is most likely to belong to (which it has highest probability of belonging to). Both minimum distance and maximum likelihood classifiers were used in this work, but minimum distance was observed to be more desirable for two reasons:

1. - The maximum likelihood classified uses much more CPU time than minimum distance (approximately three times).

2. - The differences between the outputs of the minimum distance and maximum likelihood classifiers are found in only a few pixels.

For these reasons, minimum distance was used in this work to perform the classifications. In the comparison of the classifications with the ground truth it was observed that many of the spectral classes were not representative. In order to obtain a map with homogeneous areas corresponding to the ground truth several spectral classes were represented with same symbol. Those classes were different varieties of the same information class and by joining them with the same symbol the different cover types that form the area can be visualized and identified by means of the ground truth.

VI. RESULTS.

FINAL CLASSES.

Their representation by means of the Bispectral Plot.

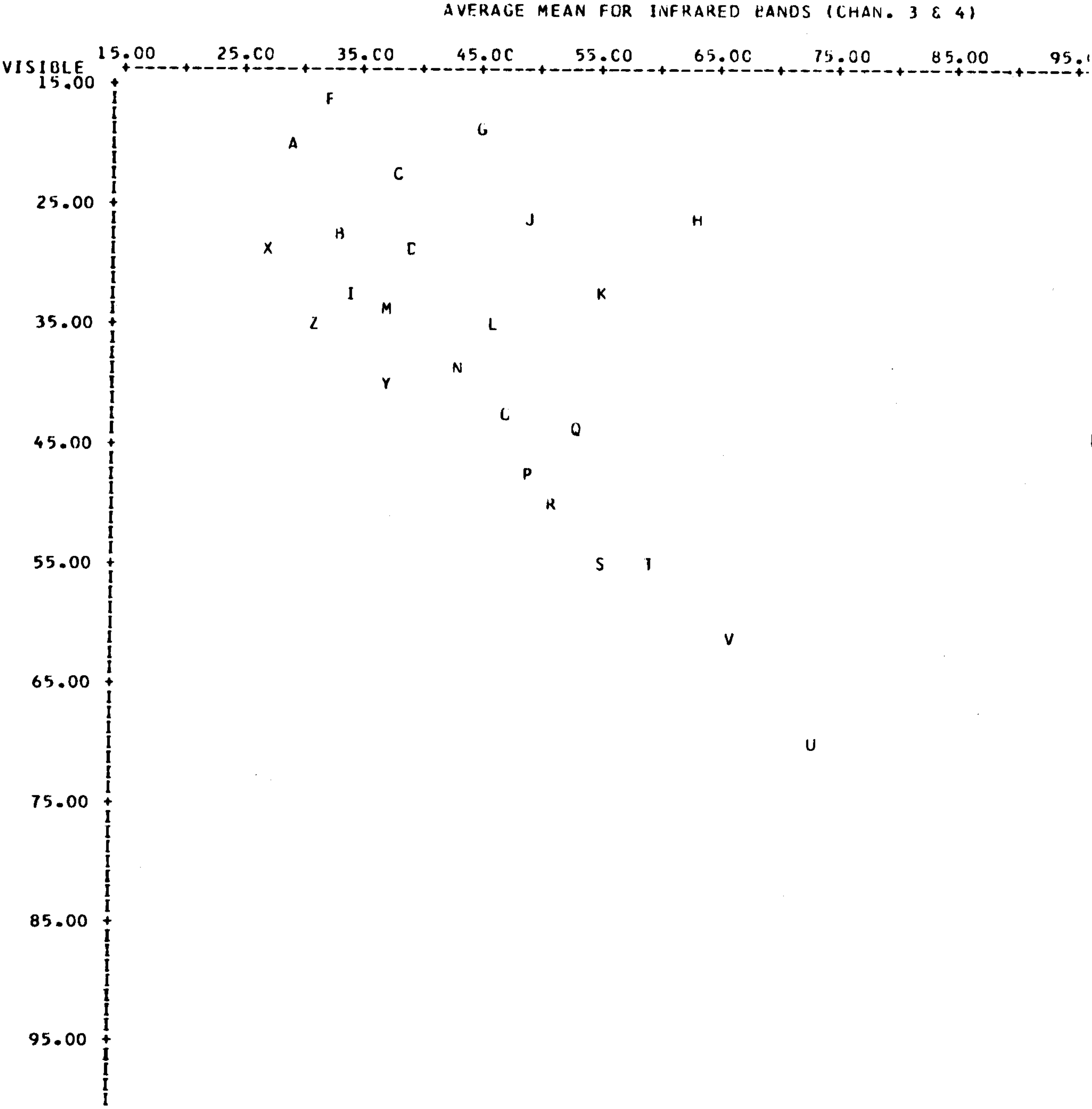
CLASS NUMBER	SYMBOL	AVERAGE VISIBLE	AVERAGE INFRARED	NOTE *	CLASS NAME
1	A	20.0	29.2		A
2	B	27.1	33.2		B
3	C	22.2	37.5		C
4	D	29.0	38.6		D
5	E	14.1	24.7	1	E
6	F	15.7	32.4		F
7	G	18.5	44.9		G
8	H	26.3	62.8		H
9	I	31.9	34.4		I
10	J	26.6	49.2		J
11	K	32.2	55.3		K
12	L	34.6	45.8		L
13	M	33.8	37.1		M
14	N	38.3	43.0		P
15	O	42.4	47.4		V
16	P	47.1	49.0		W
17	Q	43.6	52.9		X
18	R	50.5	51.5		Y
19	S	55.6	55.1		Z
20	T	54.8	59.5		$
21	U	69.4	73.5		+
22	V	61.8	66.1		=
23	W	15.4	10.1	1	/
24	X	28.9	26.9		N
25	Y	40.1	36.7		O
26	Z	34.4	30.6		Q

