VIEWING NOTES

# The Remote Sensing Information System

by David A. Landgrebe

VIDEOTAPE SERIES

Introduction to Quantitative Analysis of Remote Sensing Data

Inquiries about the VIDEOTAPE SERIES may be directed to:

CONTINUING EDUCATION ADMINISTRATION
116 Stewart Center
Purdue University
West Lafayette, Indiana 47907 (USA)



**A Continuing Education Activity of Purdue University**

## PREFACE

Remote sensing helps provide up-to-date information critical to the management of earth resources on a regional, national, and international level. As data collection and communication techniques are refined, as computational and analysis technologies evolve, and as sophisticated output devices become more economical, remote sensing will play an ever larger role as a source of information. The videotape you are about to see is one in a group of educational programs developed at Purdue University to keep scientists and administrators abreast of this rapidly evolving technology.

Purdue University has long been a leader in the development of remote sensing technology through the Laboratory for Applications of Remote Sensing (LARS). For over fifteen-years, the interdisciplinary staff, from the schools of agriculture, engineering, and science, has been responsible for much of the development in remote sensing and has gained world-wide recognition through these accomplishments. Today, activities at the Laboratory include both fundamental research and applications research, with additional major emphasis on training people in the use of quantitative remote sensing systems and developing educational materials to foster understanding of the technology.

Several key aspects of remote sensing technology are addressed in these videotapes, which make up the series Introduction to Quantitative Analysis of Remote Sensing Data. The tapes focus on such topics as scene radiation, data collection and preprocessing, and data analysis -- from both theoretical and applications perspectives -- and each from the perspective of a single scientist who is recognized as an expert in the field. In this way, when you view each tape, you are able to share personally in the insights and judgements offered on this increasingly important technology.

Videotape presentations in the series include the following titles:

The Role of Numerical Analysis in Forest Management by Roger M. Hoffer, professor of forestry.
The Role of Pattern Recognition in Remote Sensing by Philip H. Swain, associate professor of electrical engineering.
The Remote Sensing Information System by David A. Landgrebe, professor of electrical engineering and associate dean of engineering.
Correction and Enhancement of Digital Image Data by Paul E. Anuta, research engineer.
Spectral Properties of Soils by Marion F. Baumgardner, professor of Agronomy.

### Level and Prerequisites

The videotape series was prepared for graduate and advanced undergraduate students and professionals new to the field of remote sensing. Before viewing this tape, you should already have a basic understanding of remote sensing and its related terminology, such as can be gained through any of a number of introductory texts or through studying selected minicourses in the series Fundamentals of Remote Sensing (Purdue University, 1976, 1980).

Despite the mathematical and statistical nature of some aspects of the technology, the videotapes have been prepared with non-mathematical audiences in mind. Whenever possible mathematical relationships are illustrated graphically and described verbally. You may wish to consult the textbook Remote Sensing, the Quantitative Approach, edited by Philip H. Swain and Shirley M. Davis (McGraw-Hill Book Company, New York, 1978), for in-depth explanations of many of the concepts presented.

## About the Author

David A. Landgrebe is associate dean of engineering at Purdue University and director of the Engineering Experiment Station. From 1969 until 1981 he served as the director of Purdue's Laboratory for Applications of Remote Sensing while teaching in the School of Electrical Engineering and pursuing research in the area of signal representation, systems and data processing.

He has served on a number of National Academy of Sciences and NASA study groups and advisory committees and is associate editor of the journal Remote Sensing of the Environment.

He has been honored for his educational and research achievements through his selection as a recipient of the D.D. Ewing Best Teacher Award and the NASA Exceptional Scientific Achievement Medal for his work in machine analysis methods for remote sensing.

## Acknowledgements

Others at Purdue University who contributed to this videotaped program were James D. Russell, associate professor of education; Neil Sydor, producer-director; Sara Jane Coffman, instructional developer; and Susan L. Ferringer and Kathleen Z. Barash, visual designers. Music by Richard K. Thomas. Viewing notes were prepared by James Tilton.

Funds supporting the development of these videotaped programs and viewing notes were provided by the Continuing Education Administration, Purdue University. The authors express special appreciation to Mr. G.W. O'Brien for his generous support of this project.

