Information Note 110471

Remote Sensing Analysis:

A Basic Preparation

by
John Lindenlaub

The Laboratory for Applications of Remote Sensing

Purdue University
West Lafayette, Indiana

Information Note 110471

REMOTE SENSING ANALYSIS: A BASIC PREPARATION

by

John Lindenlaub

Second Edition, Revised
October, 1972

The Laboratory for Applications of Remote Sensing
Purdue University
West Lafayette, Indiana

The Laboratory for Applications of Remote Sensing is supported by the National Aeronautics and Space Administration (NASA) under grant NGL 15-005-112.

Preface

This volume was prepared as an introduction to remote sensing stressing the importance of pattern recognition in numerically oriented remote sensing systems. Its specific purpose is to provide a common background and orientation for those of you who expect to make use of the LARSYS data analysis computer software system. It is intended that this booklet be completed prior to your initial meeting with LARS personnel, although we anticipate that the materials will most likely be read "the night before" or "on the airplane."

For newcomers to remote sensing, this manual will introduce concepts and terminology which you will need later on; remote sensing veterans will be introduced to the numerical analysis and automatic data processing aspects of remote sensing. The material presented here has been drawn from two previous LARS Information Notes: 041571, "Systems Approach to the Use of Remote Sensing," by D. Landgrebe and 101866, "On Pattern Recognition," by G. Cardillo and D. Landgrebe.

The format of this manual resembles that of a programmed text. It is not intended that you read the material from cover to cover. You will find spread throughout the manual a series of pretest items and "steering" instructions. Do your best to answer these questions as you come to them. Depending upon your response, you will be directed to various sections of the manual. In that way the material can be tailored to your individual requirements. At the end of the material, you will find a post test. The questions are based on items important to persons preparing themselves to use LARSYS computer programs. If your answers to these questions are incorrect or you feel your response is too weak, review the pages indicated.*

**Test items and steering instructions are typed in a distinct format and use italic characters to offset them from the body of the manual.*

Pretest items

1. In the space below define remote sensing and give an example of a remote sensing system or an objective of a remote sensing system.

2. The feasibility of remote sensing as we know it today is dependent upon certain types of measurable variations in electromagnetic fields. These types of variations are ________________, _____________ and __________________.

Are you completely satisfied with your responses to the questions above?

If "yes," turn to page 3.

If "no," turn to page 5.

Judging by the confidence with which you answered the previous two pretest questions you must have some experience in remote sensing.

If you would like another viewpoint of what remote sensing is and how information is conveyed *turn to page 5.*

If you would like to go on, turn to page 6.

If you follow the instructions there is no way to get to this page. Take one step backwards and return to the page you just came from.

WHAT IS REMOTE SENSING? HOW IS INFORMATION CONVEYED?

Imagine that you are high above the surface of the earth looking down on it and that you want to survey the earth's surface in order to learn about its resources and thus to manage them better. How could this information be derived? What must the system to extract it look like?

The field of remote sensing provides some of the answers. Remote sensing is the science and art of acquiring information about material objects from measurements made at a distance, without coming into physical contact with the objects. In remote sensing, information may be transmitted to the observer either through force fields or electromagnetic fields, in particular, through the

- Spectral,
- Spatial, and
- Temporal

variations of these fields. Therefore, in order to derive information from these field variations, one must be able to

- Measure the variations and
- Relate these measurements to those of known objects or materials.

If, for example, one desires a map showing all of the water bodies of a certain region of the earth, it is clear that one cannot sense the water directly from spacecraft altitudes, rather only the manifestations of these water bodies which exist at that height. These manifestations, in the form of electromagnetic radiation, must therefore be measured and the measurements analyzed to determine which points on the earth contain water and which do not.

Of the two types of fields mentioned above, electromagnetic fields provide perhaps the greatest potential. The remainder of these remarks will be confined to fields of this type. Figure 1 provides a review of the spectrum of the electromagnetic fields.*

The visible portion, extending from 0.4 to 0.7 micrometers, is the most familiar to us as it is this portion of the spectrum to which our eyes are sensitive; however, sensors can be built to cover a much broader range of wavelengths. The entire portion from 0.3 to 15 micrometers, referred to as the optical wavelength portion, is particularly of interest. The wavelengths shorter than 0.4 micrometers are in the ultraviolet region. The portion above the visible spectrum is the infrared region, with 0.7 to approximately 3 micrometers referred to as the reflective infrared and the region from 3 to 15 micrometers called the emissive or thermal infrared region. In this latter portion of the spectrum, energy is emitted from the body as a result of its thermal activity or heat rather than being reflected from it.

---

* Figures appear at the end of this booklet.

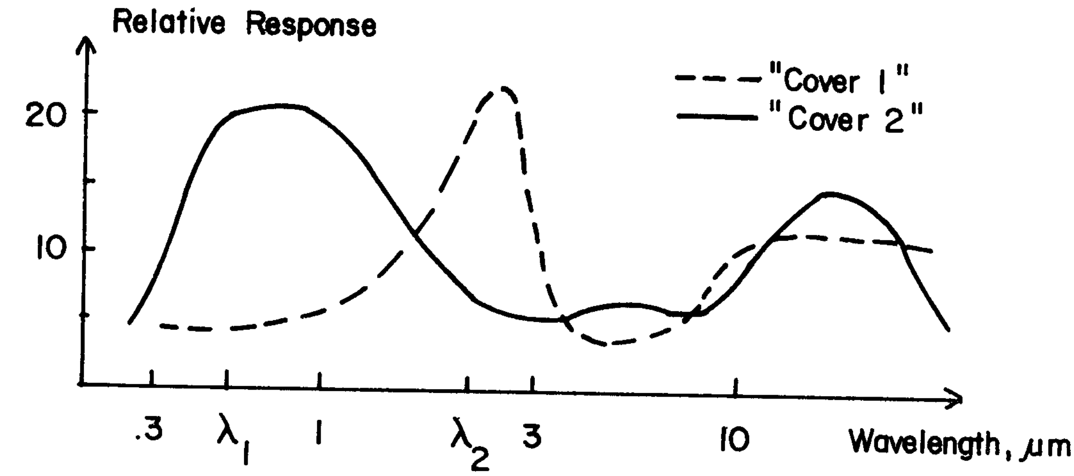
Turn to page 6

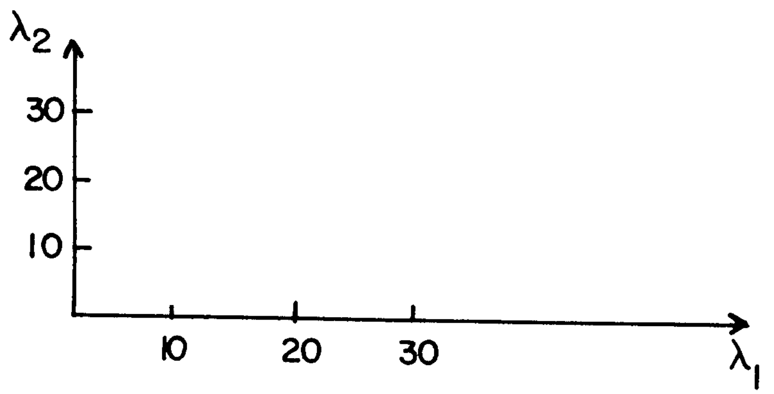
At this point we want to help you determine the level of your background in pattern recognition.

If you have had no exposure to pattern recognition please turn to page 10.

If you have some background in pattern recognition try the following question:

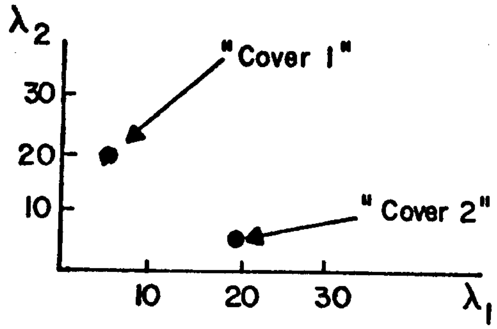
Let the relative response at wavelength λ_1 *and* λ_2 *be the features used in a remote sensing system. Locate the points in the* λ_1, λ_2 *feature plane that represent "cover 1" and "cover 2."*





Turn to page 7.

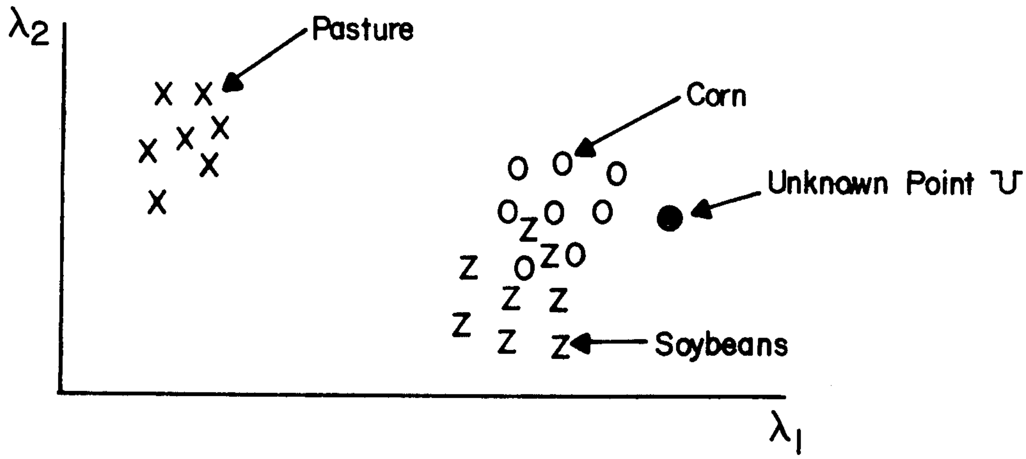
The answer to the question on page 6 is



If you did not obtain the correct answer go on to page 10.

If you did get the right answer try the following problem:

Below are shown the training samples associated with 3 different classes of ground cover as they appear in the λ_1, λ_2 plane. U *is an unknown data point. Using the classification rule: "Assign any data point* U *to that class for which the distance between* U *and the mean (center of gravity) of the training samples is minimum," classify the unknown point* U.



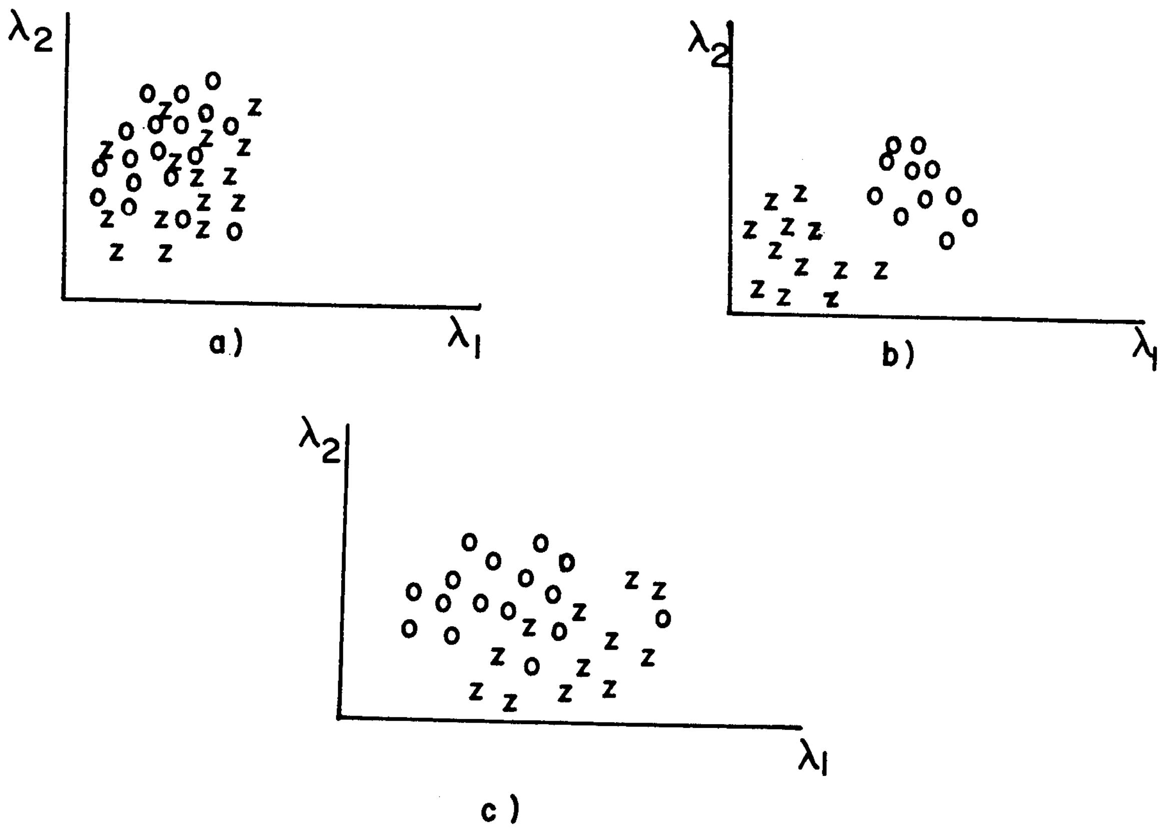
Turn to page 8.

The correct classification is "Corn."

If you did not classify U *as "Corn"* *go on to page 10.*

If you said "Corn" consider the following:

The diagrams below show several sets of training samples similar to those encountered in the 1971 Corn Blight Watch Experiment. Rate each situation as to whether the classes are distinctly separable, marginally separable or not separable.



a) b) c)

Go on to page 9.

The correct answers are

a) not separable

b) distinctly separable

c) marginally separable.

If you answered correctly, please turn to page 35.

If you were not able to answer correctly, please turn to page 10.

To arrive at this point you either indicated that you had no background in pattern recognition or you were not able to answer one or more of the pattern recognition pretest questions correctly.

This branch of the instructional materials presents a condensed introduction to remote sensing followed by a more detailed discussion of pattern recognition.

Please turn to page 11.

Figure 2 is a diagram of the organization of an earth survey system. It is necessary, of course, to have a sensor system viewing that portion of the earth under consideration. There will necessarily be a certain amount of on-board data processing. This will perhaps include the merging of data from other sources such as sensor calibration and data about where the sensor was pointed.

One must next transport the data back to earth for further analysis and processing. This may be done through a telemetry system, as is the case for the Earth Resource Technology Satellite (ERTS), or through direct package return, as will be used with SKYLAB. There usually then is a need for certain preprocessing of the data before the final processing with one or more of the data reduction algorithms. It is at this point in the system, when the data is reduced to information, that it is usually helpful to merge ancillary information, perhaps derived from sources on the surface of the earth, with the remotely sensed information.

Before leaving the matter of the organization of an information system, the necessity of having a thorough understanding of the portion of the system preceding the sensor must be pointed out. Consider Figure 3. This figure shows a simplified version of the energy exchange in a natural environment. It is possible, of course, to detect the presence of vegetation on the earth's surface by measuring the reflected and emitted radiation emanating from the vegetation. One must understand, however, that there are many experimental variables active. It is possible to deal with these experimental variables in several ways. We shall touch briefly on this point later in the discussion.

Summarizing, then, it is possible to derive information about the earth's surface and the condition of its resources by measuring the spectral, spatial, and temporal variations of the electromagnetic fields emanating from points of interest and then analyzing these measurements to relate them to specific classes of materials. To do so, however, requires an adequate understanding of the materials to be sensed and, in order to make the information useful, a precise knowledge is required about how the information will be used and by whom.

Please turn to page 13.

This page left blank intentionally.

Go on to page 13.

THE DUALITY OF SYSTEM TYPES

When we consider the state-of-the-art of remote sensing today, a duality of system types becomes readily apparent. Development in the field has had two major stems because it originated from two somewhat different types of technology. These two types of systems will be referred to here as those with

- Image orientation, and
- Numerical orientation.

An example of an image-oriented system might be simply an aerial camera and a photointerpreter. The photographic film is used to measure the spatial variations of the electromagnetic fields, and the photointerpreter relates these variations to specific classes of surface cover. Numerically oriented systems, on the other hand, tend to involve computers for data analysis. Although the photointerpreter and the computer, respectively, tend to be typical in the two types of systems, it would be an oversimplification and indeed incorrect to say that they are uniquely related to these system types. This becomes clearer upon further examination.

Figure 4 compares the organization of the two system types. Both types of systems need a sensor and some preprocessing; however, the distinction between the types can perhaps be brought out most clearly by noting the location of the *form image* block in the two diagrams. In the image-oriented type, it is in line with the data stream and must precede the analysis block. Numerically oriented systems, on the other hand, need not necessarily contain a *form image* block. If they do, and in earth resources studies they usually do, it may be at the side of the main data stream, as shown. It may thus be used to monitor the operation of the system and perhaps to do some special purpose analysis as needed. An image is, of course, the most efficient way to convey a large amount of information to a human operator. As seen, this is its principal use in both types of systems, but the use is different in the two cases.

The technology for image-oriented systems is relatively well-developed. Sensors best suited to this type of system have been in use for some time, as have appropriate analysis techniques. This type of system also has the advantage of being easily understandable to the layman or neophyte to remote sensing, an advantage important in the earth resources field since many new data users are anticipated. Similarly, it is well-suited for producing subjective information and is especially suited to circumstances where the classes into which the data are to be analyzed cannot be precisely decided upon beforehand. Thus, man with his superior intelligence is or can be deeply involved in the analysis activity. Pictorially oriented systems also have the possibility of being relatively simple and low-cost. On the other hand, it is difficult to use them for large-scale surveys over very large areas involving very large amounts of data.

In the case of numerically oriented systems, the technology is much newer and not nearly so well-developed, though very rapid progress is being made. Because the various steps involved tend to be more abstract, they tend to be less readily understandable by the layman. This type of system is best suited for producing objective information, and large-scale surveys covering large areas are certainly possible. Numerically oriented systems tend to be generally more complex, however.