* NOTE
1 = CLASS WAS OUTSIDE THE AREA DISPLAYED
2 = CLASS WOULD HAVE SUPERIMPOSED ANOTHER

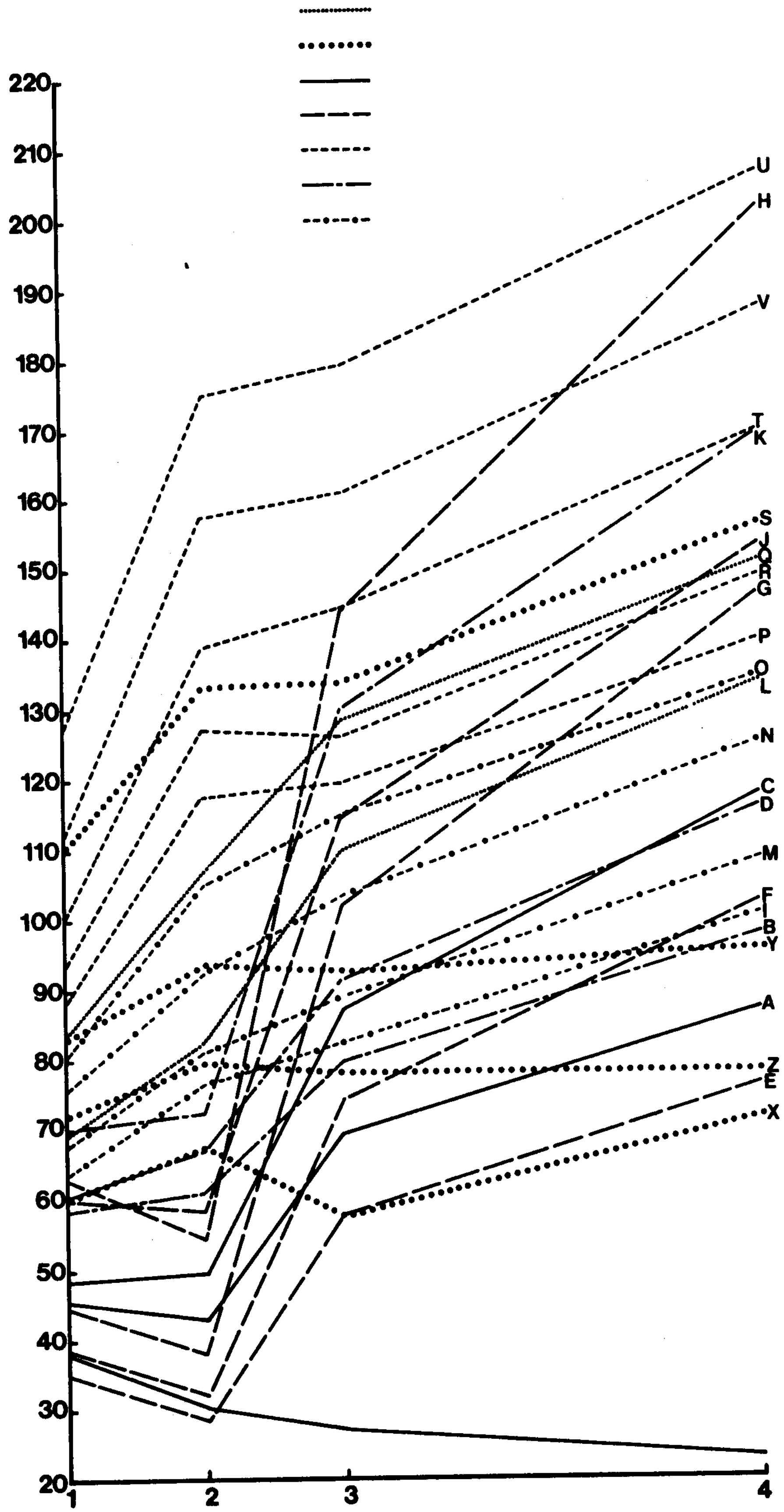
BATCNITE
RAMCN

LABORATORY FOR APPLICATIONS OF REMOTE SENSING
PURDUE UNIVERSITY

COINCIDENT BI-SPECTRAL PLOT (MEAN) FOR CLASS(ES)



SPECTRAL CURVES OF THE FINAL CLASSES.


220
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
1
2
3
4
U
H
V
T
K
S
J
Q
R
G
P
O
L
N
C
D
M
F
I
B
Y
A
Z
E
X

Means and covariance matrices of the Final Classes.

SEPARABILITY STUDY

CLASS....A

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	18.41	21.54	39.58	16.91
STD. DEV.	1.47	2.44	3.03	2.00

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.73	1.00		
0.70-0.80	0.34	0.40	1.00	
0.80-1.10	0.24	0.30	0.81	1.00

A

POOL OF 18(14),19(14),32(16),17(14),18(12),32(15),19(13)

SEPARABILITY STUDY

CLASS....B

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	23.60	30.61	45.21	21.21
STD. DEV.	1.60	2.53	2.44	1.5[illegible]

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.66	1.00		
0.70-0.80	0.10	0.13	1.00	
0.80-1.10	-0.08	-0.09	0.55	1.00

B

POOL OF 8(14),9(12),17(12),18(5),28(12),27(16),32(10)

SEPARABILITY STUDY

CLASS....C

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	19.62	24.87	49.46	25.59
STD. DEV.	1.54	2.32	1.85	1.47

CORRELATION MATRIA

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.68	1.00		
0.70-0.80	0.05	0.12	1.00	
0.80-1.10	-0.10	-0.13	0.44	1.00

C

POOL OF 19(6), 20(4)

SEPARABILITY STUDY

CLASS....D

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	24.42	33.57	52.07	25.15
STD. DEV.	1.24	2.14	1.43	1.49

CORRELATION MATRIA

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.50-0.70	0.48	1.00		
0.70-0.80	0.20	0.17	1.00	
0.80-1.10	0.10	0.04	0.38	1.00

D

11(4)

SEPARABILITY STUDY

CLASS....E

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	14.07	14.19	[illegible]	[illegible]
STD. DEV.	1.22	1.52	[illegible]	1.73

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.54	1.00		
0.70-0.80	0.23	0.14	1.00	
0.80-1.10	0.14	0.10	0.36	1.00

E

20 (12,13)

SEPARABILITY STUDY

CLASS....F

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	15.61	15.88	42.42	22.32
STD. DEV.	1.32	1.87	3.31	2.16

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.67	1.00		
0.70-0.80	0.58	0.58	1.00	
0.80-1.10	0.45	0.39	0.83	1.00

F

POOL OF 20 (11), 18 (11), 19 (12)

SEPARABILITY STUDY

CLASS....G

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	17.85	19.10	58.04	31.75
STD. DEV.	1.43	2.41	2.62	1.96

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.72	1.00		
0.70-0.80	0.22	0.21	1.00	
0.80-1.10	-0.05	-0.10	0.53	1.00

G

POOL OF 11(19), 19(7)

SEPARABILITY STUDY

CLASS....H

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	25.39	27.26	81.83	43.[illegible]
STD. DEV.	2.10	3.13	4.57	2.75

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.[illegible]	1.00		
0.70-0.80	0.[illegible]	0.14	1.00	
0.80-1.10	-0.13	-0.27	[illegible]	1.00

H

24 (14)

SEPARABILITY STUDY

CLASS....I

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	25.63	35.24	46.94	21.60
STD. DEV.	1.30	1.41	1.44	1.18

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.05	1.00		
0.70-0.80	0.18	-0.22	1.00	
0.80-1.10	-0.10	-0.02	0.32	1.00

I

11 (15)

SEPARABILITY STUDY

CLASS....J

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	24.19	29.17	65.01	33.30
STD. DEV.	1.54	3.31	3.46	2.55

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.72	1.00		
0.70-0.80	-0.07	-0.22	1.00	
0.80-1.10	-0.31	-0.43	0.78	1.00

J

11 (13)

SEPARABILITY STUDY

CLASS....K

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	28.25	36.25	73.88	36.74
STD. DEV.	2.20	4.07	3.87	2.30

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.78	1.00		
0.70-0.80	0.30	0.16	1.00	
0.80-1.10	0.14	-0.04	0.73	1.00

K

POOL OF 9(6), 24(15), 27(13)

SEPARABILITY STUDY

CLASS....L

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	27.99	41.26	62.52	24.12
STD. DEV.	1.24	2.32	2.17	1.45

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.53	1.00		
0.70-0.80	0.23	0.23	1.00	
0.80-1.10	-0.00	-0.04	0.47	1.00

L

32(5)

SEPARABILITY STUDY

CLASS....M

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	27.12	40.54	50.54	23.58
STD. DEV.	1.33	1.56	1.09	1.23

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.24	1.00		
0.70-0.80	-0.09	-0.15	1.00	
0.80-1.10	-0.29	0.00	0.23	1.00

M

11(12)

SEPARABILITY STUDY

CLASS....P

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	30.41	45.21	57.53	27.15
STD. DEV.	2.53	2.67	2.07	1.44

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.53	1.00		
0.70-0.80	0.34	0.27	1.00	
0.80-1.10	-0.03	0.13	0.47	1.00

N

11(5,6,8),27(10),3(12),11(4),24(13)

SEPARABILITY STUDY

CLASS....V

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	32.34	52.38	65.45	[illegible]
STD. DEV.	1.70	2.91	2.20	[illegible]

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.41	1.00		
0.70-0.80	-0.17	-0.41	1.00	
0.80-1.10	-0.26	-0.39	0.43	1.00

O

27(8), 11(2), 27(7).

SEPARABILITY STUDY

CLASS....W

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	35.47	56.76	67.76	30.29
STD. DEV.	2.24	1.49	1.64	1.54

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.08	1.00		
0.70-0.80	0.07	[illegible]	1.00	
0.80-1.10	[illegible]	[illegible]	0.34	1.00

P

3(7), 6(9), 12(10), 28(6), 8(8)

SEPARABILITY STUDY

CLASS....X

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	33.73	53.38	73.11	32.73
STD. DEV.	1.76	2.83	2.39	2.09

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.36	1.00		
0.70-0.80	0.13	-0.15	1.00	
0.80-1.10	0.10	-0.34	0.68	1.00

Q

27(5)

SEPARABILITY STUDY

CLASS....Y

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	37.60	63.38	71.70	31.21
STD. DEV.	1.31	1.25	1.29	1.46

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.28	1.00		
0.70-0.80	0.06	0.00	1.00	
0.80-1.10	-0.14	-0.02	0.25	1.00

R

8(6)

SEPARABILITY STUDY

CLASS....Z

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	44.44	66.71	76.16	33.95
STD. DEV.	2.26	2.91	2.36	1.52

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.34	1.00		
0.70-0.80	0.22	0.43	1.00	
0.80-1.10	0.07	0.37	0.57	1.00

S

3(5), 24(5)