Time: 25 minutes

# The Remote Sensing Information System

by David A. Landgrebe

## Synopsis

Since the earth's resources are finite, ways must be found to better manage the ones that now exist. It is now possible to use satellite-gathered data and computers to obtain the needed information on the status of many earth resources. The multispectral approach using a multidimensional feature space can provide discriminations in ground cover types which are not possible to make from imaging alone. Already feasible are discriminations between such classes of earth surface cover as some crop species, forest species mixes and various urban land use classes. Separation of more complex classes will become practical as more advanced sensors are placed in operation. This videotape provides a non-mathematical description of the multispectral technique and how it can be applied to obtain resource information. The technique is useful not only for current satellite data but even more so for future satellites with their more advanced sensors.

## Objectives

Upon completion of this videotape, you should be able to:

1. Give three examples of earth resources information obtainable through a combination of aerospace and computer technology.

2. Contrast the types of information that can be identified through visual inspection of Landsat imagery with the types of information that can be extracted by employing computer analysis.

3. Use a sketch to show the relationships between the four aspects (scene, receptor, classifier and results) of a computer pattern recognition system.

4. Describe the set of measurements used in the multispectral approach.

5. Explain why a statistical approach is generally desirable when analyzing multispectral scanner data. Name a curve that often describes the distribution of land cover class spectral responses and specify two important properties of this curve.

6. Briefly describe the maximum likelihood approach to computer classification and contrast it with the minimum distance to means approach. State when the later approach gives different results.

7. List the three main components of a remote sensing system. State how these components differ in terms of human control, complexity and amenability to optimization.

-- TURN THE PAGE --

-- BEGIN VIEWING TAPE --

## I. Introduction and Motivation

Serious problems facing our society include

- food,
- pollution, and
- energy.

Solving these problems requires efficient use of our existing finite resources through good management.

> "Good management requires good information, and the best vantage point to obtain it is from above the earth."

This space vantage point can be used through the combination of aerospace and computer technology to obtain needed information about such things as the condition of agriculture, the earth's resources, and the environment.

## II. Imagery from Aerospace Platforms (Landsat)

Simulated color infrared images can be generated from Landsat multispectral scanner data. Visual inspection of such imagery can identify such major features as:

- lakes,
- forested lands,
- croplands, and
- major cities.

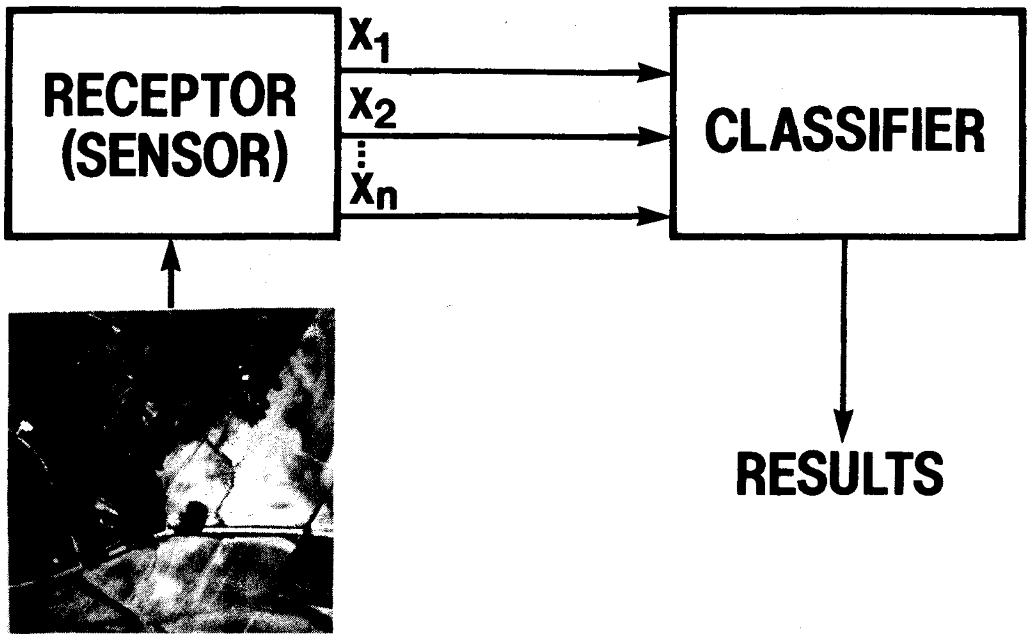
Computer analysis is required to extract more detailed information from the imagery, such as

- number of acres of a particular crop and probable yield,
- identification of soil erosion problems
- extent and condition of urban land use.