In summary to this point, the state-of-the art is such that there are two general types of systems; this duality exists primarily for historical reasons and because of the different origins from which technology development began. One type is based on imagery, and therefore a key goal of an intermediate portion of the system is the generation of high quality imagery. In the other case imagery is less important and indeed may not be necessary at all. It is not appropriate to view these two types of systems as competitive with one another since they have different capabilities and each has advantages under certain circumstances. As a matter of fact, these two stems of technology are approaching one another so that the differences between them are becoming less distinct.

We will proceed now to further consideration of numerically oriented systems since this type may be less familiar. In particular we shall examine a type of data analysis useful in this case.

Please turn to page 15.

MULTISPECTRAL ANALYSIS AND PATTERN RECOGNITION

In recent years considerable effort has been devoted to what is referred to as the multispectral approach for data analysis. An initial understanding of what is meant by the term "multispectral approach" may be obtained by considering Figure 5. Shown here in the upper left of the figure is a reproduction of a conventional color photograph of a set of color cards. The remainder of the figure shows photographs of the same color cards taken with black and white film and several different filters. The pass band of each filter is indicated beneath the particular color card set. For example, in the .62-.68 micrometer band, which is in the red portion of the visible spectrum, the red cards appear white in the black and white photo, indicating a high response or a large amount of red light energy being reflected from these cards. In essence the multispectral approach amounts to identifying any color by noting the set of gray scale values produced on the black and white photographs for that particular color rectangle.

As a very simple example of the approach, Figure 6 shows images of an agricultural scene taken in three different portions of the spectrum. Note that in the three bands alfalfa has responses which are dark, light, and dark, respectively, whereas bare soil is gray, dark, and white. Thus, alfalfa can be discriminated from bare soil by identifying the fields which are dark, light, dark in order in these three spectral bands.

One may initially think of the multispectral approach as one in which a very quantitative measure of the color of a material is used to identify it. Color, however, is a term usually related to the response of the human eye; the terminology of spectroscopy, which is more precise, is more useful in understanding the multispectral approach and is applicable beyond the visible region of the spectrum.

In order to understand this approach and to see how a numerically oriented system may be based upon it, consider Figure 7. Shown at the left is a graph of relative response (reflectance) as a function of wavelength for three types of earth-surface cover material: vegetation, soil, and water. Let two wavelengths marked $\lambda 1$ and $\lambda 2$ be selected. Shown at the right side of this figure are the data for the three materials at these two wavelengths, plotted with respect to one another. For example, in the left-hand graph, soil has the largest response at wavelength $\lambda 1$; this manifests itself in the right-hand plot in the fact that soil has the largest abscissa value (the greatest displacement to the right).

It is readily apparent that two materials whose response as a function of wavelength are different will lie in different portions of the two-dimensional space.* When this occurs one speaks of the materials involved as having unique spectral signatures. This concept will be pursued further shortly; however, at this point it is important to recognize that the concept of a spectral signature is a relative one--one cannot know that vegetation has a unique spectral signature, for example, until one sees the plots resulting from the spectral response of other materials within the scene to be analyzed. Note also that a larger number of bands can be used. The response at λ_3 could be used and the data plotted in three dimensions. Four or more dimensions indeed have meaning and utility even though an actual plot of the data is not possible.

So far, no spatial or temporal information has been involved, only spectral. Temporal information can be utilized in several ways. Consider Figure 8. Time is always a parameter of the spectral response of surface materials. As an example, consider the problem of discriminating between soybeans and corn. Under cultivation, these two plants have approximately 140-day growing cycles. Figure 8 illustrates what the two-dimensional response plot might be for fields of these two plants with time as a parameter. Upon planting and for some period thereafter, fields of soybeans and corn would merely appear to be bare soil from an observation platform above them. Eventually though, both plants would emerge from the soil and in time develop a canopy of green vegetation, mature to a brownish dry vegetation, and diminish. Thus, as viewed from above, the fields of soybeans and corn would, in fact, always be mixtures of green vegetation and soil. In addition to the vegetation of the two plants having a slightly different response as a function of wavelength, the growing cycles and plant geometries are different; thus, the mixture parameters might (and in fact do) permit an even more obvious difference between the two plants than the spectral response difference of the plant leaves themselves. This is the implication in Figure 8 as shown by the rather large difference between them 30 days from planting date (partial canopy) as compared to 75 days (full canopy). Thus, one way in which temporal information is used is simply in determining the optimum time at which to conduct a survey of given materials.

---

* This space is referred to as *feature space* in pattern recognition terminology.

Let us move now to consider how one may devise a procedure for analyzing multispectral data. In the process, one further facet of the multispectral approach must be taken into account. The radiation from all soybean fields will not have precisely the same spectral response, since all will not have had the same planting date, soil preparation, moisture conditions and so on. Indeed, response variation within a class may be expected of any earth-surface cover. The extent of response differences of this type certainly has an effect upon the existence of a distinct spectral signature, that is, the degree of separability of one material from another. Consider, for example, a scene composed of soybeans, corn and wheat fields; if five samples of each material are drawn, the two-dimensional response patterns might be as shown in Figure 9 indicating some variability exists within the three classes. Suppose now an unknown point is drawn from the scene and plotted, as indicated by the point marked U.

The design of an analysis system in this case comes down to partitioning this two-dimensional feature space in some fashion, such that each such possible unknown point is uniquely associated with one of the classes. The engineering and statistical literature of the world abounds with algorithms or procedures by which this can be done. One very simple one is shown in Figure 10. In this case the conditional centroid or center point of each class is first determined. Next the locus of points equidistant from these three centroids is plotted and results in the three segments of straight lines as shown.* These lines form, in effect, decision boundaries. In this example the unknown point "U" would be associated with the class soybeans as a result of the location of it with respect to the decision boundaries.

This technique of analysis is referred to as pattern recognition, and there are many more sophisticated procedures resulting in both linear and nonlinear decision boundaries. However, the procedure of using a few initial samples to determine the decision boundaries is common to a large number of them. The initial samples are referred to as training samples, and the general class of classifiers in which training samples are used in this way is

---

* When more than two dimensions (spectral bands) are being used, this locus would become a surface rather than a line.

referred to as supervised classifiers. Pattern recognition will be examined in more detail in a later section.

Pretest question

Imagine you are to brief a group of visitors on the operation of an airborne multispectral scanner. Figure 11 of these notes is available for reference. List in outline form the main points that you would include in your briefing.

Go on to page 19.

Are you satisfied with the work you just did?

If so, turn to page 20.

If not, continue reading below.

A TYPICAL AIRBORNE MULTISPECTRAL SCANNER SYSTEM

Up to this point, the implication has been that photography or multispectral photography is the sensor to be used in generating data for this type of an analysis procedure. While this data source can be used, a perhaps more appropriate one is a device known as a multispectral scanner. Figure 11 diagrams such a device as might be mounted in an aircraft.

Basically, the device consists of a multiband spectrometer whose instantaneous field of view is scanned across the scene. The scanning in this case is accomplished by a motor-driven scanning mirror. At a given instant the device is gathering energy from a single resolution element. The energy from this element passes through appropriate optics and may, in the case of the visible portion of the spectrum, be directed through a prism. The prism spreads out the energy from different portions of the spectrum. Detectors are located at the output of the prism. The output of the detectors can then be recorded on magnetic tape or transmitted directly to the ground. Gratings are commonly used as dispersive devices for the infrared portion of the spectrum.

A most important property of this type of system is that all energy from a given scene element in all parts of a spectrum pass through the same optical aperture. Thus, by simultaneously sampling the output of all detectors one has, in effect, determined the response as a function of wavelength in each spectral band for the scene element in view at that instant.

The rotating mirror causes the scene to be scanned across the field of view transverse to the direction of platform motion; the motion of the platform (aircraft) provides scanning of the scene in the direction of the flight so that in time every element in the scene has been in the instantaneous field of view of the instrument.

As an example of the use of a multispectral scanner system of the type shown in Figure 11 and the multispectral analysis procedure, results of the analysis of a flight line will be presented in brief form. The particular example involves the classification of a one-mile by four-mile area into classes of agricultural significance. Four-dimensional data (four spectral bands) were used for classification. The data are shown in Figure 12 along with a conventional panchromatic air photo of the scene in which the correct classification of each field has been added to the photo by hand. The symbols on the air photo and their associated classes are as follows: S-soybeans, C-corn, O-oats, W-wheat, A-alfalfa, T-timothy, RC-red clover, R-rye, SUDAN-sudan grass, P-pasture, DA-diverted acres, and H-hay.

Figure 13 shows the results of the classification. Two sample classes are shown. All points of the scene classified as row crops (either corn or soybeans) are indicated in the center of the figure. On the right are indicated all points classified as cereal grains (either wheat or oats).

A quantitative evaluation of the accuracy was conducted by designating for tabulation the correct class of a large number of fields in the flight line. The result of this tabulation is shown in Figure 14. It is seen that all results for all classes are above 80% correct.

The same procedures using aircraft data have been utilized for a wide number of classification tasks in addition to crop species identification. Some of these are as follows: tests of agricultural and engineering soils; mapping and delineating soil types, mineral content, organic content and moisture content of the soil; geologic feature mapping; water quality mapping and mensuration using both reflective and emissive spectra; forest cover identification and tree species delineation; and delineation into geographic and land-use mapping categories.

Before examining the ideas of pattern recognition in more detail it is desireable to make an important point concerning the choice of classes. There are two conditions that a class must meet in order to be useful. The class must be *separable** from all others and it must be of *informational value.* For example, it does no good to define a class called iron ore deposits if the spectral response which iron ore provides is not sufficiently distinct from all other earth-surface materials over which data are to be gathered. On the other hand, if no one is interested in locating the iron ore deposits within the region to be surveyed, there is no reason to define such a class. One may name classes of informational value and then check their separability, or vice versa.

---

* Spectrally discriminable

An opportunity is now provided for a deeper examination of pattern recognition concepts. To determine the best starting point for you, consider the following pretest item:

A block diagram of a pattern recognition system is shown in Figure 26. With this figure as an aid, explain the meaning of the terms **receptor, feature vector, categorizer** *and* **decision rule.**

Turn to page 22.

You will have to judge yourself on this one. If you feel comfortable with your answer to the last question go on to page 29.

Otherwise, turn to page 23.

What is automatic pattern recognition?

Generally, the term *pattern recognition* as used in the technical literature refers to techniques and equipment for the automatic recognition of patterns. The emphasis here is on automatic, since this field has been developed to handle problems in which the large quantity of data requires complete reliance upon a machine for classification.

There appear to be similarities between pattern recognition and photointerpretive techniques. As with photointerpretation, pattern recognition requires the development of a key, a set of tests which are to be carried out on a candidate pattern to determine its correct classification. The similarity ends at this point, however, due to the nature of the sets of tests in the two cases and the way they must be implemented. In the case of photointerpretation, the tests are usually relatively sophisticated and require human attention. On the other hand, the purpose of pattern recognition is to permit the complete removal of man from the process in order to be able to process data faster.

Thus, in comparing pattern recognition and photointerpretation, it may be said that photointerpretation is generally more suited for problems of *higher* sophistication involving *lower* data quantities, while just the reverse is true of pattern recognition.

In order to further clarify what is meant by the terms *pattern* and *pattern recognition*, a number of examples of current and important problems are presented.

Probably the first thing that comes to mind upon hearing the term *pattern recognition* is the problem of automatically recognizing various geometrical patterns. Examples of this type of problem are:

1) Reading of typed, printed, or handwritten text

2) Recognition of a person from his handwriting

3) Distinguishing manmade from natural objects on aerial photographs.

But pattern recognition is not limited to these cases, as evidenced by the following examples:

4) Recognition of the spoken word, for various speakers, *e.g.* human voice-to-computer communication

5) Recognition of a speaker regardless of the words spoken

6) Recognition of an environment or situation in which a system is placed. (Important for adaptive automatic control and learning systems)

7) Recognition of the location of faults in complex electronic systems

8) Character or signal transmission recognition over lines of communication, *e.g.* communication between computers

9) Target identification of aircraft, submarines, and missiles, and distinction from decoys using radar, sonar, etc.

10) Recognition of fields of agricultural crops, their condition, and state of growth from aerial observation.

The pattern recognition device

The problem of designing devices which classify patterns requires two main investigations. The first investigation involves the problem often referred to as *feature extraction*, *i.e.*, operations on the pattern which determine its significant characteristics. The second investigation involves the decision-making device which classifies the pattern on the basis of the comparison of its characteristics (both similarities and differences) with those of a reference set of patterns. We will now look at each of these problems in more detail.

Generally, in a pattern recognition problem a number of measurable quantities exist which are used to characterize the patterns. The optimum choice of these quantities (called features) represents the feature selection problem mentioned above. This problem, by no means a trivial one, is as yet unsolved, and, in fact, is the major stumbling block to the total unification of all the applications of pattern recognition. Often the designer must use his intuition based on some prior experience to choose what seems to be a suitable set of features. On the basis of these features, studies are undertaken to determine the best decision or classifying strategies to employ.

In accordance with this subdivision of the pattern recognition problem into two subproblems, the recognition device is generally designed in two parts, one part being called the *receptor* and the second being the *categorizer*. A simple block diagram is shown in Figure 26.

The input to the receptor is the pattern to be recognized. The receptor, using various sensors, performs the task of measuring the chosen features. The output of the receptor is a vector (called the measurement or feature vector) whose components denote the various feature measurements.

The categorizer portion of the recognition device is responsible for assigning a given input pattern to a class, on the basis of the measurement vector. The designer constructs the categorizer to obtain the "best" possible recognition of the patterns to be classified. The term "best" used here refers to the best performance as indicated by the measure chosen by the designer. It should be noted that the optimum design of the categorizer in a particular problem is carried out with respect to a given set of features. To obtain the best *overall system* it is necessary to then optimize over all sets of possible features.

An example problem

Perhaps the following example will be helpful in visualizing the operation of a pattern recognition device, and the function played by its components.

Consider the problem of remotely sensing whether a given field contains wheat, corn, or alfalfa. Assume we have decided that the percentage reflectance of electromagnetic energy in certain selected regions of the spectrum are the features. This choice of features could have been chosen, for example, by examining the characteristics of the various crops as they would appear from the air. The receptor portion of the pattern recognition device then measures the percentage reflectance in the selected frequency bands.

Let x_1 be the percentage reflectance in band one, x_2 in band two...x_n in band n where n is the total number of features measured.

The ordering of these features forms a measurement vector $\underline{x} = (x_1, x_2 \ldots x_n)$, and on the basis of this vector the categorizer is then to decide if the field is wheat, corn, or alfalfa.

We will examine this example further to introduce the concept of a measurement, or feature space, and then show how some of the common decision criteria can be represented in this space.

In order to represent a feature space easily on the plane of the paper, let us consider the situation in which we only measure two features (*i.e.*, reflectance in two spectral bands). Thus, the feature vector contains only two components $\underline{x} = (x_1, x_2)$.

The receptor then represents each field examined (at its input) by two numbers (at its output). To start the process we might examine 10 fields each of wheat, corn, and alfalfa, then plot and label the classification of each field.

Figure 27 shows a set of results which might be obtained. These patterns of known classification then constitute the reference set of patterns to which the patterns of unknown fields are compared. Obviously this reference set must be large enough and carefully selected, so that the set is typical of all future patterns to be classified. In practice, the selection of this set is crucial, and requires great care and judgment. The point labeled U in Figure 27 represents the feature vector of an unknown field whose classification is to be determined.

The job of the categorizer begins at this point. That is, to classify the unknown field on the basis of its representation in the chosen feature space. Many decision rules for making the classification have been proposed and studied in the technical literature. We will mention only a few here to illustrate the approach.

1. Minimum distance to the means criterion - According to this approach the mean vector of each known class is found and representated as a point in the feature space. See Figure 28. The pattern is classified into the class whose mean is closest. This criterion divides the space into three regions for classification as shown in Figure 28. Then, depending on whether the feature vector for the unknown field falls in regions A, C, or W, it is classified as alfalfa, corn, or wheat, respectively.

2. Minimum distance to the nearest member of a class - According to this criterion, the distance from the unknown pattern to each reference pattern for each class is determined, and the minimum distance found. The unknown pattern is then classified into the same class as that of the reference pattern nearest it. As before, this decision criterion divides the feature space into decision regions. A graphical representation of this is shown in Figure 29.

In these first two classification schemes, the assumption is made that the classes are sufficiently represented (characterized) by the limited number of known reference patterns. Specifically, in the example being considered it is assumed that the 10 reference patterns of each class are sufficient to characterize the various classes. The justification of this assumption is a problem in its own right. As will be seen, this same assumption is employed in the third method of classification to be discussed, but in a slightly different way.

Continue by turning to page 29.

3. <u>Statistical pattern recognition</u> - Assume for the moment that we have three joint probability density functions, one each for the classes wheat, corn, alfalfa. Let each represent the probability that the representation in feature space of a pattern of the particular class falls in a given region of the feature space. We might have the three density functions shown in Figure 30. For each category of interest, a set of likelihood ratios can be computed which express the relative probabilities that a candidate pattern belongs to the category of interest rather than to any of the others. Thus, points in the feature space are assigned to the class for which the probability of occurrence of that point is the largest. Figure 31 shows decision regions which might be obtained by this approach. It should be noted that most, if not all, optimum statistical decision criteria can be put in the form of ratio criteria such as this.