SEPARABILITY STUDY

CLASS....$

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	40.29	69.40	82.13	36.85
STD. DEV.	2.00	3.35	2.59	1.68

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.64	1.00		
0.70-0.80	0.47	0.51	1.00	
0.80-1.10	0.25	0.39	0.55	1.00

T

13(6), 6(5), 8(3), 12(5), 13(7), 6(3)

SEPARABILITY STUDY

CLASS....+

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	51.47	67.34	[illegible]	44.[illegible]3
STD. DEV.	2.39	3.56	[illegible]	[illegible]

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.[illegible]0	1.00		
0.70-0.80	0.43	0.59	1.[illegible]	
0.80-1.10	0.40	0.65	[illegible].71	1.00

U

12(1)

SEPARABILITY STUDY

CLASS....=

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	44.96	78.71	91.49	40.62
STD. DEV.	1.44	1.78	1.73	1.34

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.31	1.00		
0.70-0.80	-0.08	-0.20	1.00	
0.80-1.10	-0.16	-0.02	0.31	1.00

V

13(3)

SEPARABILITY STUDY

CLASS..../

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	15.43	15.30	15.21	4.90
STD. DEV.	1.43	3.13	4.39	2.41

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.71	1.00		
0.70-0.80	0.53	0.68	1.00	
0.80-1.10	0.43	0.55	0.91	1.00

W

20(14)

SEPARABILITY STUDY

CLASS....[illegible]

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	24.23	33.64	3[illegible]	15.51
STD. DEV.	2.35	2.42	2.25	1.54

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.80	1.00		
0.70-0.80	0.40	0.27	1.00	
0.80-1.10	0.17	-0.01	0.92	1.00

X

6(14)

SEPARABILITY STUDY

CLASS....0

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	33.40	46.87	52.66	20.80
STD. DEV.	2.56	2.62	2.46	1.69

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.54	1.00		
0.70-0.80	0.16	0.36	1.00	
0.80-1.10	-0.04	0.04	0.52	1.00

Y

40(10)

SEPARABILITY STUDY

CLASS....[illegible]

CHANNEL	1	2	3	4
SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
MEAN	29.08	39.75	44.19	16.96
STD. DEV.	1.86	2.68	3.05	1.94

CORRELATION MATRIX

SPECTRAL BAND	0.50 - 0.60	0.60 - 0.70	0.70 - 0.80	0.80 - 1.10
0.50-0.60	1.00			
0.60-0.70	0.54	1.00		
0.70-0.80	0.29	0.45	1.00	
0.80-1.10	0.12	0.12	0.76	1.00

Z

40(14)

Transformed divergence between pairs of classes.

BATONITE
RAMON

LABORATORY FOR APPLICATIONS OF REMOTE SENSING
PURDUE UNIVERSITY

APR 30,1979
08 32 28 PM
LARSYS VERSION 3

RETENTION LEVEL .. 1 MAXIMUM 30000
MINIMUM 0

DIVERGENCE **WITH** SATURATING TRANSFORM

	CHANNELS	DIJ(MIN)	D(AVE)	AB (10)	AC (10)	AD (10)	AE (10)	AF (10)	AG (10)	AH (10)	AI (10)	AJ (10)	AK (10)
1.	1 2 3 4	1523.	1977.	1771	1907	1996	1844	1764	1999	2000	2000	2000	2000

WEIGHTED INTERCLASS DIVERGENCE (DIJ)

	CHANNELS	AL (10)	AM (10)	AN (10)	AO (10)	AP (10)	AQ (10)	AR (10)	AS (10)	AT (10)	AU (10)	AV (10)	AW (10)	AX (10)	AY (10)	AZ (10)
1.	1 2 3 4	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	1997	1985	2000	2000

WEIGHTED INTERCLASS DIVERGENCE (DIJ)

	CHANNELS	BC (10)	BD (10)	BE (10)	BF (10)	BG (10)	BH (10)	BI (10)	BJ (10)	BK (10)	BL (10)	BM (10)	BN (10)	BO (10)	BP (10)	BQ (10)
1.	1 2 3 4	1719	1585	2000	1999	2000	2000	1916	1999	2000	2000	1997	1999	2000	2000	2000

WEIGHTED INTERCLASS DIVERGENCE (DIJ)

	CHANNELS	BR (10)	BS (10)	BT (10)	BU (10)	BV (10)	BW (10)	BX (10)	BY (10)	BZ (10)	CD (10)	CE (10)	CF (10)	CG (10)	CH (10)	CI (10)
1.	1 2 3 4	2000	2000	2000	2000	2000	2000	1832	1996	1869	1815	2000	1934	1942	2000	2000

WEIGHTED INTERCLASS DIVERGENCE (DIJ)

	CHANNELS	CJ (10)	CK (10)	CL (10)	CM (10)	CN (10)	CO (10)	CP (10)	CQ (10)	CR (10)	CS (10)	CT (10)	CU (10)	CV (10)	CW (10)	CX (10)
1.	1 2 3 4	1958	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000

CONTINUED...
RETENTION LEVEL .. 1 MAXIMUM 30000
MINIMUM 0

DIVERGENCE **WITH** SATURATING TRANSFORM

	CHANNELS	WEIGHTED INTERCLASS DIVERGENCE (DIJ)														
		CY (10)	CZ (10)	DE (10)	DF (10)	DG (10)	DH (10)	DI (10)	DJ (10)	DK (10)	DL (10)	DM (10)	DN (10)	DO (10)	DP (10)	DQ (10)
1.	1 2 3 4	2000	2000	2000	2000	1999	2000	1847	1984	2000	1970	1817	1981	2000	2000	2000
		DR (10)	DS (10)	DT (10)	DU (10)	DV (10)	DW (10)	DX (10)	DY (10)	DZ (10)	EF (10)	EG (10)	EH (10)	EI (10)	EJ (10)	EK (10)
1.	1 2 3 4	2000	2000	2000	2000	2000	2000	1999	1998	1950	1547	2000	2000	2000	2000	2000
		EL (10)	EM (10)	EN (10)	EO (10)	EP (10)	EQ (10)	ER (10)	ES (10)	ET (10)	EU (10)	EV (10)	EW (10)	EX (10)	EY (10)	EZ (10)
1.	1 2 3 4	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	1996	2000	2000	2000
		FG (10)	FH (10)	FI (10)	FJ (10)	FK (10)	FL (10)	FM (10)	FN (10)	FO (10)	FP (10)	FQ (10)	FR (10)	FS (10)	FT (10)	FU (10)
1.	1 2 3 4	1965	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000
		FV (10)	FW (10)	FX (10)	FY (10)	FZ (10)	GH (10)	GI (10)	GJ (10)	GK (10)	GL (10)	GM (10)	GN (10)	GO (10)	GP (10)	GQ (10)
1.	1 2 3 4	2000	2000	2000	2000	2000	2000	2000	1874	1998	2000	2000	2000	2000	2000	2000

CONTINUED...
RETENTION LEVEL .. 1 MAXIMUM 30000
MINIMUM 0

DIVERGENCE **WITH** SATURATING TRANSFORM

	CHANNELS	GR (10)	GS (10)	GT (10)	GU (10)	GV (10)	GW (10)	GX (10)	GY (10)	GZ (10)	HI (10)	HJ (10)	HK (10)	HL (10)	HM (10)	HN (10)
1.	1 2 3 4	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	1876	1679	2000	2000	2000

WEIGHTED INTERCLASS DIVERGENCE (DIJ)

	CHANNELS	HO (10)	HP (10)	HQ (10)	HR (10)	HS (10)	HT (10)	HU (10)	HV (10)	HW (10)	HX (10)	HY (10)	HZ (10)	IJ (10)	IK (10)	IL (10)
1.	1 2 3 4	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000

WEIGHTED INTERCLASS DIVERGENCE (DIJ)

	CHANNELS	IM (10)	IN (10)	IO (10)	IP (10)	IQ (10)	IR (10)	IS (10)	IT (10)	IU (10)	IV (10)	IW (10)	IX (10)	IY (10)	IZ (10)	JK (10)
1.	1 2 3 4	1685	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	1998	1999	1868	1523

WEIGHTED INTERCLASS DIVERGENCE (DIJ)

	CHANNELS	JL (10)	JM (10)	JN (10)	JO (10)	JP (10)	JQ (10)	JR (10)	JS (10)	JT (10)	JU (10)	JV (10)	JW (10)	JX (10)	JY (10)	JZ (10)
1.	1 2 3 4	1928	2000	1998	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000