> "One approach to answer these questions is to use a technique called pattern recognition."

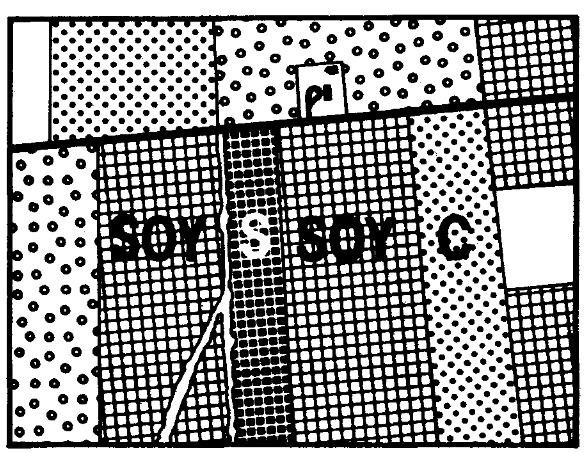
III. Pattern Recognition

"In pattern recognition, a series of measurements are taken of the scene by the sensor system. These measurements are submitted to a classification algorithm which assigns the particular measurements to one of several candidate classes."

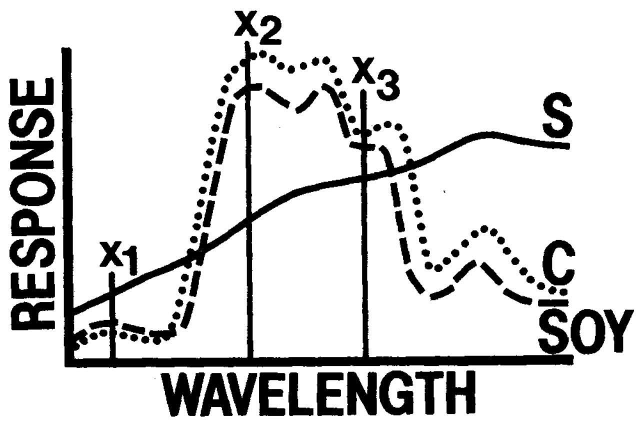



A. Multispectral Approach

One possible set of measurements exploited in pattern recognition is the spectral response in each of several wavelength bands. When this set of measurements is used, we say we are using the <u>multispectral approach</u> to pattern recognition. As an example of this approach, let's look at an idealized agricultural scene:



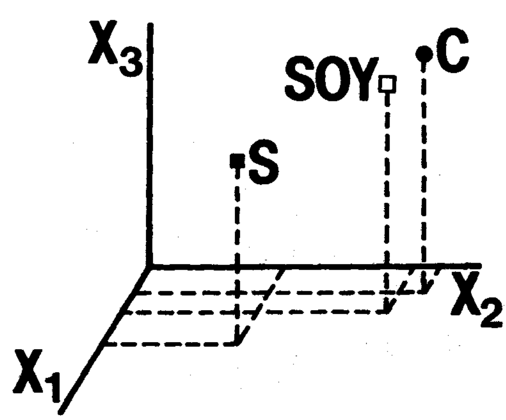
A sensing instrument can precisely measure how a particular ground cover reflects different amounts of energy at different wavelengths of visible light. The energy distribution for the three areas indicated above look like this:




These curves contain the information needed to discriminate between the three areas.

B. Discrimination Process

"The discrimination process can be accomplished by sampling these three responses at different wavelengths and imaging them to be plotted in a three-dimensional feature space."



A computer can be programmed to determine where a sample point occurs in such a space and, in principle, discriminate among types of surface cover on this basis.