We return for a moment to the problem of obtaining the joint distribution functions, and see how this is connected with the basic assumption discussed above. The knowledge of the density functions could come about in one of two ways: 1) the probability densities are actually known, say through some theoretical study; 2) if the probability densities are not known, this knowledge must be gained by taking samples of each class. Again, a basic assumption employed is that the samples are sufficient to characterize the classes. In this case this means that the sample size is large enough to construct the probability densities required.

Using the geometrical interpretation of making classification decisions outlined above, we can summarize the discussion of the operation of the categorizer as follows. The feature space in which patterns are represented by points is divided into non-overlapping regions, one region corresponding to each of the categories. A classification decision consists of assigning to each candidate pattern the name of the category associated with the region in which the pattern is located.

Again, in practice, considerable judgment and experimentation are usually required to select the best categorizer approach for a given recognition task. A given categorizer may work well on one problem and not on another. Or two different ones may have the same error rates, but one may make different types (more costly) of errors and for different reasons.

A more realistic example

The above example was a hypothetical one designed to illustrate the approach. Let us now consider a more realistic situation. It is not usually possible in pattern recognition problems to conceive and design a receptor which is so effective that the various pattern classes are as obviously separable as they appear to be in Figure 27, at least not in only two dimensions.

Figure 32 illustrates a more typical situation, but again only in two dimensions in order to preserve the illustrative simplicity. The data for this illustration was obtained from actual measurements of the reflectance of wheat and oats in two spectral bands. These two bands, numbered 6 and 8, were selected from nine for which data happened to have been available. Notice that the two pattern classes are somewhat overlapping in two dimensional space, indicating that perfect classification will not be possible.

An even worse situation is illustrated in Figure 33. The data plotted here is that from spectral bands 3 and 5 of the same wheat and oat reflectance measurements used in the previous figure.

In order to design a pattern recognition system, it is necessary to have a quantitative measure of the effectiveness of the receptor. For example, in the above case data is available from nine spectral bands. Suppose that we are limited to the use of only two of the nine bands. A limitation of this type could come about due to limits in computer speed or memory size.* The question then arises: what pair of features would be the best to use?

A few such measures of receptor effectiveness are already available in the technical literature. One, the divergence criterion, was used in the generation of this example. Divergence is a quantity defined for pairs of probability density functions. It yields a number which is a function of the distance or separation between two densities; that is, the larger the separation, the greater the divergence.

In preparing this example, the divergence for each possible pair of the nine available features was computed. The feature pair with the highest divergence (bands 6 and 8) and a feature pair with a somewhat lower divergence (bands 3 and 5) were selected for presentation here.

---

* Actually, the size and speed of the average commercially available computer usually permits maximum dimensionalities of from 30 to several thousand, depending upon the particular categorizer algorithm.

To complete this example, patterns of oats and wheat were classified using a maximum likelihood ratio categorizer with the above two feature pairs, and assuming Gaussian statistics. The results are shown in Tables I and II, and the decision boundaries (which turned out to be two-sheeted hyperbolas) are shown in Figure 32 and 33. To improve the classification performance over that of Table I, the designer could (a) further optimize the receptor in some fashion, (b) find a categorizer more suited to this specific problem, and/or (c) go to a higher dimensionality. Actually, all nine features could easily have been used.

Table I. Feature 6 & 8

True Class	No. of Samples	No. classed as Oats	No. classed as Wheat	Percent Correct Classification
Oats	99	76	23	76.8
Wheat	78	11	67	85.9
Overall				80.1

Table II. Features 3 & 5

True Class	No. of Samples	No. classed as Oats	No. classed as Wheat	Percent Correct Classification
Oats	99	73	26	73.7
Wheat	78	33	45	57.7
Overall				66.6

Please turn to page 33.

You have finished the reading materials prepared to provide a background in remote sensing and pattern recognition. The post test given below will help determine your mastery of the material. Where unique answers exist they are given. Page references direct you to the sections of the reading material should review be necessary.

1a. *Define remote sensing.*

- determining the characteristics of an object from afar (i.e. not having direct contact)

b. *Give an example of a remote sensing application.*

- creating maps using aerial photographs etc

(Reference for question 1, page 5.)

2. *Remote sensing systems depend on one or more of several types of variations in electromagnetic fields. List at least 3 of these variations.*

a. temporal

b. spacial

c. spectral

(Reference for question 2, page 5.)

Continue by turning to page 34

Post Test (continued)

3. As they have developed to date, remote sensing systems tend to fall into one of two categories: image-oriented systems and numerically oriented systems.

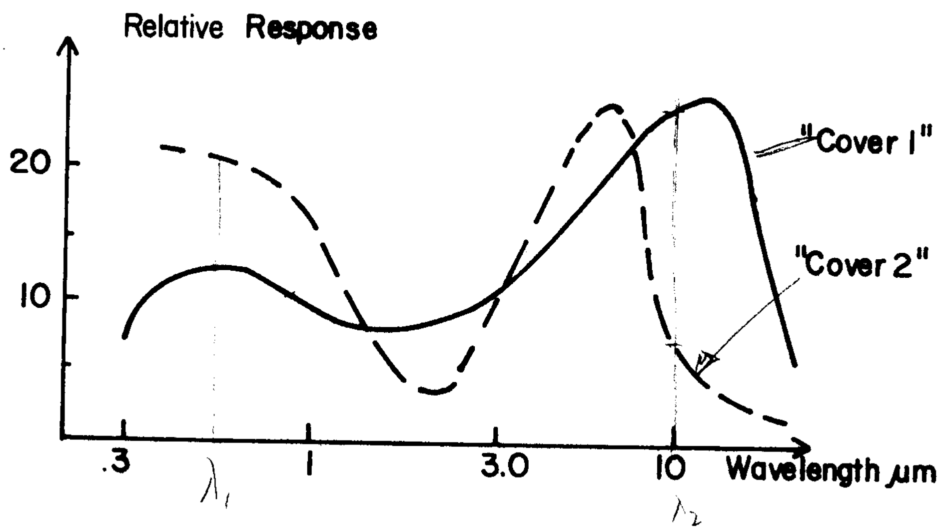
   Cite an example of each type of system.

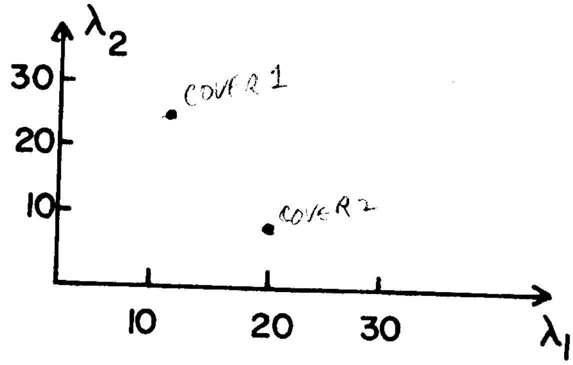
   a. photography & photointerpretation

   b. a census

(Reference for question 3, page 13, 14.)

4. Select two wavelengths λ_1 and λ_2 to serve as features for distinguishing "cover 1" from "cover 2" and plot the "cover 1" and "cover 2" points on the λ_1, λ_2 plane.





(Reference for question 4, pages 15 and 16.)

Please turn to page 52.

To reach this point you must have a reasonable background in pattern recognition. This branch of the instructional materials discusses remote sensing systems emphasizing the important role played by pattern recognition.

WHAT IS REMOTE SENSING? HOW IS INFORMATION CONVEYED? *(continued)*

In addition to the optical wavelengths, the microwave range is also useful in remote sensing. Preliminary results using both passive microwave and radar sensors indicate considerable promise for this microwave portion of the spectrum. For reasons of simplicity and in the interest of time, however, we shall limit our considerations in the remainder of this discussion to the optical portion of the spectrum.

Figure 2 is a diagram of the organization of an earth survey system. It is necessary, of course, to have a sensor system viewing that portion of the earth under consideration. There will necessarily be a certain amount of on-board data processing. This will perhaps include the merging of other data from sources such as sensor calibration and orientation.

One must next transport the data back to earth for further analysis and processing. This may be done through a telemetry system as is the case for the Earth Resource Technology Satellite (ERTS), or through direct package return as will be used with SKYLAB. There usually then is a need for certain preprocessing of the data before the final processing with one or more of the data reduction algorithms. It is at this point in the system, when the data is reduced to information, that it is usually helpful to merge ancillary information, perhaps derived from sources on the surface of the earth.

An important part of the system which must not be overlooked is indicated by the last block in Figure 2, that of information consumption, for there is no reason to go through the whole exercise if the information produced is not to be used. In the case of an earth resource information system, this last portion can prove to be the most challenging to design and organize since many potential consumers of this information are not accustomed to receiving it from a space system and may indeed know very little about the information-providing capabilities.

Before leaving the matter of the organization of an information system, the necessity of having a thorough understanding of the portion of the system preceding the sensor must be pointed out. Consider Figure 3. This figure shows a simplified version of the energy exchange in a natural environment. It is possible, of course, to detect the presence of vegetation on the earth's surface by measuring the reflected and emitted radiation emanating from the vegetation. One must understand, however, that there are many experimental variables active. For example, the sun provides a constant source of illumination from above the atmosphere, but the amount of radiation which is reflected from the earth's surface depends upon the condition of the atmosphere, the existence of surrounding objects, and the angle between the sun and the earth's surface as well as the angle between the earth's surface and the point of observation. Even more important is the variation which will exist in the vegetation itself. It is possible to deal with these experimental variables in several ways. We shall touch briefly on this point later in the discussion.

Summarizing, then, it is possible to derive information about the earth's surface and the condition of its resources by measuring the spectral, spatial, and temporal variations of the electromagnetic fields emanating from points of interest and then analyzing these measurements to relate them to specific classes of materials. To do so, however, requires an adequate understanding of the materials to be sensed and, in order to make the information useful, a precise knowledge is required about how the information will be used and by whom.

Go on to page 37.

THE DUALITY OF SYSTEM TYPES

When we consider the state-of-the-art of remote sensing today, a duality of system types becomes readily apparent. Development in the field has had two major stems because it originated from two somewhat different types of technology. These two types of systems will be referred to here as those with

- Image orientation, and

- Numerical orientation.

An example of an image-oriented system might be simply an aerial camera and a photointerpreter. The photographic film is used to measure the spatial variations of the electromagnetic fields, and the photointerpreter relates these variations to specific classes of surface cover. Numerically oriented systems, on the other hand, tend to involve computers for data analysis. Although the photointerpreter and the computer, respectively, tend to be typical in the two types of systems, it would be an oversimplification and indeed incorrect to say that they are uniquely related to these systems types. This becomes clearer upon further examination.

Figure 4 compares the organization of the two system types. Both types of systems need a sensor and some preprocessing; however, the distinction between the types can perhaps be brought out most clearly by noting the location of the *form image* block in the two diagrams. In the image-oriented type, it is in line with the data stream and must precede the analysis block. Numerically oriented systems, on the other hand, need not necessarily contain a *form image* block. If they do, and in earth resources studies they usually do, it may be at the side of the main data stream, as shown. It may thus be used to monitor the operation of the system and perhaps to do some special purpose analysis as needed. An image is, of course, the most efficient way to convey a large amount of information to a human operator. As seen, this is its principal use in both types of systems, but the use is different in the two cases.

In considering the design of information gathering systems, it is of great importance that the type of sensor as well as the means of analysis to be used are well-mated to the type of system orientation. Thus, let us briefly consider the types of imaging space sensors available.

Perhaps the single most distinguishing characteristic of earth resources information systems is that a very large amount of data can be, and indeed must be, gathered in order to derive the desired information. Since an image is a very efficient way to communicate large quantities of data to man, let us arbitrarily restrict ourselves to sensors which are capable of creating images. Shown in the table on the next page is a categorization of imaging sensors into three broad classes: photographic, television, and scanner. The table also provides examples of the advantages and disadvantages of each.

Table III

Sensor Type	Advantages	Disadvantages
Photography	Spatial Resolution	Data Return
Television	Size/Weight	Spectral Range
Scanner	Spectral Range	Mechanical Complexity

Types of imaging space sensors

In the case of photography, the great advantage, of course, is the very high spatial resolution which can be achieved, but to maintain this high resolution, data return by direct package return is required. Also, photography as a sensor is useful only in the visible and in a small part of the reflective infrared portion of the spectrum.

Television has the advantage that the signal occurs in electrical form and thus is immediately ready to be transmitted back to the earth; storage of the data, however, is not inherently present in the system in a permanent form, as it is in the case of photography. Thus, for space systems purposes one is not necessarily faced with the task of carrying along a large quantity of the storage medium (photographic film for the photography case) when using a television sensor. One may view this either as an advantage of size and weight or as one of efficiency in that a satellite may be operated a very long time with a single servicing. Television sensors are restricted to approximately the same spectral range as photography, however.

Scanners can be built to operate over the entire optical wavelength range. They can also provide a greater photometric dynamic range. In order to achieve these advantages, however, they tend to be more mechanically complex.

It is important to realize that the advantages and disadvantages here must be considered only as examples since the advantages and disadvantages in any specific instance will depend upon the precise details of the system. General statements are also difficult relative to the type of sensors which will be best for image-oriented and numerically oriented systems. There is a clear tendency to favor photography for image-oriented systems due to its high spatial resolution capability, while multiband scanners tend to be favored for numerically oriented systems since they make available greater spectral and dynamic ranges.

The technology for image-oriented systems is relatively well-developed. Sensors best suited to this type of system have been in use from some time, as have appropriate analysis techniques. This type of system also has the advantage of being easily understandable to the layman or neophyte to remote sensing, an advantage important in the earth resources field since, as it was pointed out above, many new data users are anticipated. Similarly, it is well-suited for producing subjective information and is especially

suited to circumstances where the classes into which the data are to be analyzed cannot be precisely decided upon before hand. Thus, man with his superior intelligence is or can be, deeply involved in the analysis activity. Image-oriented systems also have the possibility of being relatively simple and low-cost. On the other hand, it is difficult to use them for large-scale surveys over very large areas involving very large amounts of data.

In the case of numerically oriented systems, the technology is much newer and not nearly so well-developed, though very rapid progress is being made. Because the various steps involved tend to be more abstract, they tend to be less readily understandable by the layman. This type of system is best suited for producing objective information, and large-scale surveys covering large areas are certainly possible. Numerically oriented systems tend to be generally more complex, however.

In summary, the state-of-the-art is such that there are two general types of systems; this duality exists primarily for historical reasons and because of different origins from which technology development began. One type is based on imagery, and therefore a key goal of an intermediate portion of the system is the generation of high quality imagery. In the other case imagery is less important and indeed may not be necessary at all. It is not appropriate to view these two types of systems as competitive with one another since they have different capabilities and each has advantages under certain circumstances. As a matter of fact, these two stems of technology are approaching one another so that the differences between them are becoming less distinct.

We will proceed now to a further consideration of numerically oriented systems since this type may be less familiar. In particular we shall examine a type of data analysis useful in this case.

Please turn to page 40

THE MULTISPECTRAL APPROACH AND PATTERN RECOGNITION

In recent years considerable effort has been devoted to what is referred to as the multispectral approach for data analysis. An initial understanding of what is meant by the term "multispectral approach" may be obtained by considering Figure 5. Shown here in the upper left of the figure is a reproduction of a conventional color photograph of a set of color cards. The remainder of the figure shows photographs of the same color cards taken with black and white film and several different filters. The pass band of each filter is indicated beneath the particular color card set. For example, in the .62-.68 micrometer band, which is in the red portion of the visible spectrum, the red cards appear white in the black and white photo, indicating a high response or a large amount of red light energy being reflected from these cards. In essence the multispectral approach amounts to identifying any color by noting the set of gray scale values produced on the black and white photographs for that particular color rectangle.

As a very simple example of the approach, Figure 6 shows images of an agricultural scene taken in three different portions of the spectrum. Note that in the three bands alfalfa has responses which are dark, light, dark respectively whereas bare soil is gray, dark, white. Thus, alfalfa can be discriminated from bare soil by identifying the fields which are dark, light, dark in order in these three spectral bands.

One may initially think of the multispectral approach as one in which a very quantitative measure of the color of a material is used to identify it. Color, however, is a term usually related to the response of the human eye; the terminology of spectroscopy which is more precise is more useful in understanding the multispectral approach and is applicable beyond the visible region of the spectrum.

In order to understand this approach and to see how a numerically oriented system may be based upon it, consider Figure 7. Shown at the left is a graph of relative response (reflectance) as a function of wavelength for three types of earth-surface cover material: vegetation, soil, and water. Let two wavelengths marked λ_1 and λ_2 be selected. Shown in the righthand part of this figure is

the data for the three materials at the two wavelengths, plotted with respect to one another. For example, in the left graph soil has the largest response at wavelength λ_1; this manifests itself in the righthand plot in the fact that soil has the largest abscissa value (the greatest displacement to the right).

It is readily apparent that two materials whose response as a function of wavelength are different will lie in different portions of the two-dimensional space.* When this occurs one speaks of the materials involved as having unique spectral signatures. This concept will be pursued further presently; however, at this point it is important to recognize that the concept of a spectral signature is a relative one--one cannot know that vegetation has a unique spectral signature, for example, until one sees the plots resulting from the spectral response of other materials within the scene to be analyzed. Note also that a larger number of bands can be used. The response at λ_3 could be used and the data plotted in three dimensions. Four or more dimensions indeed have meaning and utility even though an actual plot of the data is not possible.