WEIGHTED INTERCLASS DIVERGENCE (DIJ)

	CHANNELS	KL (10)	KM (10)	KN (10)	KO (10)	KP (10)	KQ (10)	KR (10)	KS (10)	KT (10)	KU (10)	KV (10)	KW (10)	KX (10)	KY (10)	KZ (10)
1.	1 2 3 4	1965	2000	1999	1975	2000	1958	2000	2000	2000	2000	2000	2000	2000	2000	2000

BATONITE
RAMON

LABORATORY FOR APPLICATIONS OF REMOTE SENSING
PURDUE UNIVERSITY

APR 30,1979
08 32 28 PM
LARSYS VERSION 3

CONTINUED...
RETENTION LEVEL .. 1 MAXIMUM 30000
MINIMUM 0

DIVERGENCE **WITH** SATURATING TRANSFORM

WEIGHTED INTERCLASS DIVERGENCE (DIJ)

	CHANNELS	LM (10)	LN (10)	LO (10)	LP (10)	LQ (10)	LR (10)	LS (10)	LT (10)	LU (10)	LV (10)	LW (10)	LX (10)	LY (10)	LZ (10)	MN (10)
1.	1 2 3 4	2000	1550	1927	2000	1994	2000	2000	2000	2000	2000	2000	2000	1997	2000	1998

WEIGHTED INTERCLASS DIVERGENCE (DIJ)

	CHANNELS	MO (10)	MP (10)	MQ (10)	MR (10)	MS (10)	MT (10)	MU (10)	MV (10)	MW (10)	MX (10)	MY (10)	MZ (10)	NO (10)	NP (10)	NQ (10)
1.	1 2 3 4	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	1961	1985	1837	1998	1997

WEIGHTED INTERCLASS DIVERGENCE (DIJ)

	CHANNELS	NR (10)	NS (10)	NT (10)	NU (10)	NV (10)	NW (10)	NX (10)	NY (10)	NZ (10)	OP (10)	OQ (10)	OR (10)	OS (10)	OT (10)	OU (10)
1.	1 2 3 4	2000	2000	2000	2000	2000	2000	2000	1830	1997	1674	1704	2000	2000	2000	2000

WEIGHTED INTERCLASS DIVERGENCE (DIJ)

	CHANNELS	OV (10)	OW (10)	OX (10)	OY (10)	OZ (10)	PQ (10)	PR (10)	PS (10)	PT (10)	PU (10)	PV (10)	PW (10)	PX (10)	PY (10)	PZ (10)
1.	1 2 3 4	2000	2000	2000	1998	2000	1715	1809	1989	2000	2000	2000	2000	2000	2000	2000

WEIGHTED INTERCLASS DIVERGENCE (DIJ)

	CHANNELS	QR (10)	QS (10)	QT (10)	QU (10)	QV (10)	QW (10)	QX (10)	QY (10)	QZ (10)	RS (10)	RT (10)	RU (10)	RV (10)	RW (10)	RX (10)
1.	1 2 3 4	1991	1988	1988	2000	2000	2000	2000	2000	2000	1958	1999	2000	2000	2000	2000

CONTINUED...
RETENTION LEVEL .. 1 MAXIMUM 30000
MINIMUM C

DIVERGENCE **WITH** SATURATING TRANSFORM

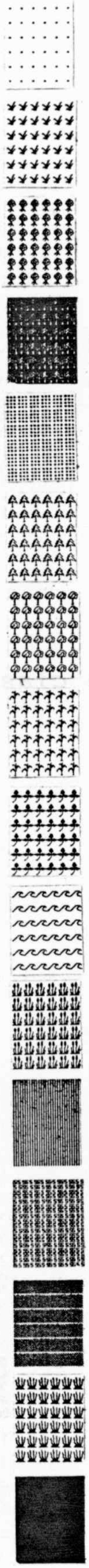
	CHANNELS	RY (1C)	RZ (10)	ST (10)	SU (10)	SV (10)	SW (10)	SX (10)	SY (10)	SZ (1C)	TU (10)	TV (10)	TW (10)	TX (10)	TY (10)	TZ (10)
1.	1 2 3 4	2000	2000	1753	2000	2000	2000	2000	2000	2000	1994	1994	2000	2000	2000	2000

WEIGHTED INTERCLASS DIVERGENCE (DIJ)

	CHANNELS	UV (1C)	UW (10)	UX (1C)	UY (10)	UZ (10)	VW (10)	VX (10)	VY (10)	VZ (1C)	WX (10)	WY (1C)	WZ (10)	XY (1C)	XZ (10)	YZ (1C)
1.	1 2 3 4	1998	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	2000	1844	1627

WEIGHTED INTERCLASS DIVERGENCE (DIJ)

Maps of several areas obtained with minimumdistance classifier as well as photographs and color composites of them.