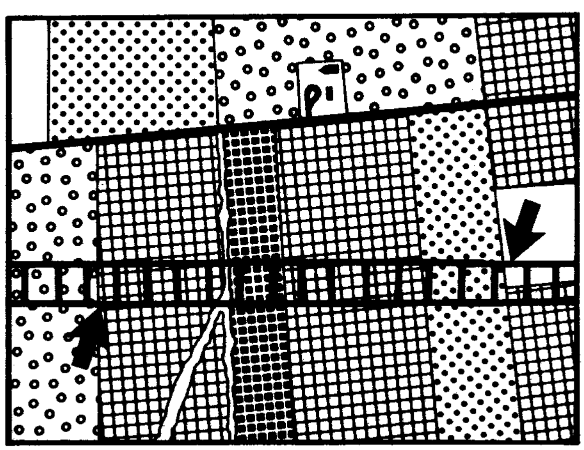
Higher dimensional spaces can be utilized by a computer even though we cannot draw such spaces.

Measurements made using non-visible wavelengths can also be used.

> "To discriminate among different materials, the only requirement is that measurements be chosen so the resulting points are separated from one another in the multidimensional space."

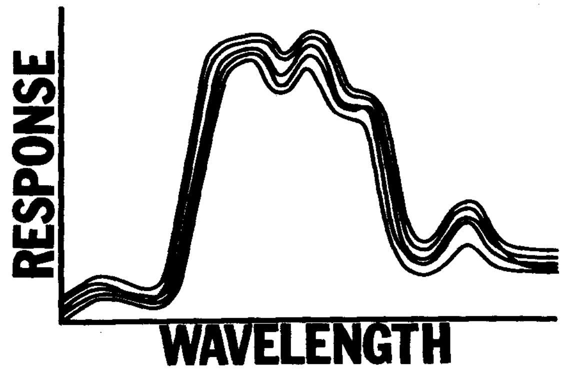
C. Statistical Variability of Spectral Responses

A "pixel" is an individual rectangular picture element that makes up an image in digital format. Pixels occurring wholly within a uniform area do not necessarily have an identical spectral response. A natural statistical variability occurs because of a large number of scene variables, environmental factors, and measurement circumstances.

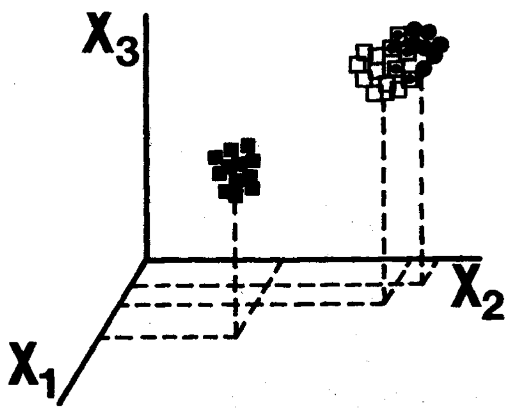


Another cause of statistical variability is that some pixels are actually a composite of two or more classes.

All of these sources of statistical variability combine together with the result that a given class is represented by a family of responses, called a statistical ensemble, as below.

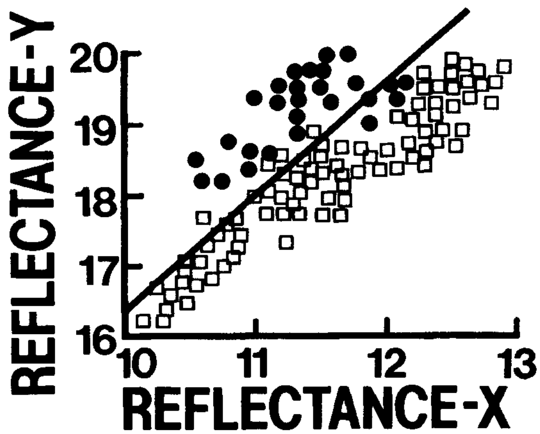



Thus, a more suitable sampling of the three classes mentioned earlier results in the following plot in three-dimensional feature space:




"It becomes apparent that a class exists not at a point in a feature space but in a region, and it is best modeled statistically."