So far no spatial or temporal information has been involved, only spectral. Temporal information can be utilized in several ways. Consider Figure 8. Time is always a parameter of the spectral response of surface materials. As an example, consider the problem of discriminating between soybeans and corn. Under cultivation, these two plants have approximately 140-day growing cycles. Figure 8 illustrates what the two-dimensional response plot might be for fields of these two plants with time as a parameter. Upon planting and for some period thereafter, fields of soybeans and corn would merely appear to be bare soil from an observation platform above them. Eventually though, both plants would emerge from the soil and in time develop a canopy of green vegetation, mature to a brownish dry vegetation, and diminish. Thus, as viewed from above, the fields of soybeans and corn would, in fact, always be mixtures of green vegetation and soil. In addition to the vegetation of the two plants having a slightly different response as a function of wavelength, the growing cycles and plant geometries are different; thus, the mixture parameters might (and in fact do) permit an even more obvious difference between the two plants than the spectral response difference of the plant leaves themselves. This is the implication in Figure 8 as shown by the rather large difference between them 30 days from planting date (partial canopy) as compared to 75 days (full canopy). Thus, one way in which temporal information is used is simply in determining the optimum time at which to conduct a survey of given materials.

A second use of temporal information is perhaps less obvious. Consider the situation of Figure 8 at the 75- and 100-day point.

---

* This space is referred to as *feature space* in pattern recognition terminology.

In this case, the separation of the two materials is relatively slight. However, if this data is replotted in four dimensional space, λ_1 and λ_2 at 75 days as dimensions one and two and λ_1 and λ_2 at 100 days as dimensions three and four, the small separabilities at the two times may augment one another.

A third use of temporal information is simply that of change detection. In many earth resources problems it is necessary to have an accurate historical record of the changes taking place in a scene as a function of time.

Let us move now to consider how one may devise a procedure for analyzing multispectral data. In the process, one further facet of the multispectral approach must be taken into account. The radiation from all soybean fields will not have precisely the same spectral response, since all will not have had the same planting date, soil preparation, moisture conditions and so on. Indeed, response variation within a class may be expected of any earth-surface cover. The extent of response differences of this type certainly has an effect upon the existence of a distinct spectral signature, that is, the degree of separability of one material from another. Consider, for example, a scene composed of soybean, corn and wheat fields; if five samples of each material are drawn, the two dimensional response patterns might be as shown in Figure 9 indicating some variability exists within the three classes. Suppose now an unknown point is drawn from the scene and plotted, as indicated by the point marked U.

The design of an analysis system in this case comes down to partitioning this two dimensional feature space in some fashion, such that each such possible unknown point is uniquely associated with one of the classes. The engineering and statistical literature of the world abounds with algorithms or procedures by which this can be done. One very simple one is shown in Figure 10. In this case the conditional centroid or center point of each class is first determined. Next the locus of points equidistant from these three centroids is plotted and results in the three segments of straight lines as shown.* These lines form, in effect, decision boundaries. In this example the unknown point "U" would be associated with the class soybeans as a result of the location of it with respect to the decision boundaries.

This technique of analysis is referred to as pattern recognition, and there are many more sophisticated procedures resulting in both linear and nonlinear decision boundaries. However, the procedure of using a few initial samples to determine the decision boundaries is common to a large number of them. The initial samples are referred to as training samples, and the general class of classifiers in which training samples are used in this way is referred to as supervised classifiers.

---

* When more than two dimensions (spectral bands) are being used, this locus would become a surface rather than a line.

Pretest question

Imagine you are to brief a group of visitors on the operation of an airborne multispectral scanner. Figure 11 of these notes is available for reference. List in outline form those points that you would include in your briefing.

Go on to page 44

Are you satisfied with the work you just did?
If so, turn to page 45.
If not, continue reading below.

A TYPICAL AIRBORNE MULTISPECTRAL SCANNER SYSTEM

Up to this point, the implication has been that photography or multispectral photography is the sensor to be used in generating data for "this type" of an analysis procedure. While indeed this data source can be used, a perhaps more appropriate one is a device known as a multispectral scanner. Figure 11 diagrams such a device as might be mounted in an aircraft.

Basically, the device consists of a multiband spectrometer whose instantaneous field of view is scanned across the scene. The scanning in this case is accomplished by a motor-driven scanning mirror. At a given instant the device is gathering energy from a single resolution element. The energy from this element passes through appropriate optics and may, in the case of the visible portion of the spectrum, be directed through a prism. The prism spreads out the energy according to the portion of the spectrum; detectors are located at the output of the prism. The output of the detectors can then be recorded on magnetic tape or transmitted directly to the ground. Gratings are commonly used as dispersive devices for the infrared portion of the spectrum.

A most important property of this type of system is that all energy from a given scene element in all parts of a spectrum pass through the same optical aperture. Thus, by simultaneously sampling the output of all detectors one has, in effect, determined the response as a function of wavelength in each spectral band for the scene element in view at that instant.

The rotating mirror causes the scene to be scanned across the field of view transverse to the direction of platform motion, and the motion of the platform (aircraft) provides scanning of the scene in the direction of the flight so that in time every element in the scene has been in the instantaneous field of view of the instrument.

As an example of the use of a multispectral scanner system of the type shown in Figure 11 and the multispectral analysis procedure, results of the analysis of a flight line will be presented in brief form. The particular example involves the classification of a one-mile by four-mile area into classes of agricultural significance. Four-dimensional data (four spectral bands) were used for the classification. The data are shown in Figure 12 along with a conventional panchromatic air photo of the scene in which the correct classification of each field has been added to the photo by hand. The symbols on the air photo and their associated classes are as follows: S-soybeans, C-corn, O-oats, W-wheat, A-alfalfa, T-timothy, RC-red clover, R-rye, SUDAN-sudan grass, P-pasture, DA-diverted acres, and H-hay.

Figure 13 shows the results of the classification. Two sample classes are shown. All points of the scene classified as row crops (either corn or soybeans) are indicated in the center of the figure. On the right are indicated all points classified as cereal grains (either wheat or oats).

A quantitative evaluation of the accuracy was conducted by designating for tabulation the correct class of a large number of fields in the flightline. The result of this tabulation is shown in Figure 14. It is seen that all results for all classes are above 80% correct.

The same procedures using aircraft data have been utilized for a wide number of classification tasks in addition to crop species identification. Some of these are as follows: tests of agricultural and engineering soils; mapping and delineating soil types, mineral content, organic content and moisture content of the soil; geologic feature mapping; water quality mapping and mensuration using both reflective and emissive spectra; forest cover identification and tree species delineation; and delineation into geographic and land-use mapping categories.

Please turn to page 46

SOME PROCEDURAL DETAILS IN THE USE OF PATTERN RECOGNITION IN REMOTE SENSING

With the basic concept of pattern recognition in mind, it is possible to proceed to some further details on how it may be applied. One of the more important of these details is the definition of the classes into which the data are to be categorized.

There are two conditions that a class must meet in order to be useful. The class must be *separable** from all others and it must be of *informational value*. For example, it does no good to define a class called iron ore deposits if the spectral response which iron ore provides is not sufficiently distinct from all other earth-surface materials over which data are to be gathered. On the other hand, if no one is interested in locating the iron ore deposits within the region to be surveyed, there is no reason to define such a class. We shall see presently that one may name classes of informational value and then check their separability, or vice versa.

A second matter is determining the point at which a class actually becomes defined. In an agricultural survey, simply naming a class soybeans does not define it precisely enough. For example, what percent ground cover is required before a given resolution element should have its classification changed from bare soil to soybeans? What percent of a resolution element may be covered with weeds and so on? The fact of the matter is that the class becomes precisely defined only by the training samples to be used for it. Thus, an important step in the procedure is the selection of training samples which are sufficiently typical of the whole class in question.

---

* Spectrally discriminable

One must also recognize that the definition of a class is always a relative matter. That is, it is relative to the other classes used in the same classification. The effect of the decision boundaries is to divide up the feature space (see Figure 9) into non-overlapping regions depending on the relative location of the class training sets relative to one another.

It should also be noted, however, that as a result, every point in the space automatically becomes associated with one of the named classes. It is therefore necessary that the list of classes be exhaustive so that there is a logical class to which every point in the scene to be analyzed may be assigned.

As a result of these factors, it is apparent that the selection of training samples is especially important. There are two approaches to obtaining training data; we shall refer to them here as the signature bank approach and the extrapolation mode.

Using the signature bank approach, the researcher first decides on a list of appropriate classes and then draws from a signature bank previously collected data on the classes of material identical to those selected. This approach has a considerable amount of aesthetic appeal. Presumably one could accumulate a very large bank of data from typical classes and thereafter always have training data available for any situation without further effort.

However, such an approach would place stringent constraints on the sensor system since absolute measurements of scene radiance would be necessary if they are to be referenced to a future data-gathering mission. Further, the extent to which detailed and sophisticated classes could be utilized would be limited by the ability to determine and adjust for the instantaneous values of all the other experimental parameters, such as the condition of the atmosphere, the sun and view angle, possible seasonal variations in the vegetation, the natural statistical distribution of the data for various classes, etc. In short, while such a procedure is possible, it will result in more stringent requirements on the sensor system and considerable data preprocessing in order to achieve this maximum utility. Alternatively, it would have to be restricted to cases in which only relatively simple classes were necessary.

The extrapolation mode, on the other hand, has somewhat different characteristics. In this case, training data for each of the classes are obtained by locating within the data to be analyzed specific examples of each of the classes to be utilized. The classification procedure, therefore, will amount to an extrapolation from points of known classification within the scene to the remaining portions of the data. This approach has the advantage of requiring less exactness in the calibration capability of the sensor system and in the knowledge of the other experimental variables since only variation of these factors within the data-gathering mission, and not variations from mission to mission must be accounted for. On the other hand, it has the disadvantage of requiring some knowledge about the scene to be analyzed before the analysis can proceed. In the case of populated or accessible areas, this knowledge usually comes from ground observations. In the case of inaccessible areas as well as accessible and/or populated ones, it could perhaps also come from a very limited, low-altitude aircraft mission. The relative cost of this additional information often turns out to be relatively low. The extrapolation mode was used in both the preceding example and the one to follow.

To illustrate these details and procedures, an example follows in which a pattern-recognition system was trained and then used to classify a relatively large amount of data. Data for this experiment were collected aboard the Apollo 9 space vehicle as a part of an experiment known as S065. This example was selected because in addition to illustrating the steps described above, it provides the first indication of how these techniques may perform on space data. Both the ground resolution and the spectral resolution of these data are similar to those which will be obtained by the Earth Resources Technology Satellite; however, since the S065 experiment involved photographic sensors, the results obtained may be on the conservative side of those from ERTS since, as previously indicated, photography does not ordinarily provide the optimal type of data for this analysis procedure. Further, since the sensors were photographic, some preliminary processing steps to prepare the data for analysis were necessary. These steps involved first scanning the photography (on a rotating drum microdensitracer) to convert it to digital form, then bringing the images gathered over the same scene in the different spectral bands into spatial alignment with one another. These steps are not typical and are beyond the scope of the discussion at hand. We will proceed from the point at which the preprocessing steps provided four dimensional (four spectral bands) data for analysis. The four spectral bands involved were .47 to .61 micrometers, .59 to .71 micrometers, .68 to .89 micrometers, and .51 to .89 micrometers. These bands were determined by the film and filter combinations used on the four cameras.

Figure 15 shows a black and white print of a color infrared version of the particular frame used, a portion of Southern California, Arizona, and Northern Mexico. In the lower left of the frame is the Imperial Valley, an irrigated area of very great importance agriculturally. Also shown in the figure is a computer-generated gray scale printout of one band of the data. The scene covers about 10,000 square miles and contains about five million points.

In order to test the separability of various classes, two analysis tasks were carried out. The first, involving agricultural classes, was carried out in the area designated by the small rectangle in the lower left of the printout. Figure 16 shows a high resolution printout of this same area. The individual fields of the scene are clearly evident in this printout. To begin, some relatively simple classes were defined. These were green vegetation, bare soil, water, and salt flats. Figure 17 shows the result of classifying the data into these categories. The accuracy of this classification was judged to be very high and as a result it was decided to attempt a classification with more detailed categories. The result of this classification is shown in Figure 18. It is seen that the classes used were barley, alfalfa, sugar beets, bare soil, salt flat, and water. A quantitative assessment of the accuracy in this case indicated an average accuracy of approximately 70%.

The second analysis task carried out on this data set was done over the whole frame. Classes of geologic interest were defined in this case and an attempt was made to achieve what amounts to a geologic map of the area. The result of this classification is shown in Figure 19. Some, but not all, of the classes used are indicated at the bottom of this figure. The result of this classification was compared with existing geologic maps of the area by a professional geologist, and again the results were judged to be highly satisfactory.

Now, with an overview of the experiment and the results achieved in mind, let us examine the procedures used to obtain the results. In the case of an agricultural problem the classes of interest usually exist in well-defined fields. It is thus, relatively easy to locate sample fields of each class from which to draw training samples. In this case, ground observations from a relatively small region on the ground can be used to derive a sufficient number of training samples for each class. The number of training samples necessary for each class depends upon the number of spectral bands to be used among other things. But generally no more than a few hundred are required, fewer in simpler situations. Thus, Figure 20 depicts a typical situation for this type of classification. The fields outlined here are a typical set of training fields for such a classification task.

The classification of a natural area presents a slightly different situation, however, In this case it may be more difficult to manually locate training samples since boundaries between different materials will be more difficult to locate. Over the last year or two research has been directed towards devising some machine-aided procedures for deriving training samples in this circumstance. One such procedure involves the use of a type of classifier not utilizing training samples and thus referred to as a nonsupervised classifier. The basic idea behind nonsupervised classifiers becomes apparent by considering the next several figures.

Assume that one has some two-dimensional data as shown in Figure 21. Assume also that one knows there are three classes of material represented in this data, but the correct association of the individual points with the three classes is unknown. The approach is to initially assume that the three classes are separable and check this hypothesis subsequently.

There are algorithms (computational procedures) available which will automatically associate a group of such points with an arbitrary number of mode centers or cluster points. These procedures, known as clustering techniques, can be used to so divide the data and the result of applying such a procedure might be as shown in Figure 22. There remains, then, the matter of checking to be sure that the points assigned to a single cluster all belonged to the same class of material. In passing it is worth noting a comparison between supervised and nonsupervised classifiers. In the supervised case, one generally names classes of informational value and then checks to see if the classes are separable. The reverse is the case in the nonsupervised scheme. One separates the data and then checks to see if the clusters resulting are indeed associated with the classes of informational value.

Figure 23 shows the result of applying such a clustering technique to some multispectral data. The algorithm was instructed to form five cluster points. Comparison of the clustering results with the data in image form shows that the clusters indeed were associated with individual fields. Cluster four, for example, was associated with fields in the upper left and lower right, clusters two and three with the field in the lower left, and so on. Such a technique is used to speed the training phase of the classifier by aiding the human operator in obtaining points grouped according to the class that they came from; the statistics of each cluster point can be immediately computed from the clustered results so that decision boundaries are located. The operator is thus relieved of the necessity of locating and separating individual fields for training each class.

The value of such a procedure is even greater in cases where the boundary between classes is not so distinct in the data. Figure 24 shows the result of clustering data for a soils mapping classification. Here it would be more difficult to select samples associated with specific soil types. As a result of the clustering, the operator has only to associate the soil type with each cluster point and training samples are immediately available for further processing.

It was this latter procedure which was used in deriving training for the geologic map in the Apollo 9 data. Figure 25 shows the outline of cluster plots from which training was derived. In this case it was only necessary to quickly mark regions containing at least the samples of the classes desired, thus greatly simplifying and speeding the training of the classifier in this case. The specific steps to be followed then are:

1) decide upon the list of classes, and determine the general locality of examples of these classes based on limited ground observation. This information may be from a low-altitude aircraft pass, information available from a perhaps out-of-date or inaccurate map or a limited ground survey.

2) designate these regions to the clustering algorithm and after clustering identify the specific clusters associated with the classes of interest.

3) compute the statistics of each class from appropriate clusters and proceed with the classification.

It must be emphasized that the techniques are still evolving and are very experimental in nature. Many questions about the utility of the extrapolation mode and machine-aided training procedures remain to be answered. Not the least of these is the extent over which a given classifier can generalize or extrapolate from its training areas and the extent to which machine-aided training procedures do make the training of classifiers fast for practical situations. The Earth Resources Technology Satellite will, for the first time, provide a data base with which answers to these questions may be sought.

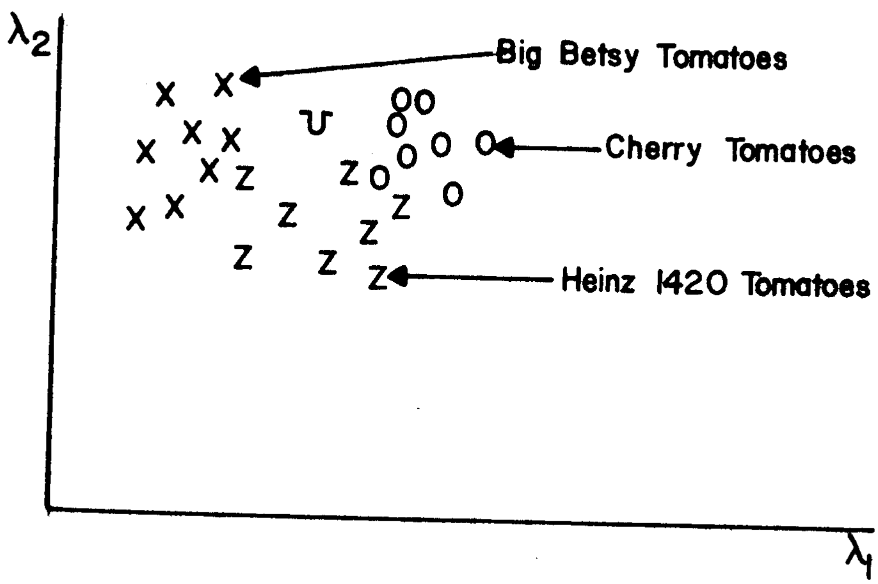
Please turn to page 21.