CROPFIELDS 1

SOIL

FOREST 1

CROPFIELDS 3

FOREST 2

IRRIGATED LAND

CROPFIELDS 2

WATER

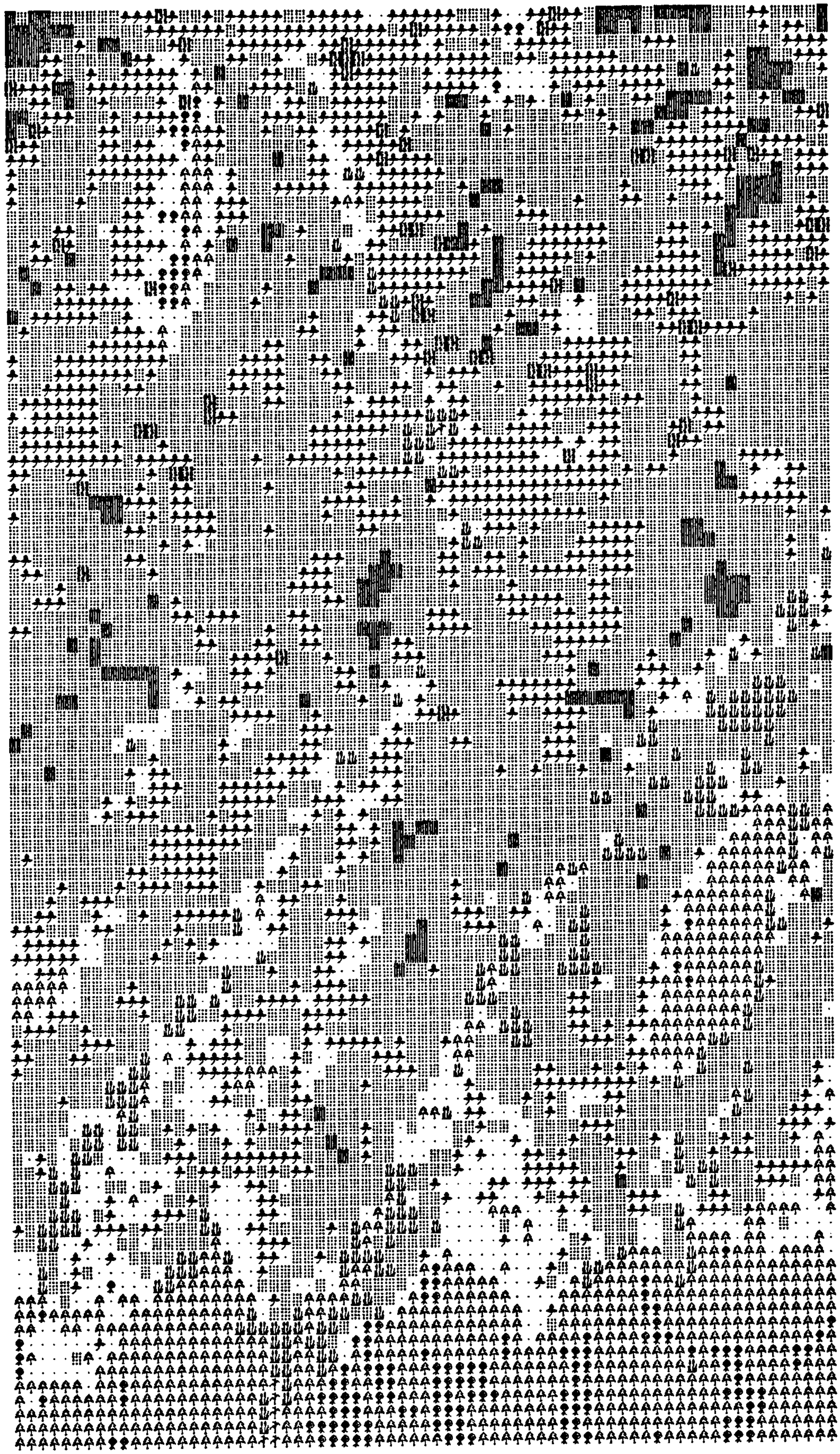
VEGETATION SURROUNDING
A RIVER

SOILS/CROPS

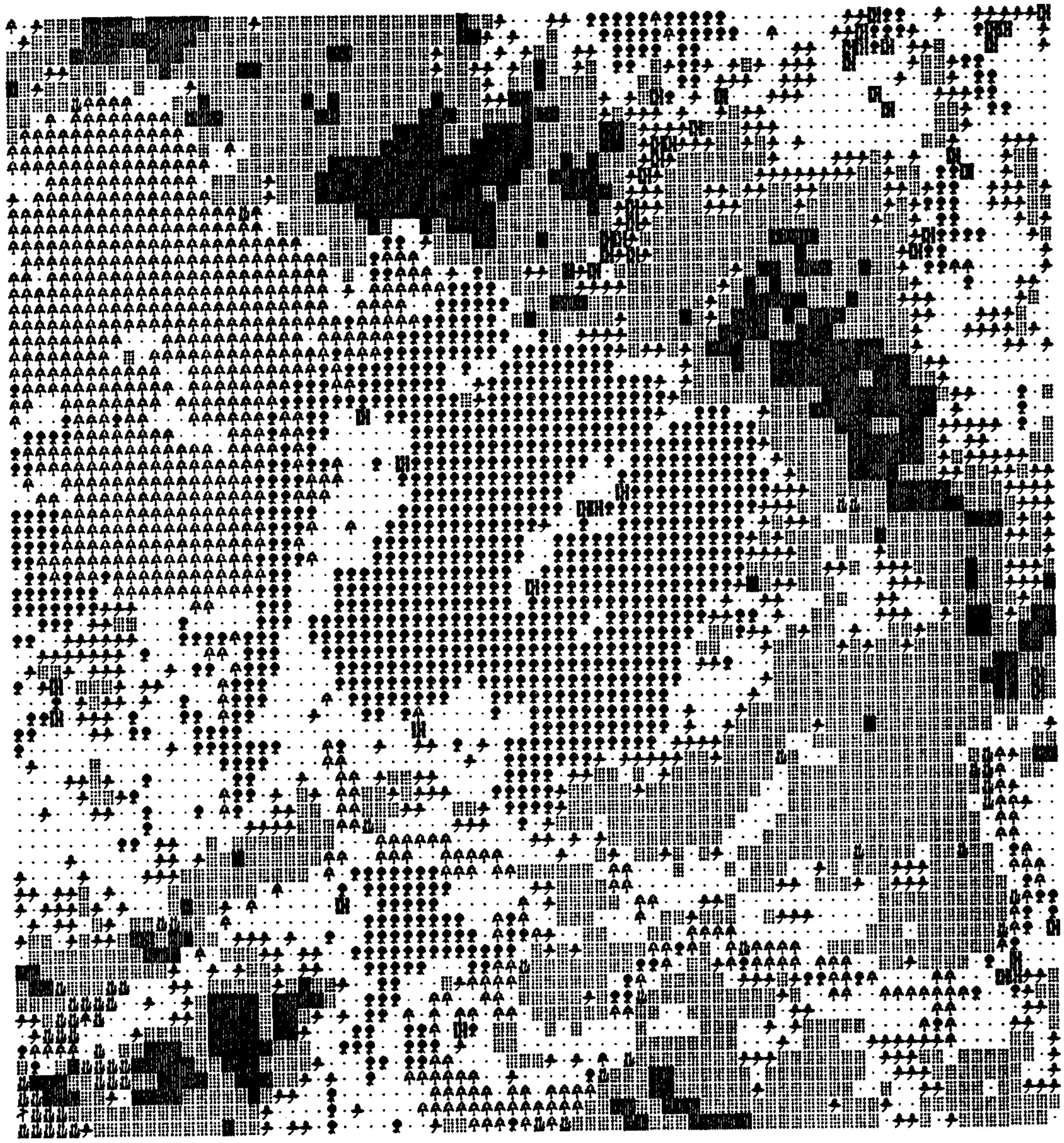
URBAN

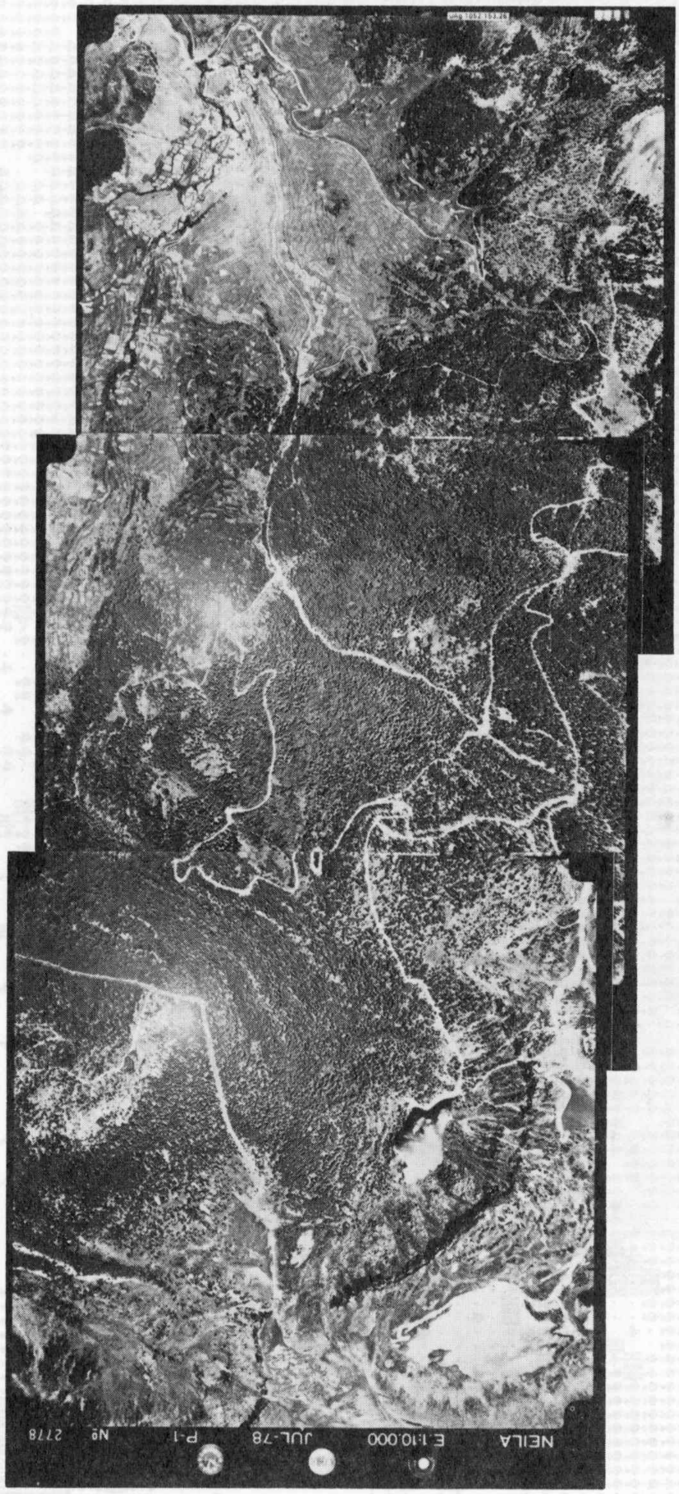

CASTILDELGADO E. 1:10.000 JUL-78 P-1 Nº 2750

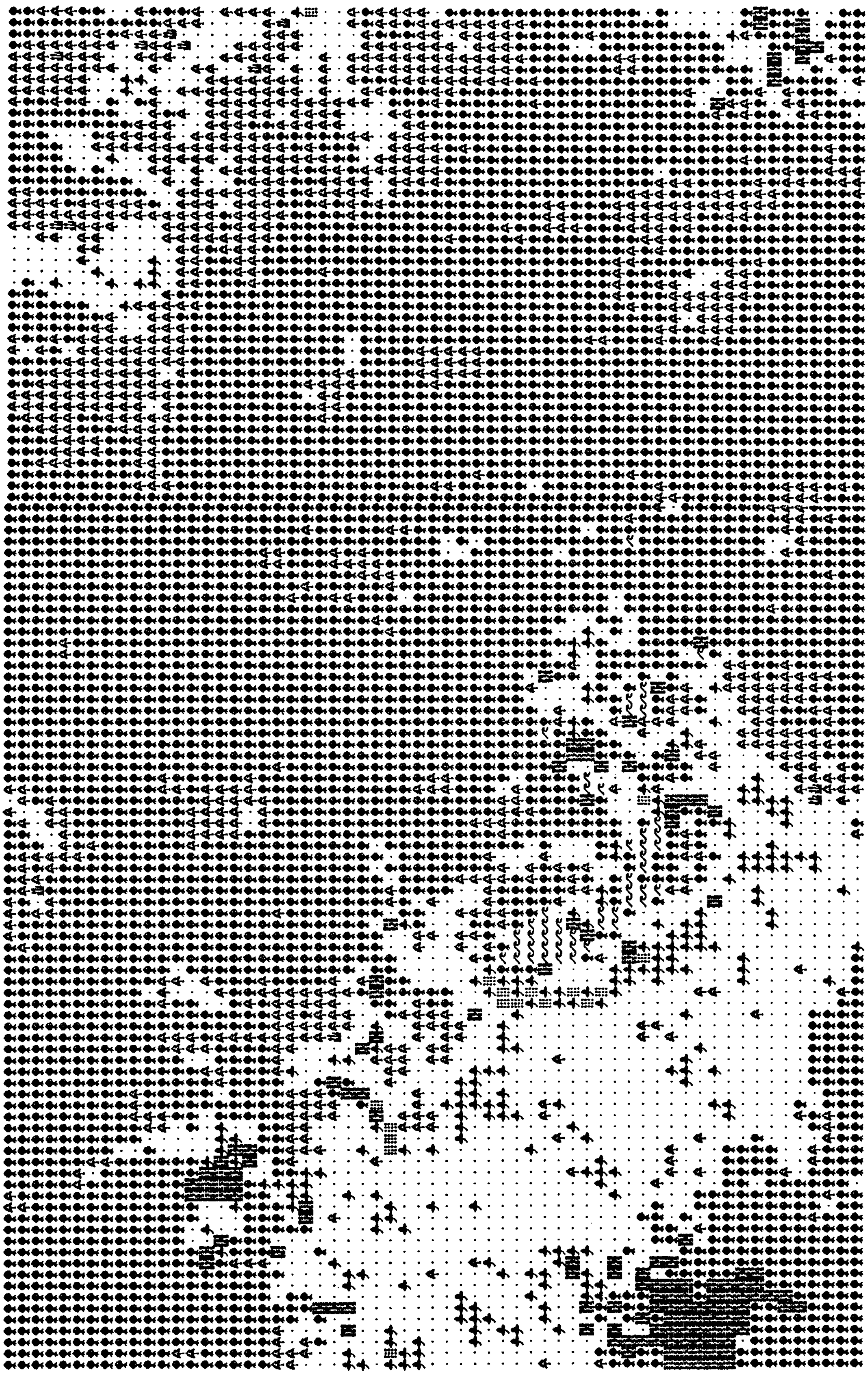

VASCUNANA
E 1:10.000
JUL-78
P-1
Nº

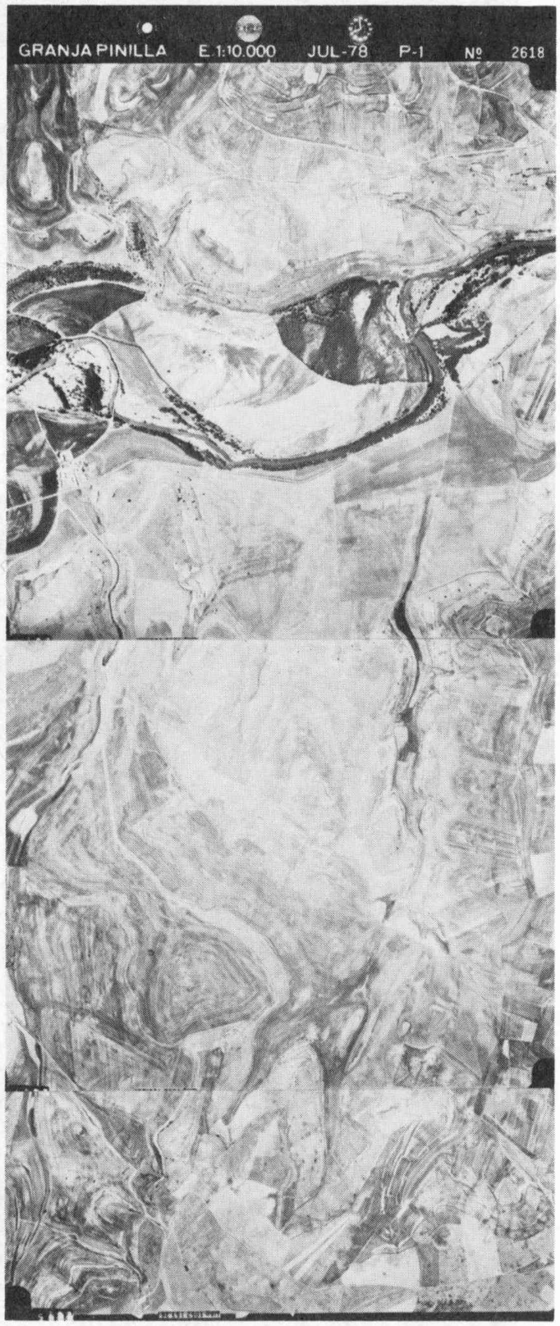



CAVARRUBIAS
E.1:10.000
JUL-78
P-1
Nº 2637



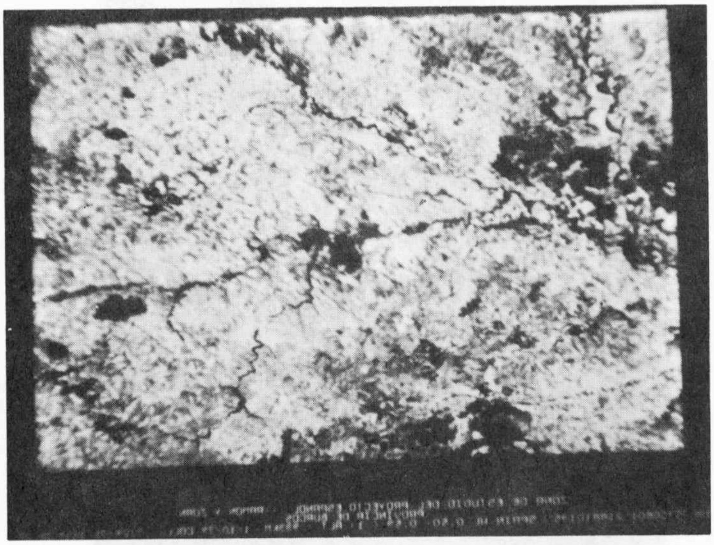

NEILA
E.1:10.000
JUL-78
P-1
Nº 2778

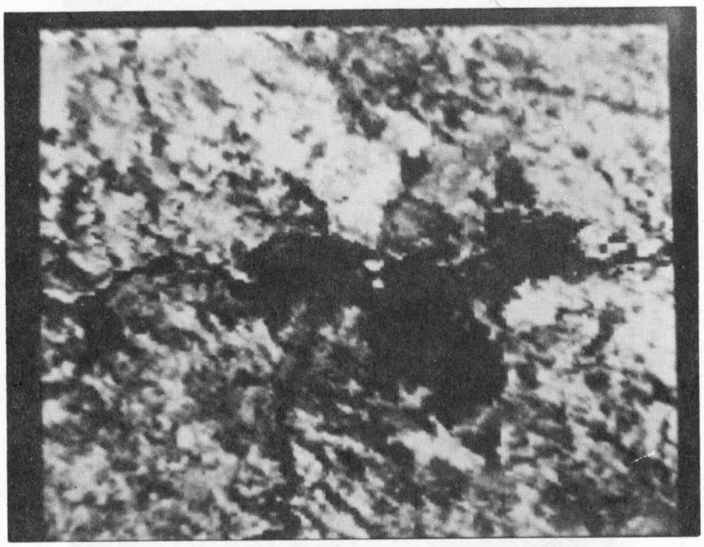


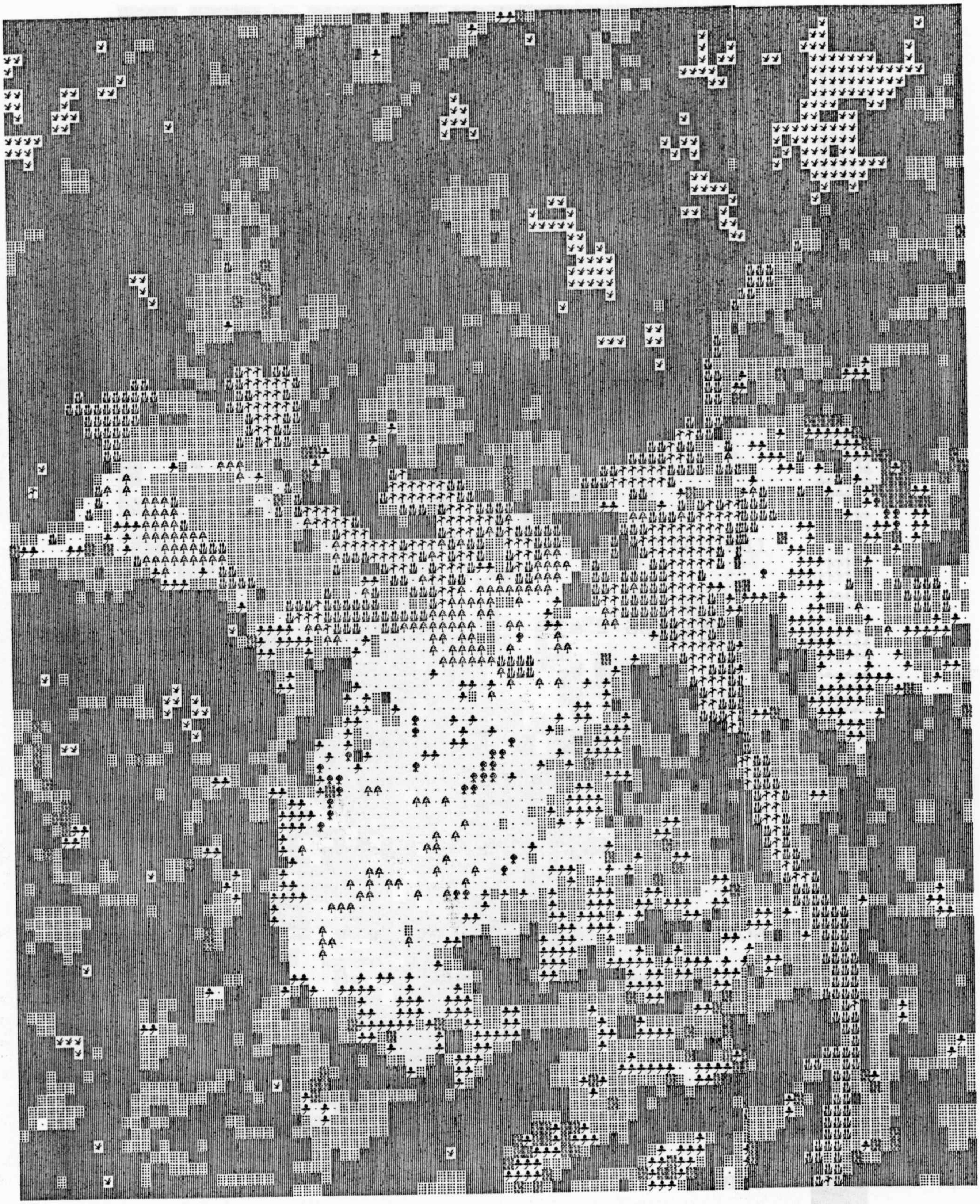

GRANJA PINILLA
E.1:10.000
JUL-78
P-1
Nº 2618


EL RETORTILLO E 1:10.000 JUL-78 P-2 Nº 2614


EL RETORTILLO E:1:10.000 JUL-78 F-1 Nº 2609







CROP FIELDS 1

SOIL

FOREST 1

CROPFIELDS 3

FOREST 2

IRRIGATED LAND

CROPFIELDS 2

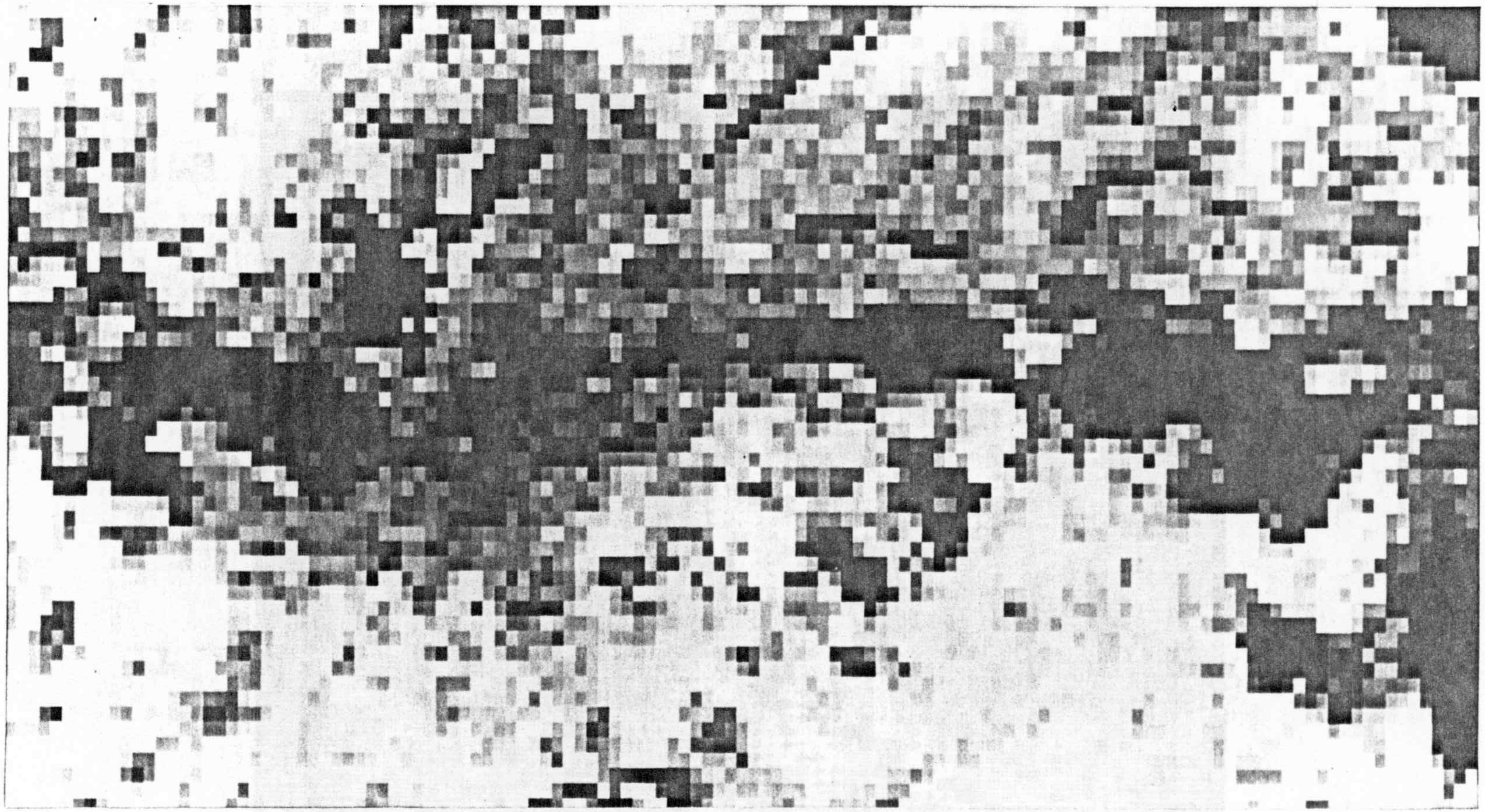
VEGETATION SURROUNDING
A RIVER

URBAN

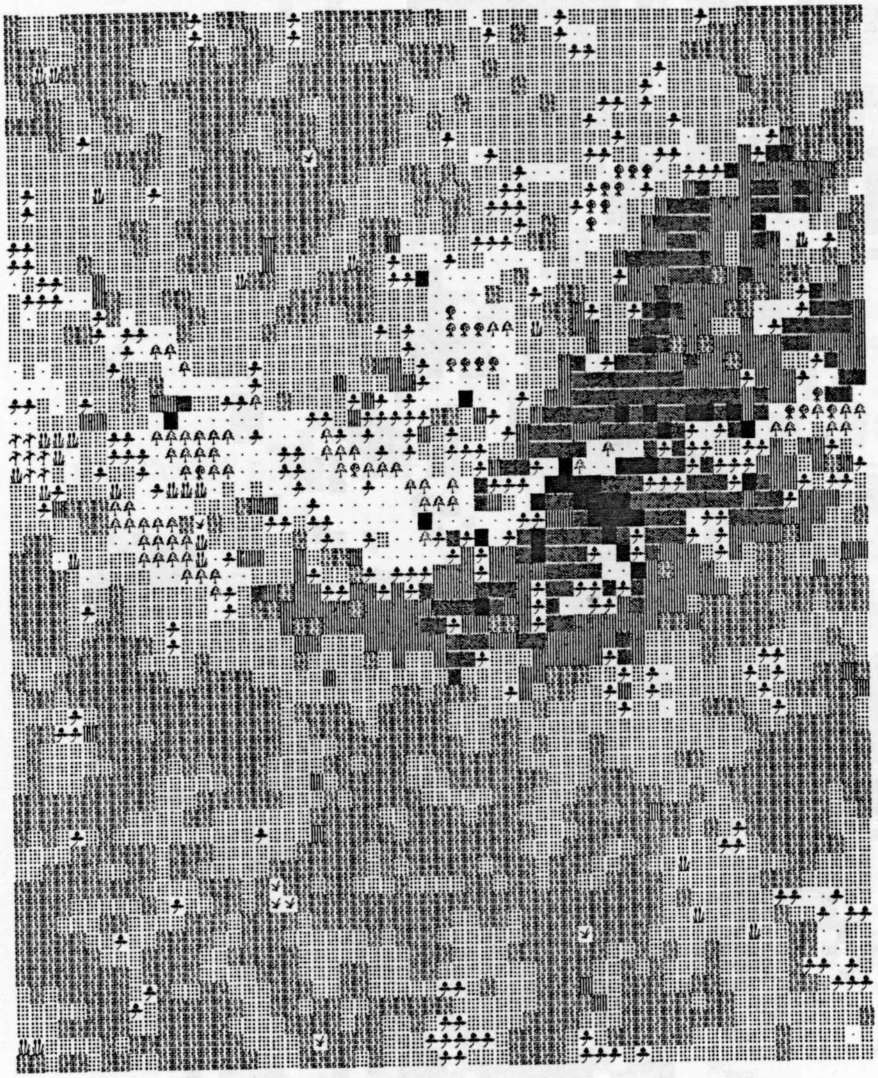
SOIL/CROP

URBAN

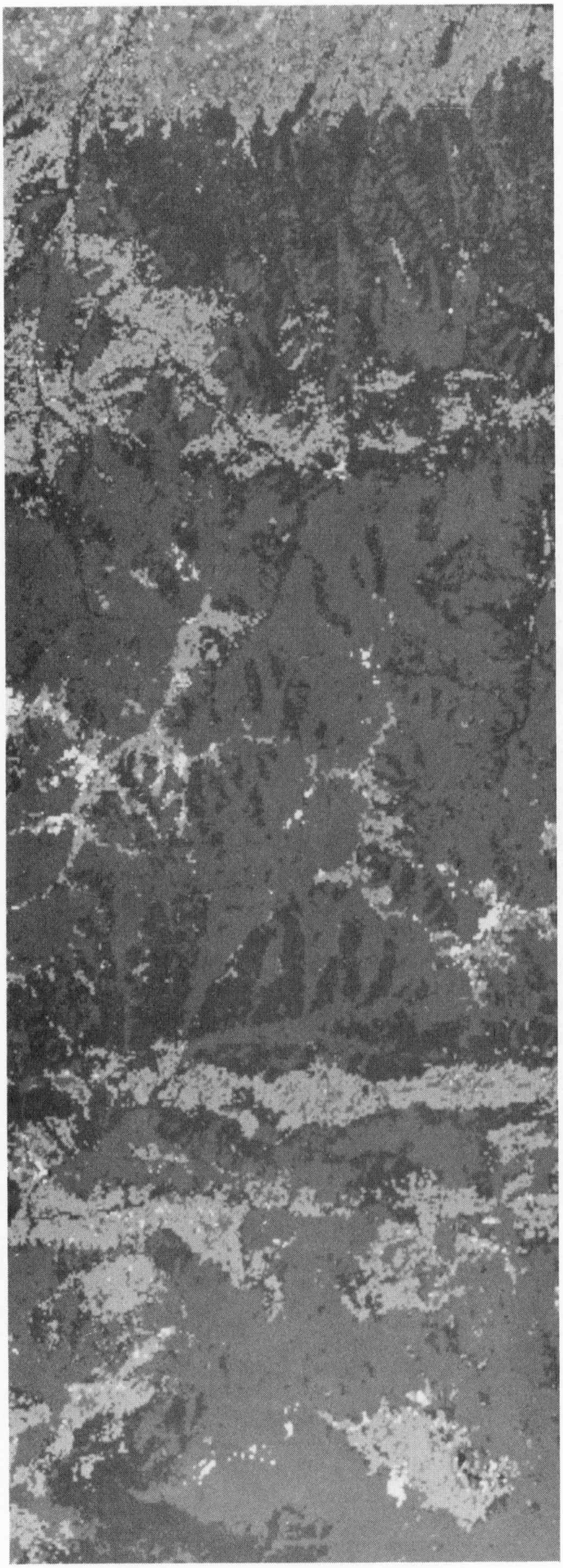
URBAN