Using a statistical model, a computer program can be written to assign data points in the feature space to an appropriate class. Even when the spectral responses of two classes overlap, they often can be satisfactorily separated as shown in the two-dimensional example below:



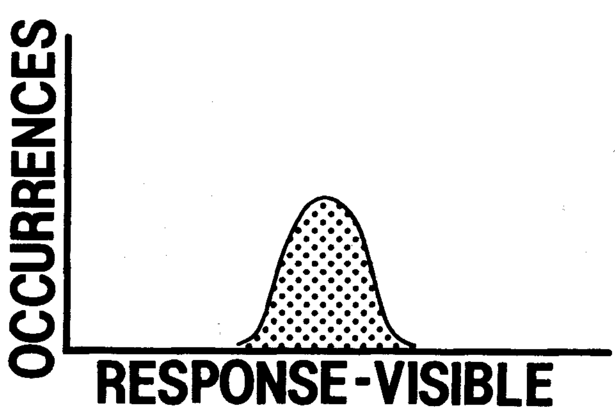

D. Statistical Characterization of Classes

We need not examine all of the possible spectral responses in a given geographic region when defining the spectral responses of the classes in the region.

> "We can collect a small set of samples which are representative and typical of the entire data set.... This set of samples is usually referred to as the training set.

1. Statistical Model of Spectral Response

Because of natural statistical variability, we know that spectral responses for a class are distributed over a range of values as below:



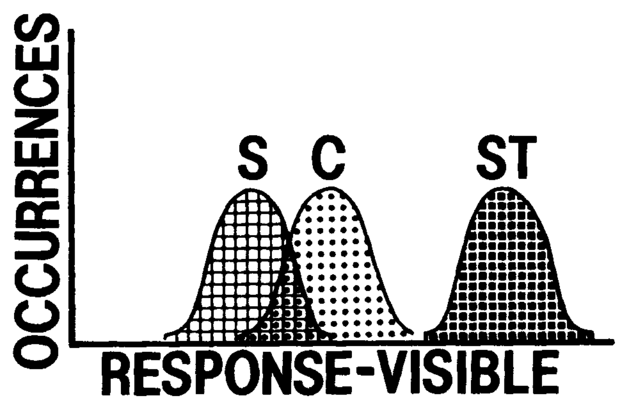

> "The distribution of such data often follows the bell-shaped or so-called Gaussian curve."

The Gaussian curve has a number of important properties:

- it is symmetrical about its mean (most frequently occuring) value, and
- the standard deviation (the relative size of deviation from the means) and the mean completely specify the entire distribution.

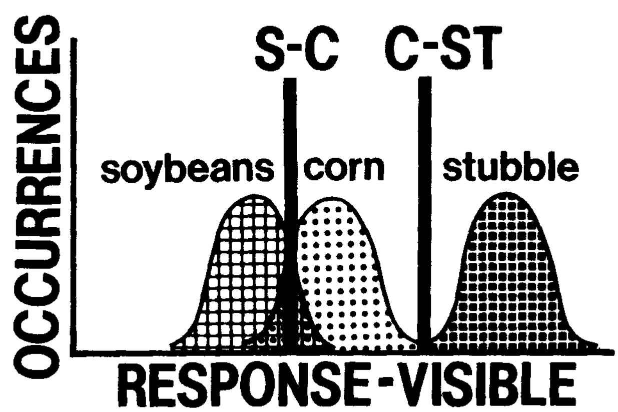
2. Maximum Likelihood Concept

The distributions corresponding to our earlier three spectral class example would appear as below:

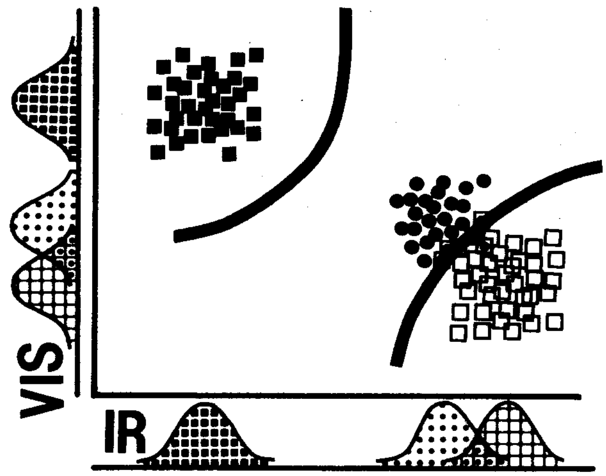



"The three curves are called probability density functions. Their value at any given x-value or response value gives the relative likelihood of occurrence of that response value."