You have just come from page 34

Post Test (continued)

5. A fictitious pattern recognition system using two features uses the classification rule "assign data points to the class for which the distance between the point being classified and the closest training point is minimized." A typical situation showing training points for 3 classes and an unknown data point U *are shown below. Using the classification rule just stated predict which class the pattern recognition system would assign to* U.

HEINZ 1420 TOMATOES



Please turn to the next page.

The correct answer to question 5 is Heinz 1420 tomatoes. If you did not get this answer refer to pages 17 and 26 through 28.

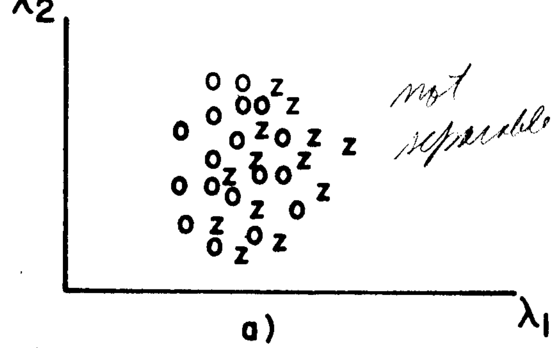
6. Figure 11 of these notes is a diagram of an airborne multispectral scanner system. Could you prepare a 5-minute talk on how this system works?

If your answer is a confident "yes" turn to the next page.

If your answer is no, reread page 19, then turn to page 54.

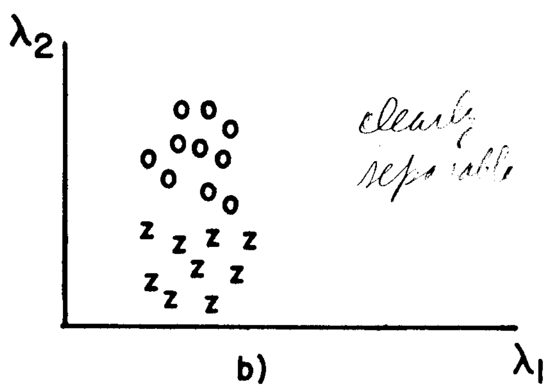
Post Test (continued)

7. *The success of a pattern recognition system is highly data-dependent. Below are three examples of data. Predict the relative success of a pattern recognition system subjected to these data sets by rating the data as clearly separable, marginally separable or not separable.*



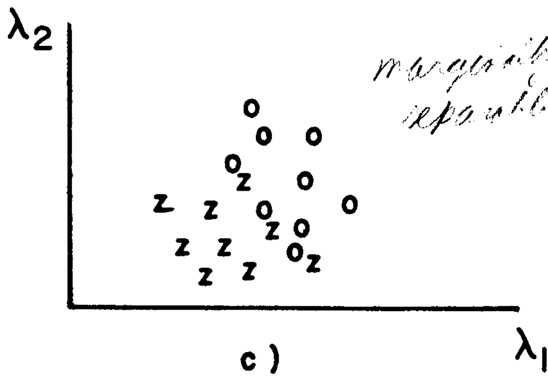
not separable

a)



clearly separable

b)



marginally separable

c)

Please turn to page 55

The answers to question 7 are

- *a) not separable*
- *b) clearly separable*
- *c) marginally separable.*

8. *What are the conditions one looks for in order for a class to be considered a useful class?*

- the information is distinguishable from other info
- it can be used

Please turn to page 56

The answer to question 8 is

a) the class must be spectrally distinguishable from all other classes

b) it must be of informational value.

References for question 8 are pages 20 or 46.

9. *Explain the meaning of the terms* receptor, feature vector, feature space, categorizer *and* decision rule *as they apply to the system shown in Figure 26 of these notes.*

Please turn to page 57.

Pages 23 to 28 serve as reference for question 9.

If you have answered all of the post test questions you have completed this portion of the training materials.

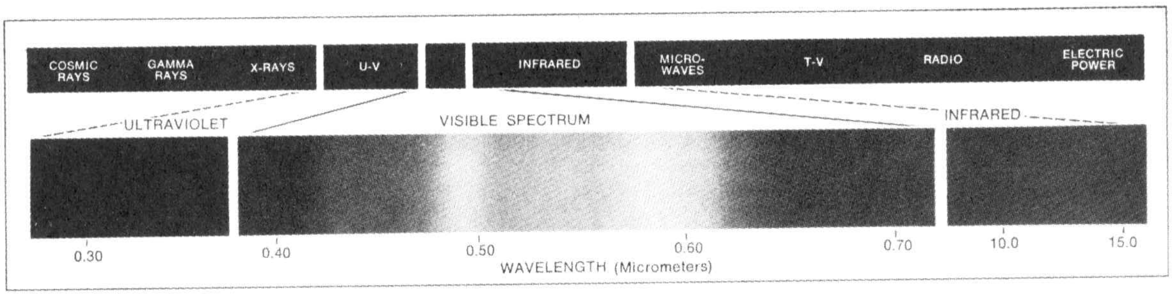


Figure 1. The Electromagnetic Spectrum.

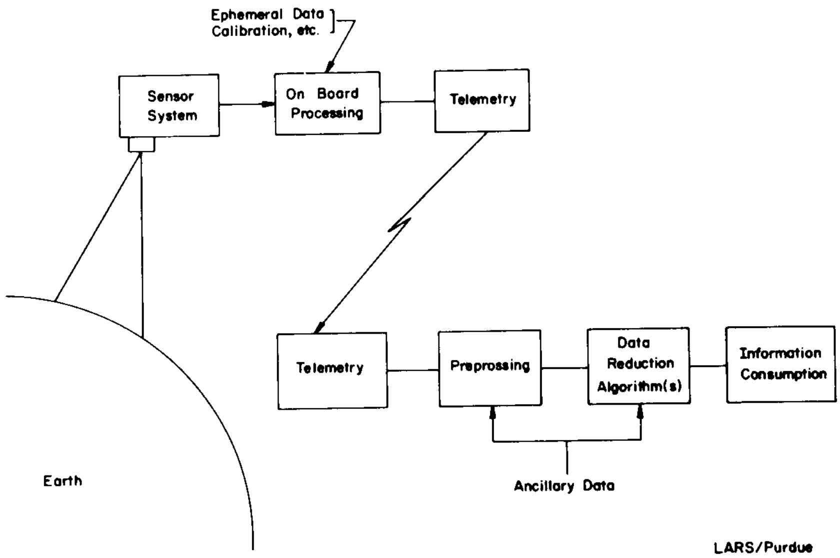


Figure 2. Organization of an Earth Survey System.

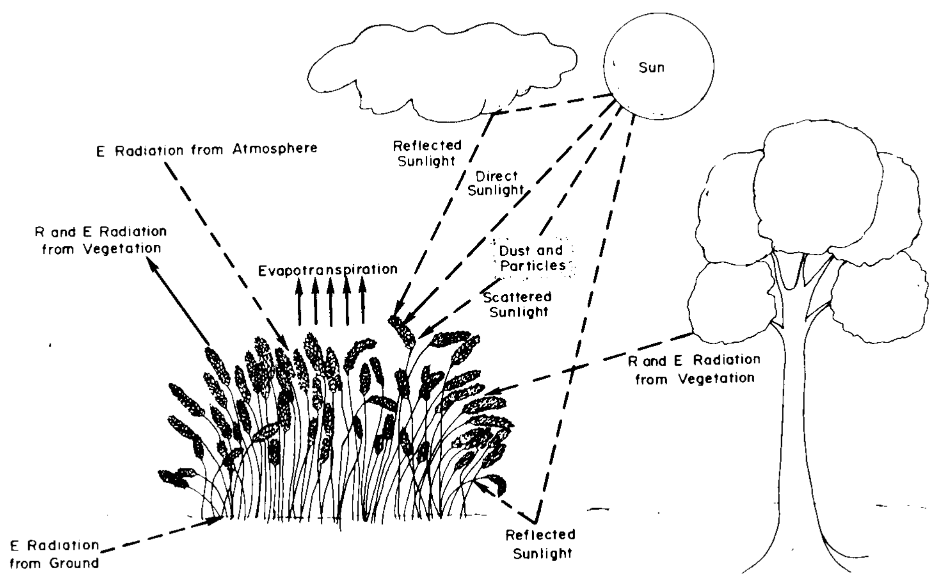


Figure 3. Reflected (R) and Emitted (E) Radiation Energy Exchange in a Natural Environment.

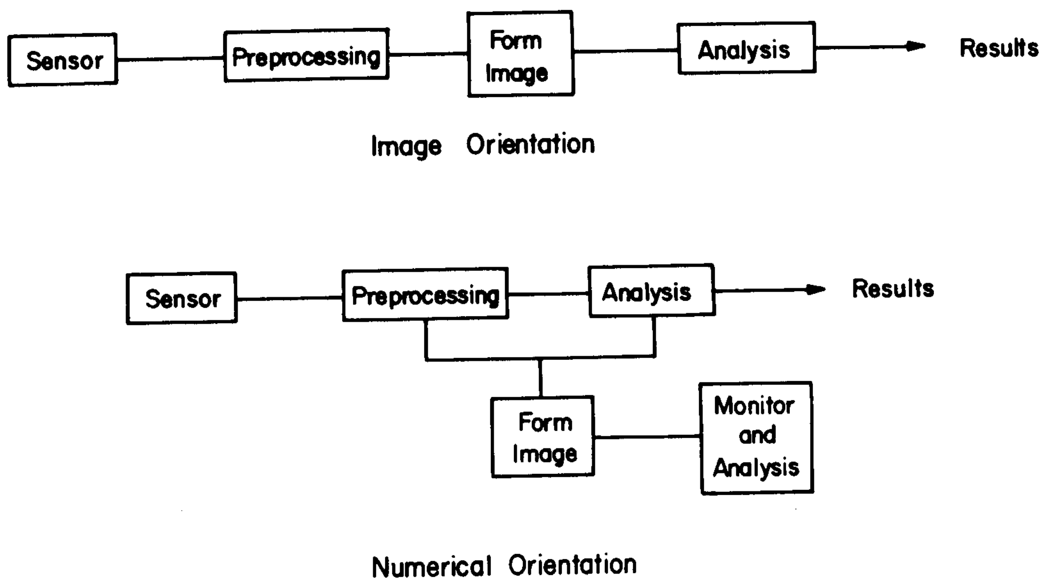


Figure 4. Organization of Image and Numerically Oriented Systems.

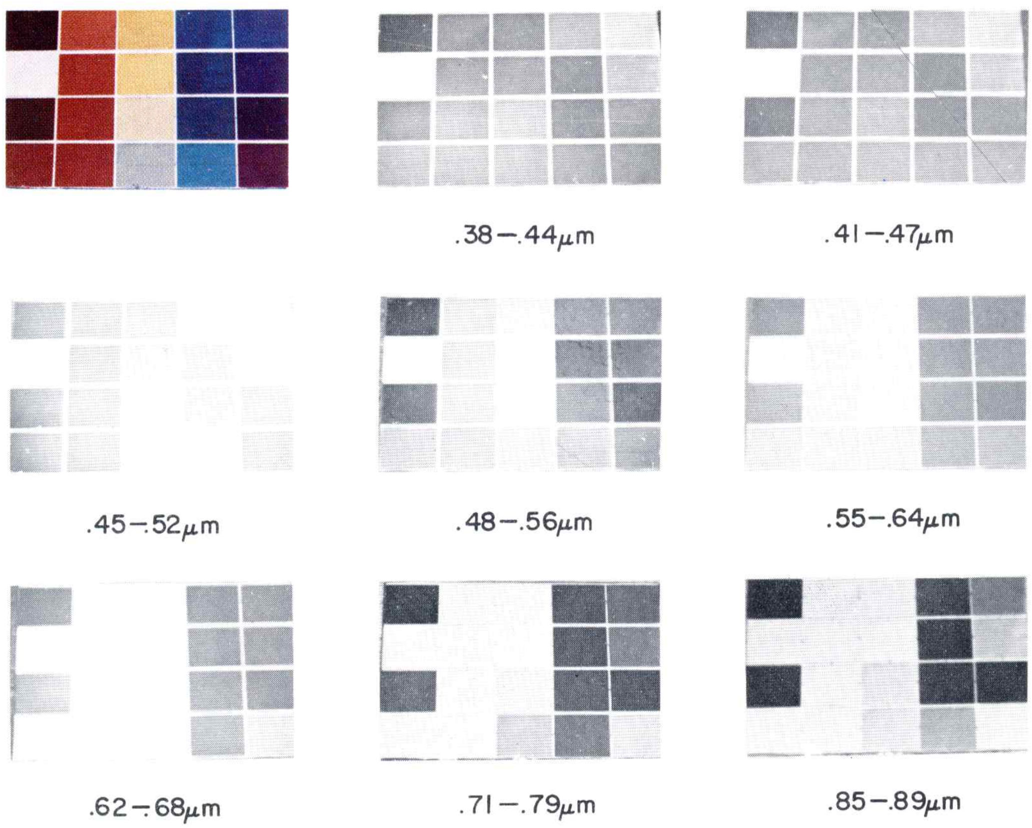
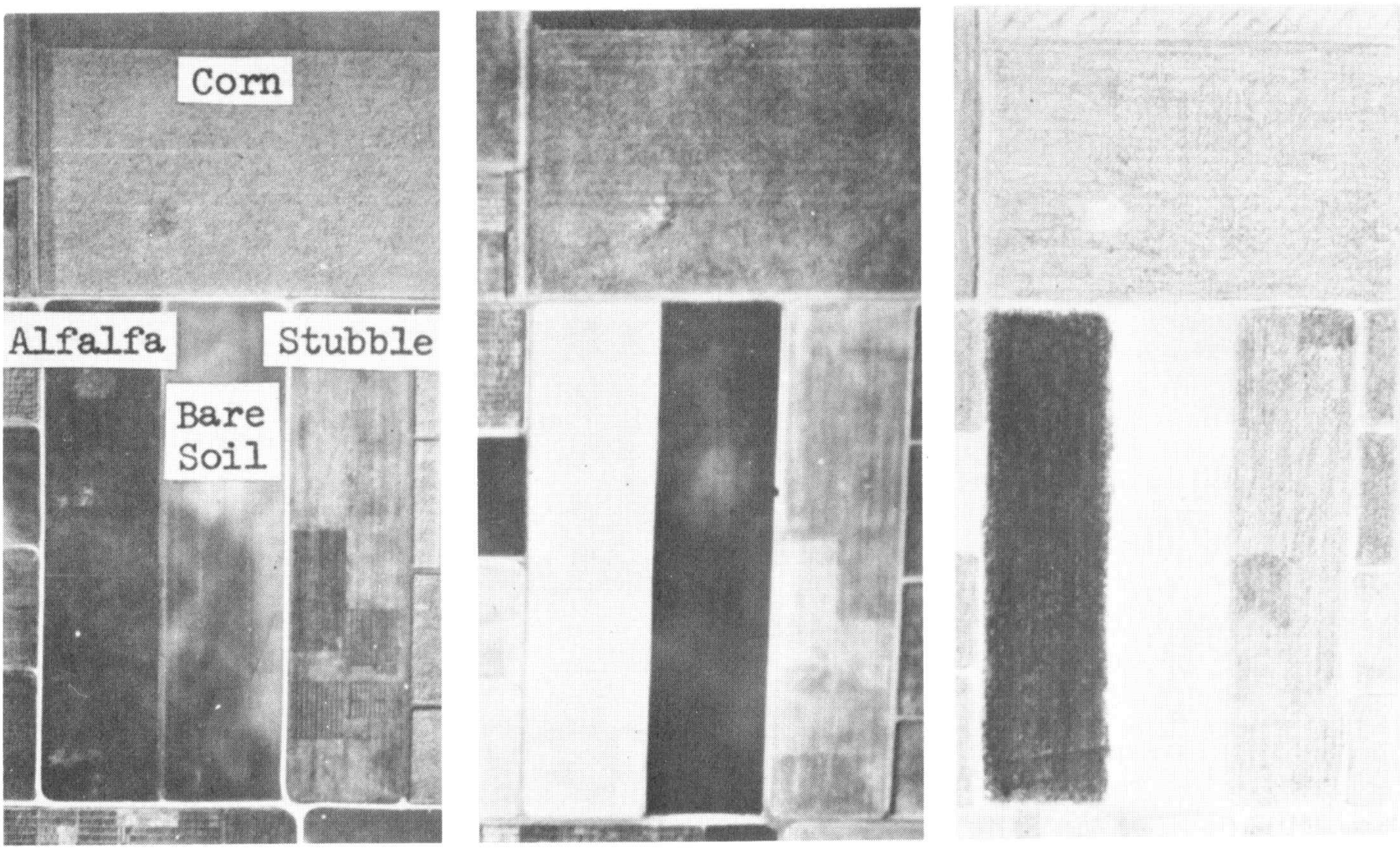


Figure 5. Multispectral Photography of Color Cards.



.4 - .7μm

.7 - .9μm

4.5 - 5.5μm
(Artist's Concept)

Figure 6. Multispectral Response of Corn Alfalfa, Stubble, and Bare Soil.