Color Composite of the City of Burgos.



Classification of a forest area.



VII. CONCLUSIONS AND RECOMMENDATIONS

1. - Although photointerpretation of the Landsat imagery can provide adequate information for mapping broad categories of land cover, computer aided analysis techniques of the Landsat data in digital format give more information at a higher level of mapping detail with mapping units of approximately half a hectare.

2. - The selection of adequate training areas, and the number of spectral classes for every area (training the classifier) is absolutely necessary to perform a good classification.

3. - The calibration of the data, and further on the interpretation of the spectral curves, gives the analyst a very important tool to control this unsupervised classification at every step of the process and facilitate the decision making process.

4. - Good ground truth, gathered at the same time of the satellite pass, is always necessary in order to identify the different classes and especially to evaluate the classification.

RECOMMENDATIONS

1. - Since analysis means extraction of information from a given data set and in this case the data are spectral responses of ground cover types, it is advisable to think in terms of physical behavior in order to know what is done in every analysis step and not to fall in just a play of numbers with the computer. The computer is an indispensable tool when dealing with large amounts of data, but no more than a tool. The man is the one who makes decisions.

Pooling two classes or deleting one without knowing the physical meaning of what we are doing could lead to catastrophic results in the analysis.

For these reasons, the use of spectral families (calibrated curves)is advisable in order to have a good criteria for pooling or deleting. Think in two hypothetical responses completely different in shape, but with a small transformed divergency. If we think only in numbers we would pool those classes or delete one of them and we would commit a mistake. According to this pooling or deleting should only be allowed among members of a family, because doing it in that way, we are not only taking into account the measurement of separability, but also the physical interrelation of the data and its spectral response.

2. - Landsat spatial resolution is only half a hectare, and for this reason has its limitations as well as its advantages. Landsat resolution is good for forestry but causes problems in agricultural uses when dealing with crop fields of about the same size as the land cover resolution.

We can't try to get from Landsat more information than what is really contained in the data. The Landsat information should not be either overestimated or underestimated.

3. - Before an analysis is undertaken planning of the work should be done.

It is convenient to know beforehand the objectives in order to gather useful reference data and to select the method to extract the information in both the most economic and effective way.

VIII. REFERENCES

1. - LARSYS User's Manual Volumes 1, 2, and 3.

2. - "Remote Sensing - The Quantitative Approach", Philip H. Swain and Shirley M. Davis.

3. - Landsat Data Users Handbook.