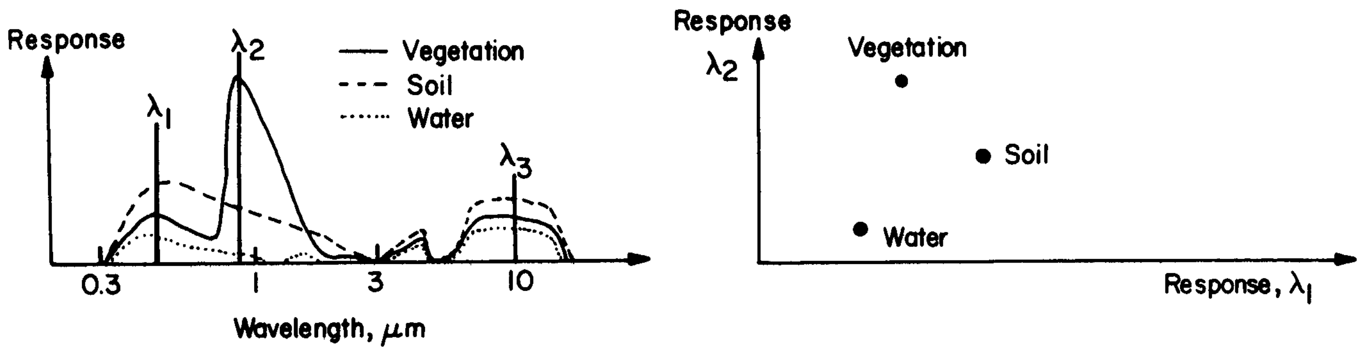


Figure 7. Spectral Data in Two-Dimensional Feature Space.

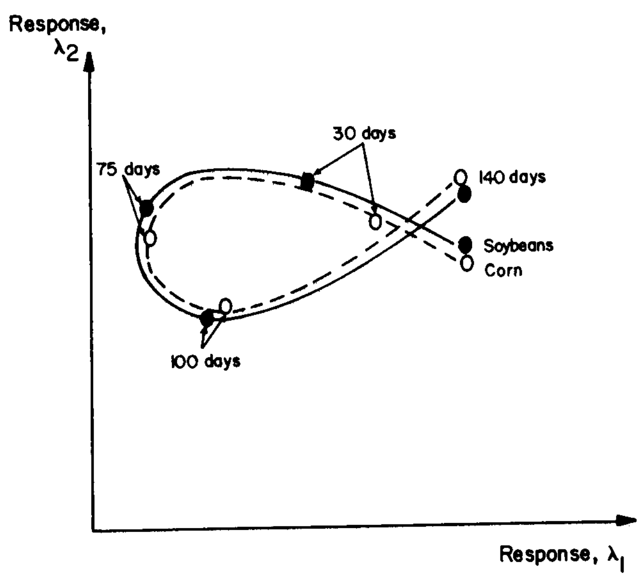


Figure 8. Temporal Change in Two-Dimensional Feature Space.

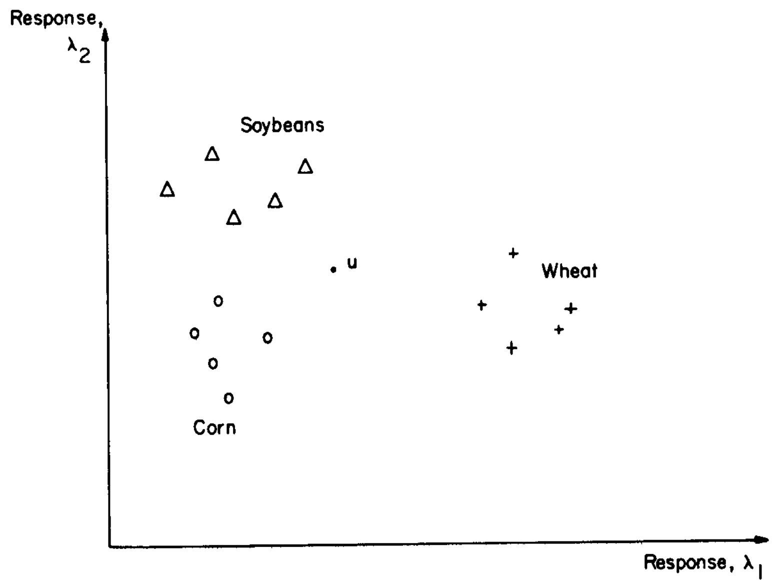


Figure 9. Samples in Two-Dimensional Feature Space.

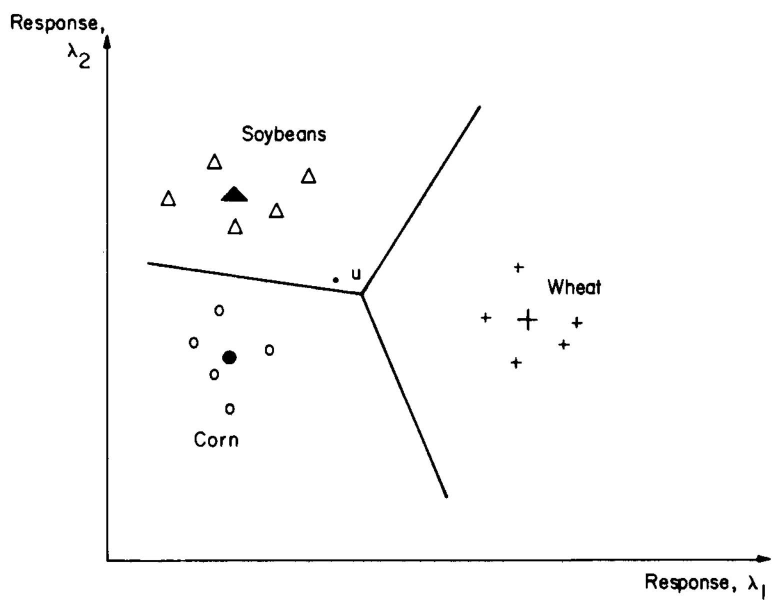


Figure 10. Minimum Distance to Means Classification.

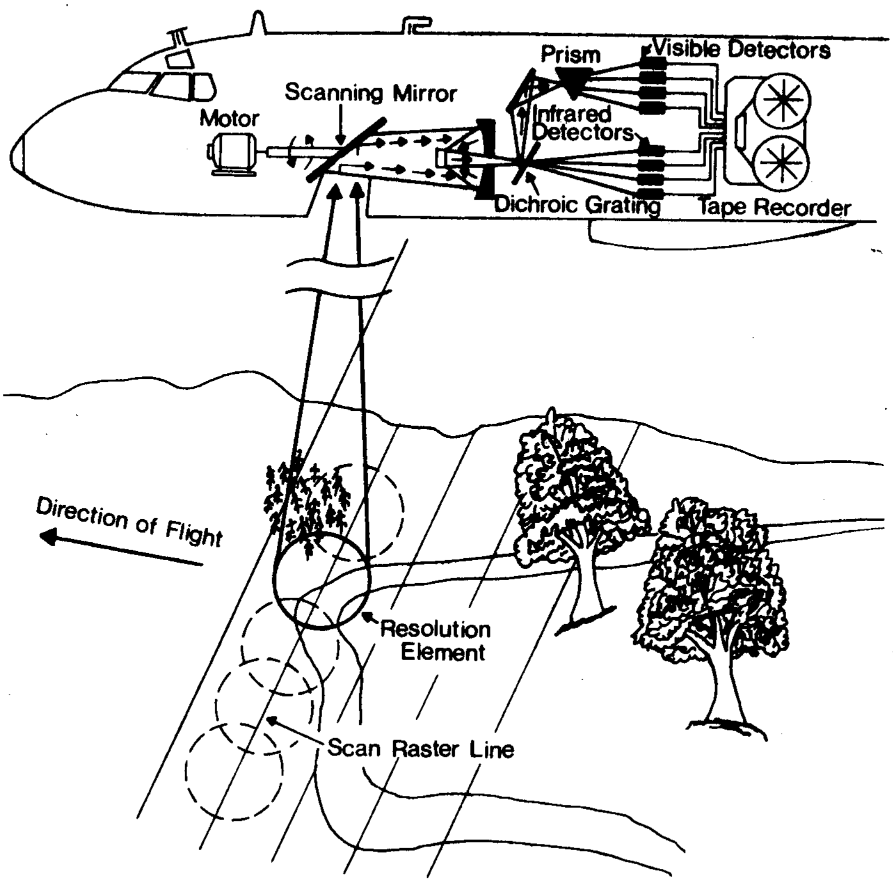
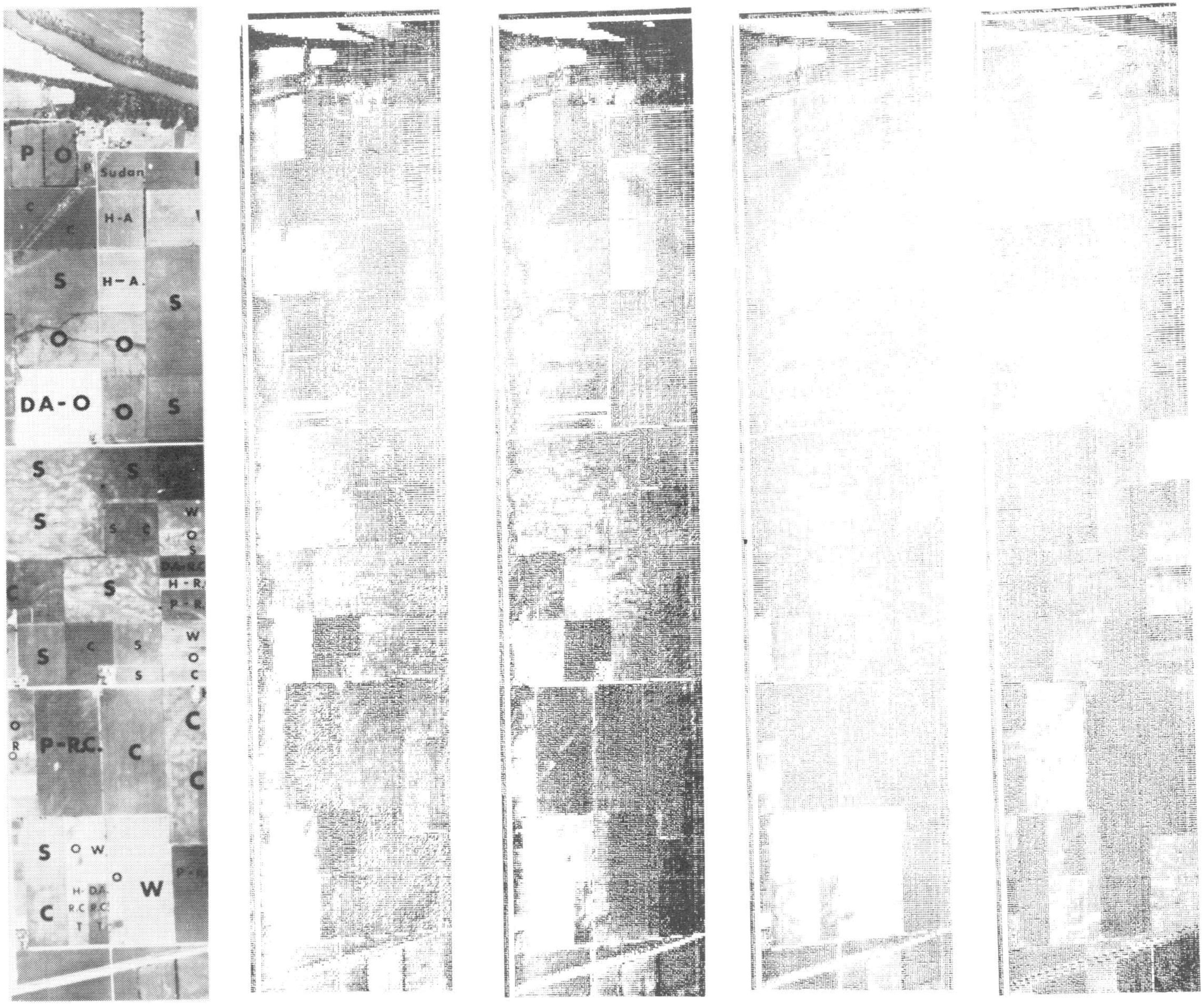


Figure 11. An Airborne Multispectral Scanner.



Air Photo .40-.44µm .52-.55µm .66-.72µm .72-.80µm

Figure 12. Data in Four Wavelengths

Figure 13. Spectral Pattern Recognition of Row Crops and Cereal Grains.

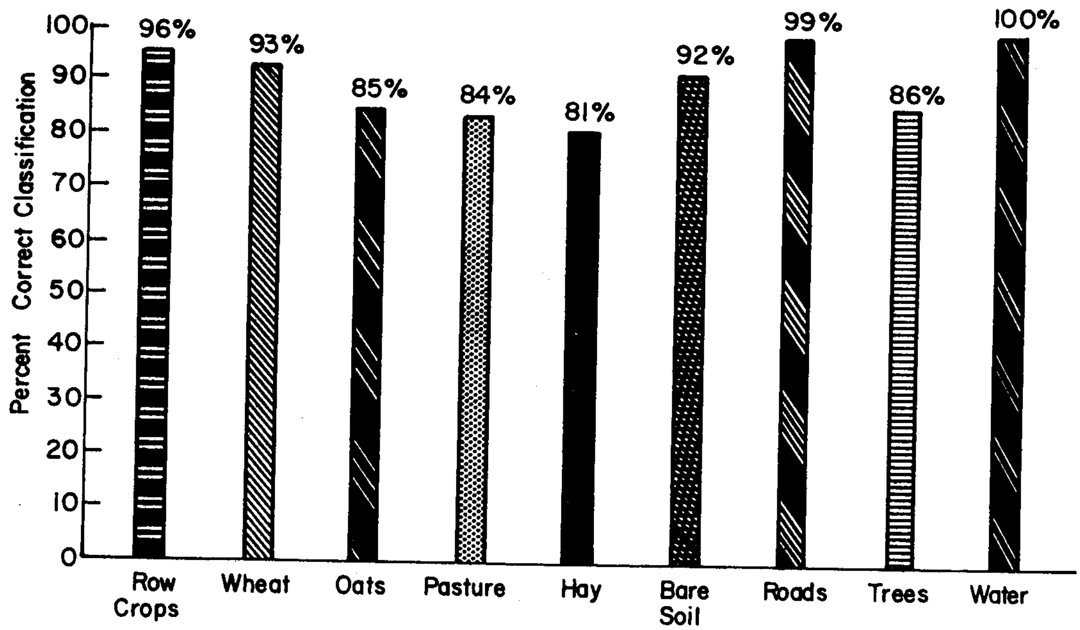


Figure 14. Classification Results for Test Samples.

15. (Left) Black and white print of Color Infrared Frame 3698A (Apollo 9) and (Right) Gray Scale Panchromatic Computer Printout of the Imperial Valley, California.

Figure 16. High Resolution Printout of a Section of the Digitized Apollo 9 Frame 3698A, Imperial Valley, California.

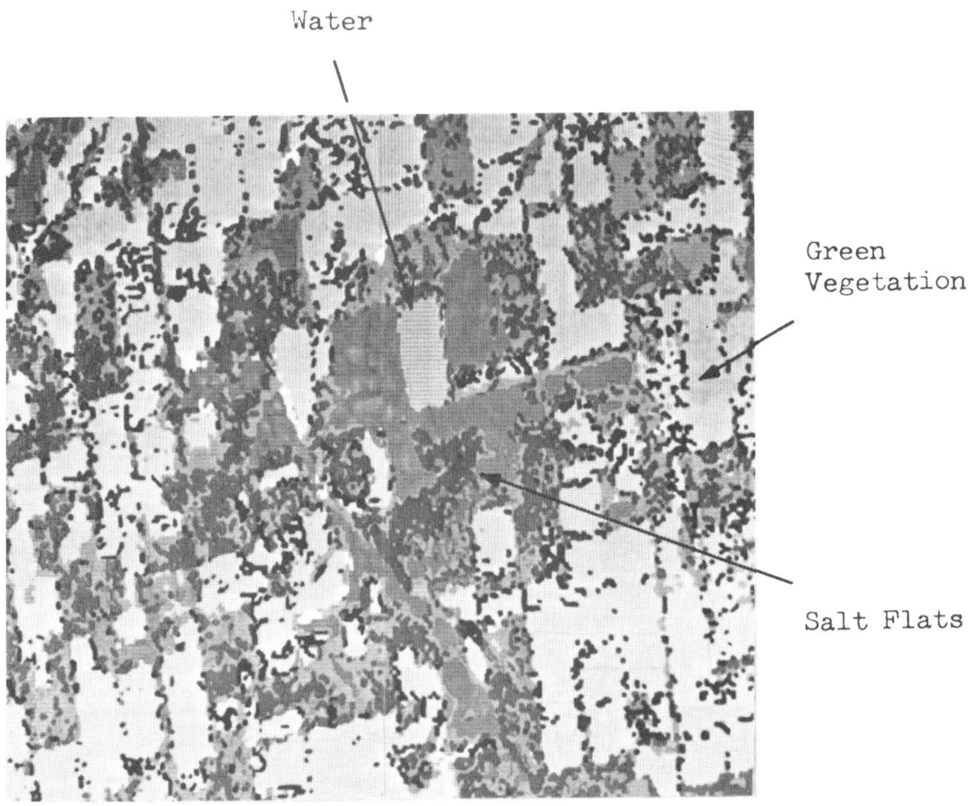
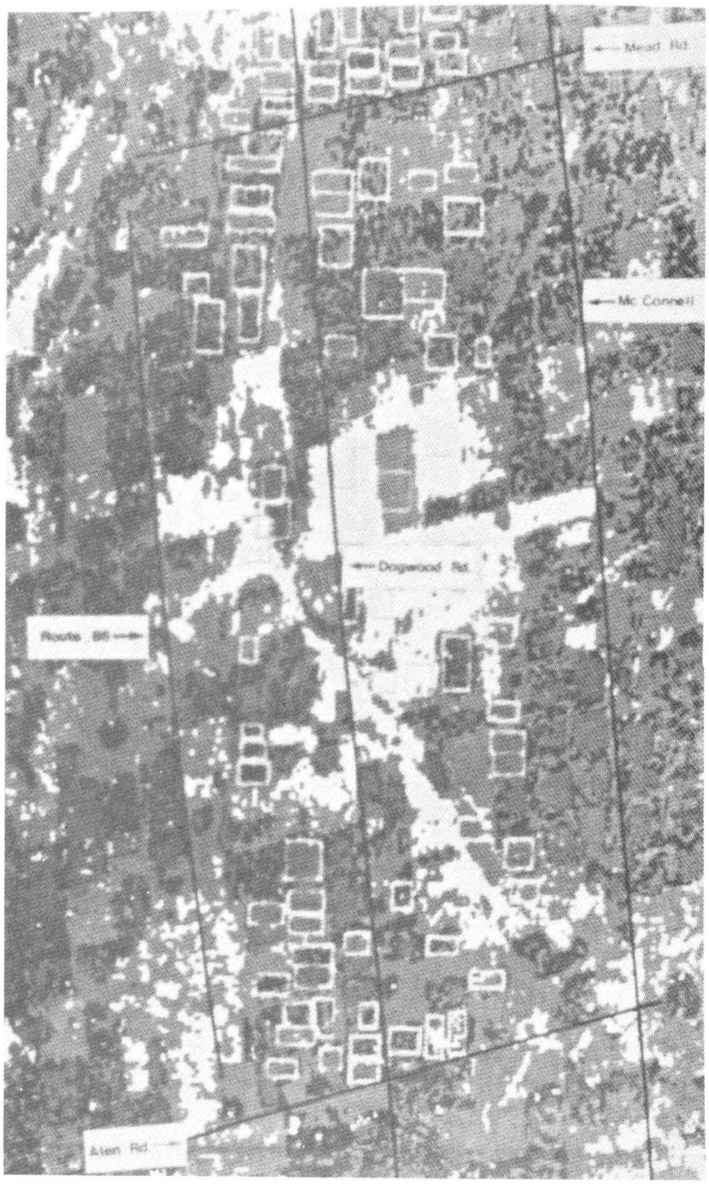


Figure 17. Classification of Apollo 9 Data Into Green Vegetation, Bare Soil, Salt Flats, and Water Classes (Bare Soil is Illustrated by All Other Colors).



Barley

Alfalfa

Water

Sugar Beets

Bare Soil

Salt Flat

Figure 18. Classification Results for Alfalfa, Barley, Sugar Beets, Bare Soil, Salt Flats and Water Classes.

Apollo 9 Computer Map of Imperial Valley Region

Clouds		Cloud Shadow	
Basalt		Vegetation	
Alluvium		Sand Dunes	

Figure 19.

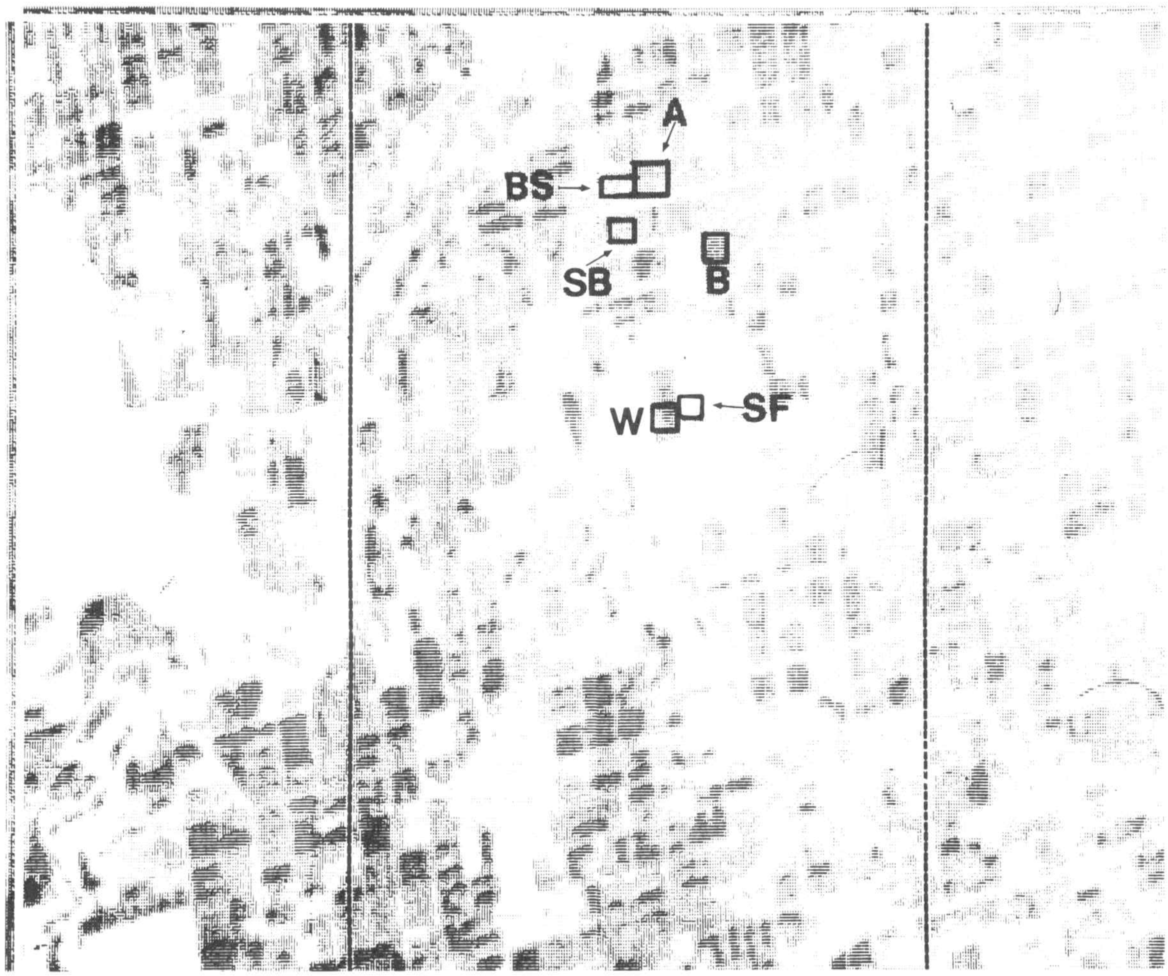


Figure 20. High Resolution Printout with Training Fields.

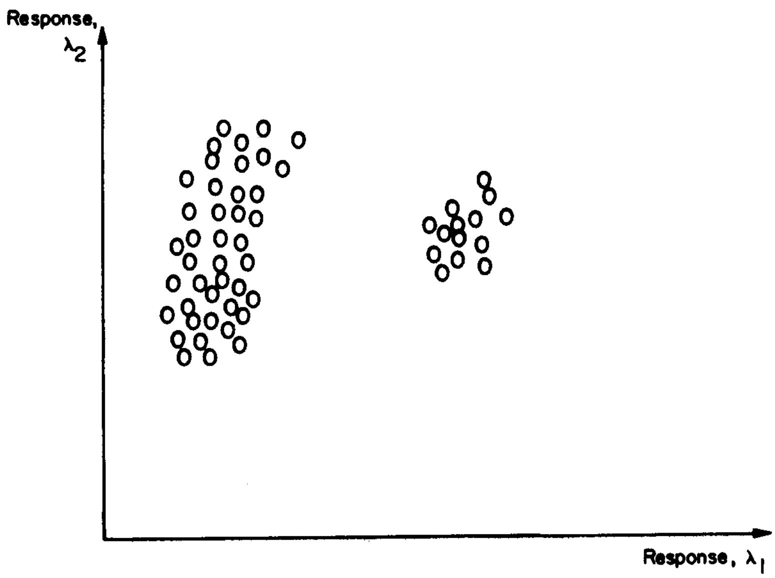


Figure 21. Samples in Two-Dimensional Feature Space.

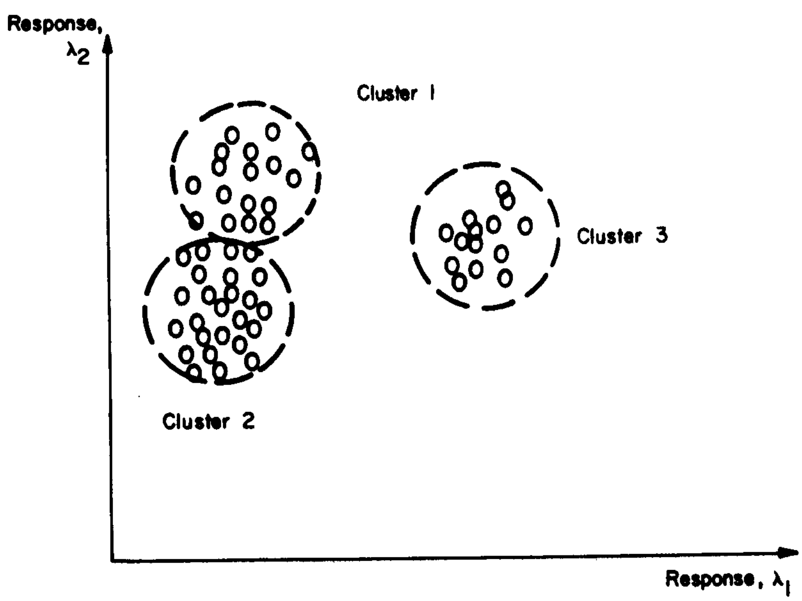
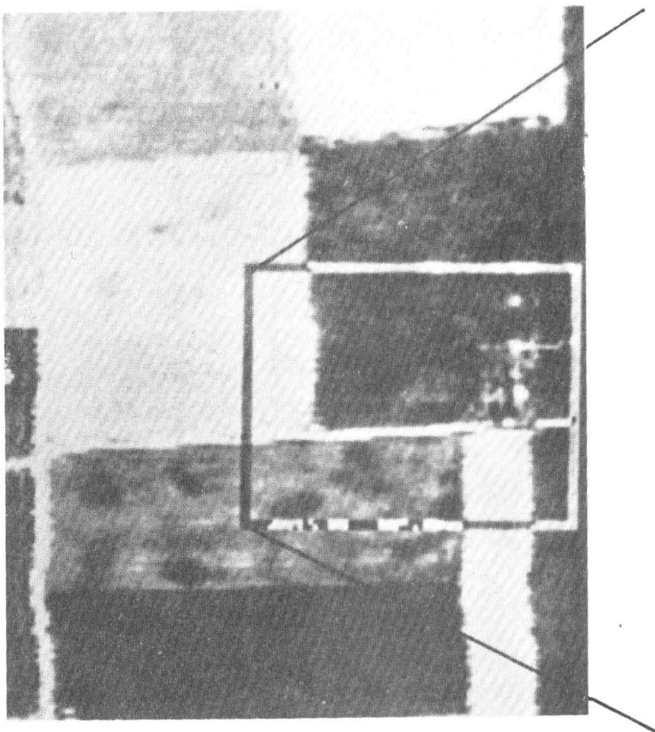


Figure 22. Clustering in Two-Dimensional Feature Space.



```
444432111111111221111332352325
444432111111121112212223553323
444442111111112122121212333323
444442111111112111112222223555
444443211111111111112222355555
444442111111111111122233555555
444442111111111111111235545555
444442111111111121112555555555
444442111111111122112255545555
444443111111111112223235555555
444442111111112323321253545555
444442111111122222311355555555
444443211111122111112353555555
444442111111121121113355333555
444434322333333334444455553355
444444444433233222334444444335
33323332222233323334444444355
332233332213332233234444444335
33322333221232323234444444333
333323221322333222224444444333
3333323222212232332234444444333
323222333322222233224444444333
322222333232222221134444444333
3222232223332222221134444444333
3222233223322222222124444444333
2222332223332322222234444444333
2232322222232222222334444444333
2223323222223211222234444444333
2232222222233222233223 4444444333
22222222222222232222234444444333
```

Figure 23. Clustered Data Using Four Spectral Bands.

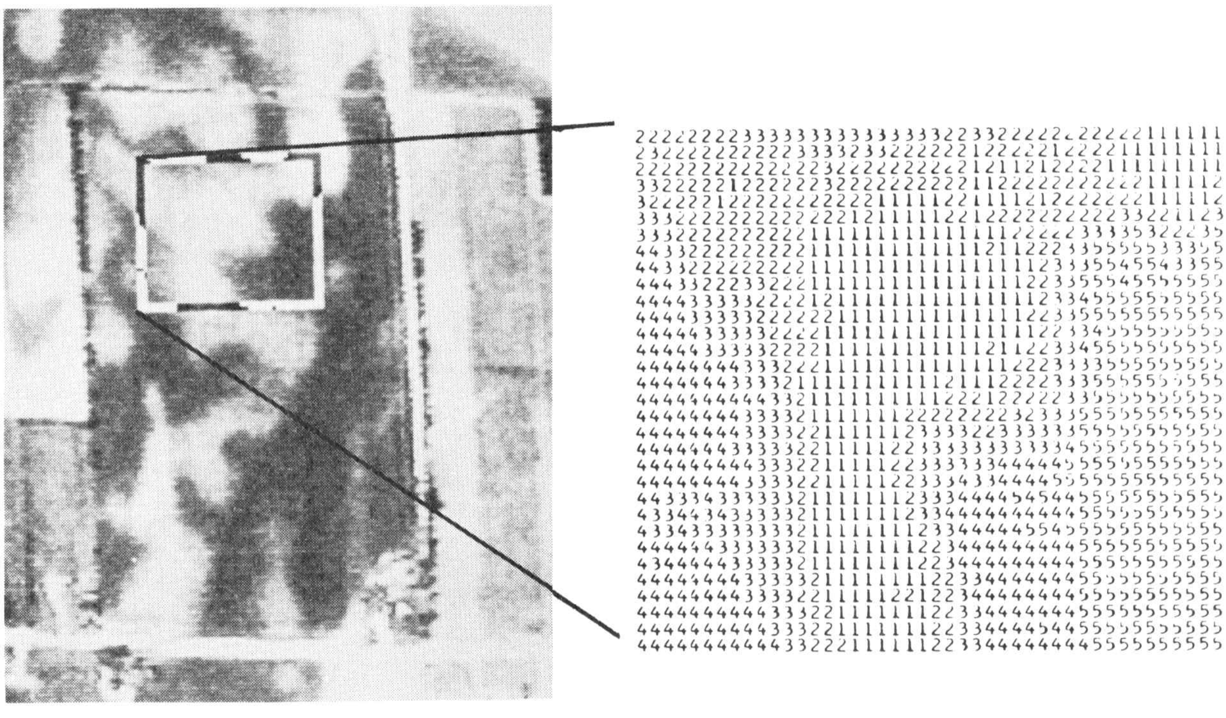


Figure 24. Clustered Data Using Four Spectral Bands.

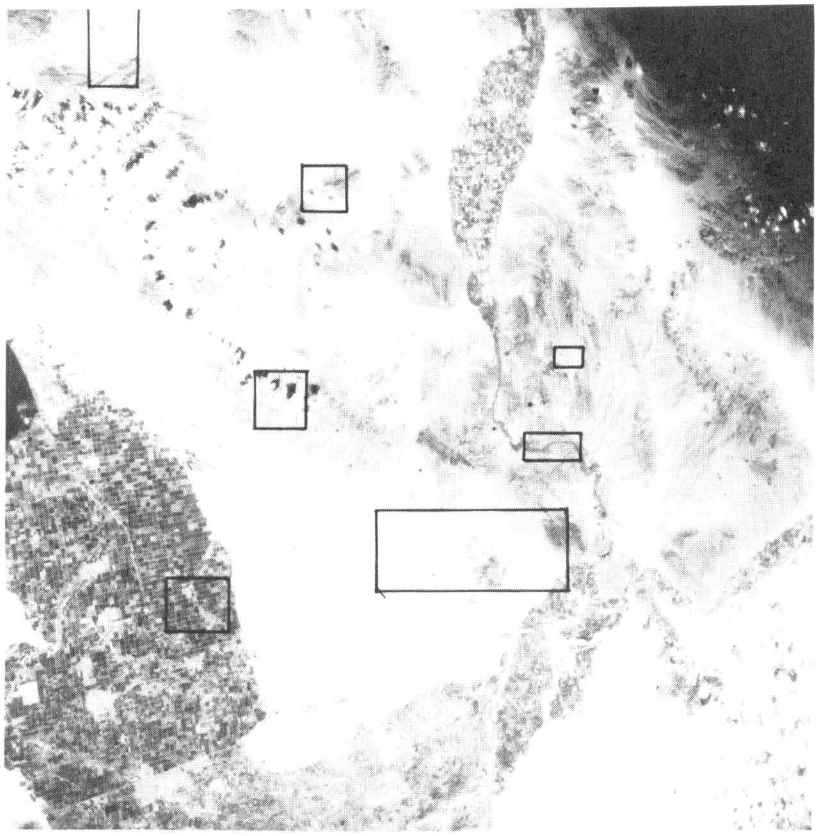


Figure 25. Cluster Fields for Machine Aided Training.

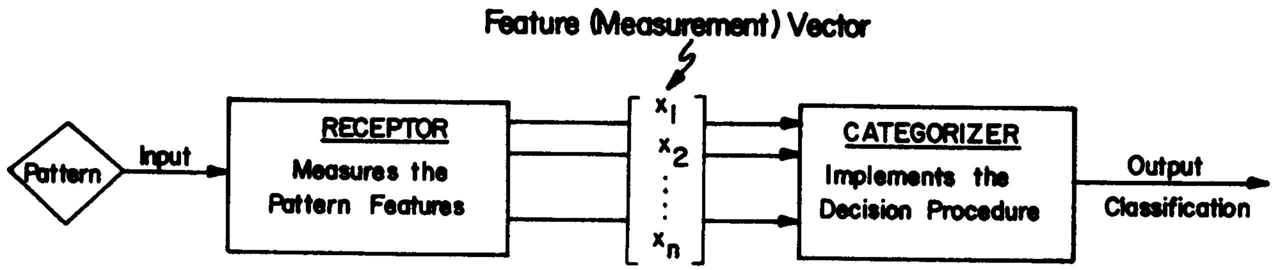


Figure 26. Block Diagram of a Pattern Recognition Device.

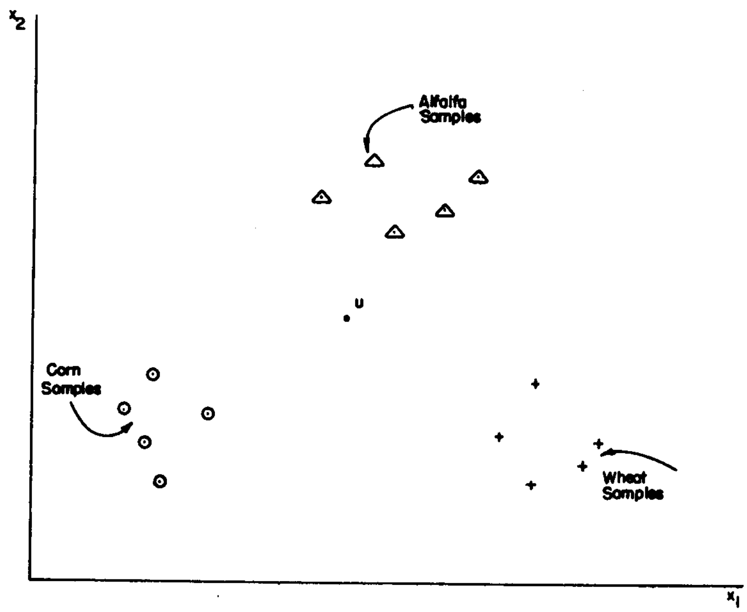


Figure 27. Samples in Two-Dimensional Feature Space.

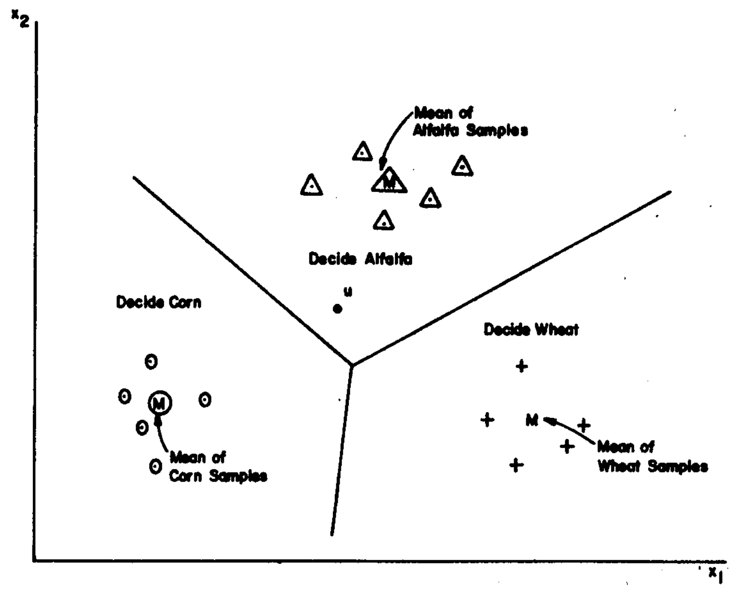


Figure 28. Decision Regions Based on Minimum Distance to Means.

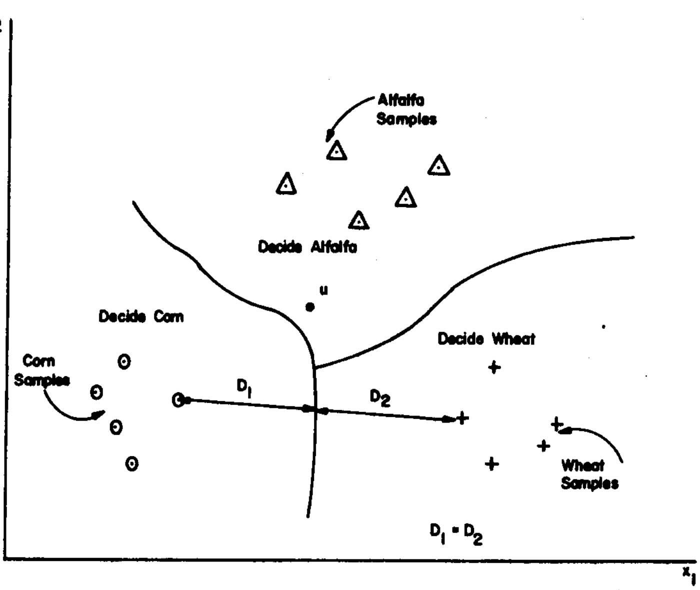


Figure 29. Decision Regions Based on Minimum Distance to Nearest Class.

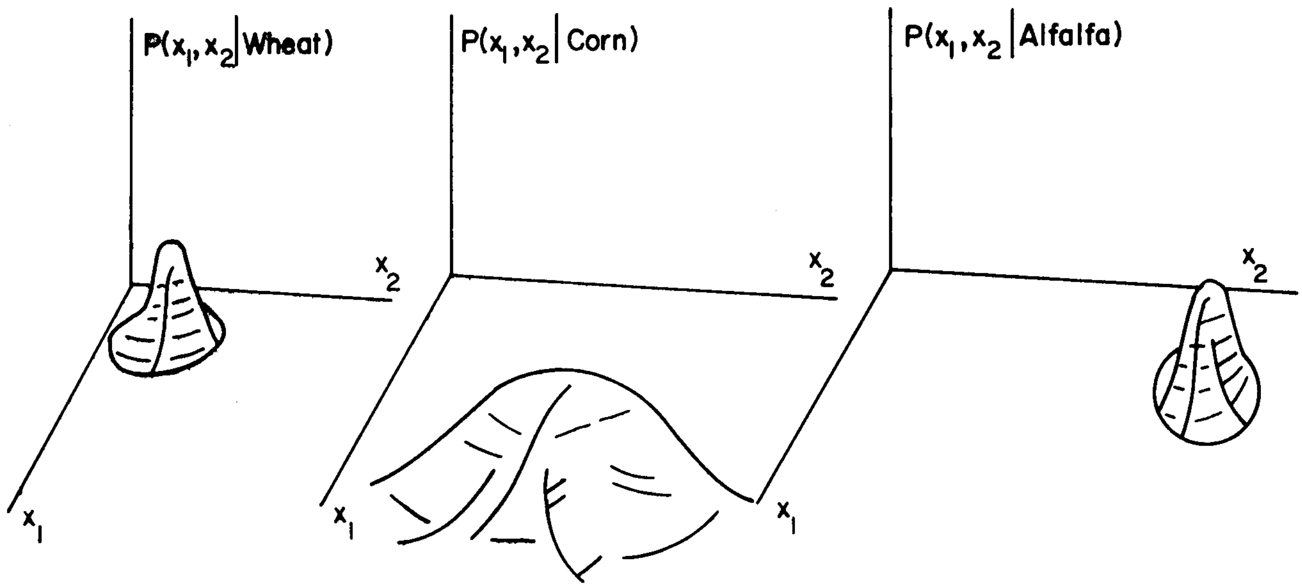


Figure 30. Probability Density Functions for Each Class.

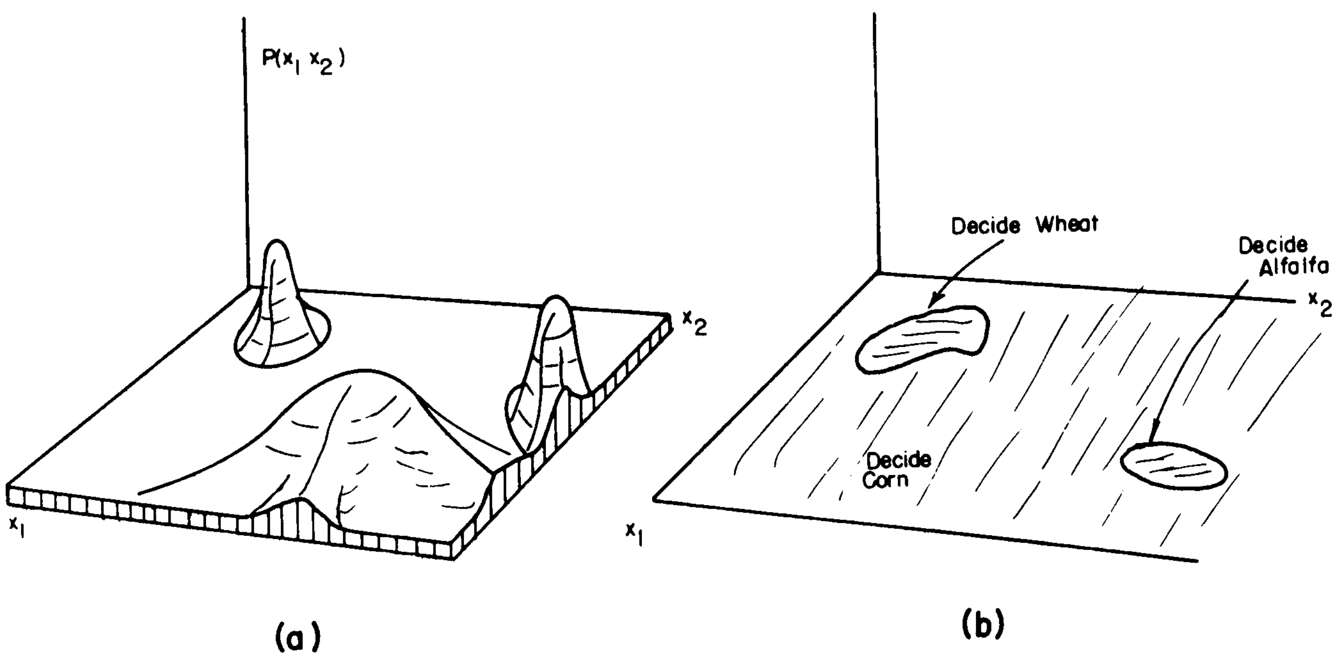


Figure 31. Density Functions (a) and Decision Regions for a Statistical Approach (b).

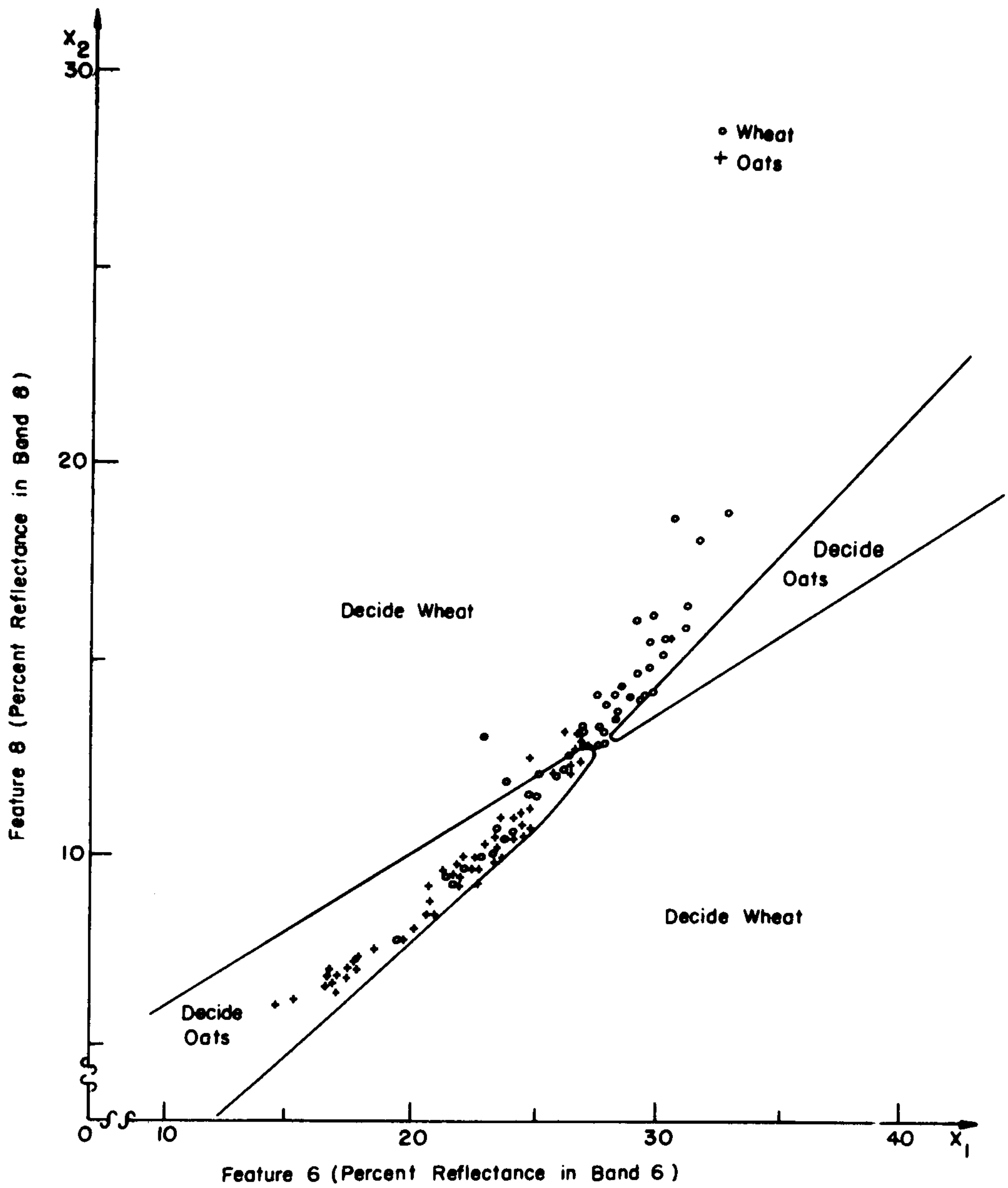


Figure 32. Feature Space: Bands 6 and 8.

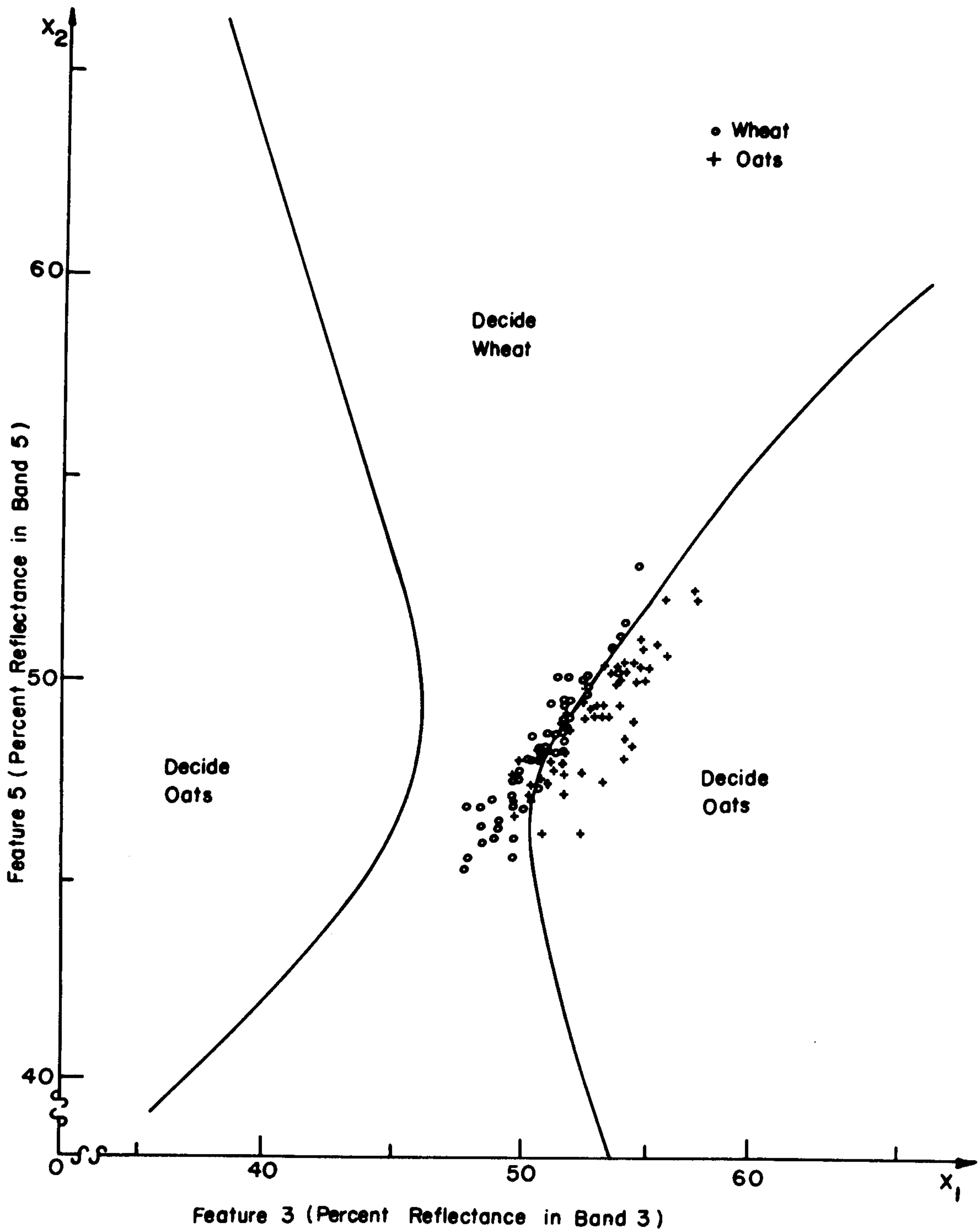


Figure 33. Feature Space: Bands 3 and 5.