Decision boundaries can now be drawn so that a pixel with a given response value will be assigned to it's most likely class:



The maximum likelihood concept can be readily generalized to a two-dimensional or even higher dimensional cases. Here is a two-dimensional example:




The maximum likelihood procedure is only one of several other ways to arrive at a classification assignment for each point in a feature space. Another frequently used scheme is the minimum distance to means procedure.

3. Minimum Distance to Means Procedure

In this procedure, the mean of each class is first determined. Then decision boundaries are drawn by plotting the locus of points equidistant from each pair of mean points.

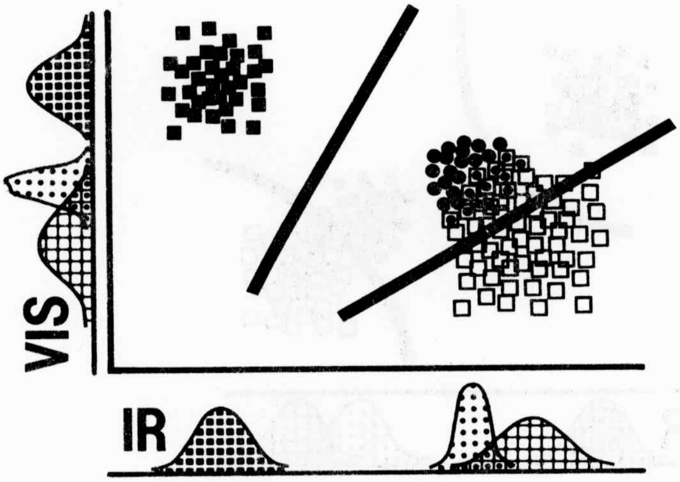
Advantages over the maximum likelihood case:

- less computer time (decision boundaries are linear)
- fewer training samples

Disadvantage:

- often less accurate

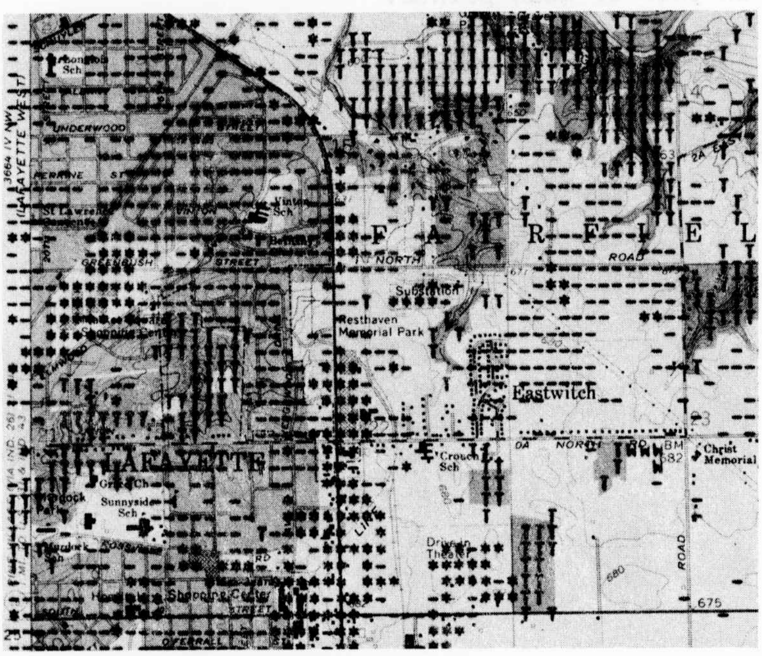
When class variances are significantly different, the minimum distance to means procedure is less accurate:




Linear decision boundaries produced by the minimum distance to means procedure.

The minimum distance to means procedure produces less accurate results because the decision boundary is placed too close to the class of larger variance.

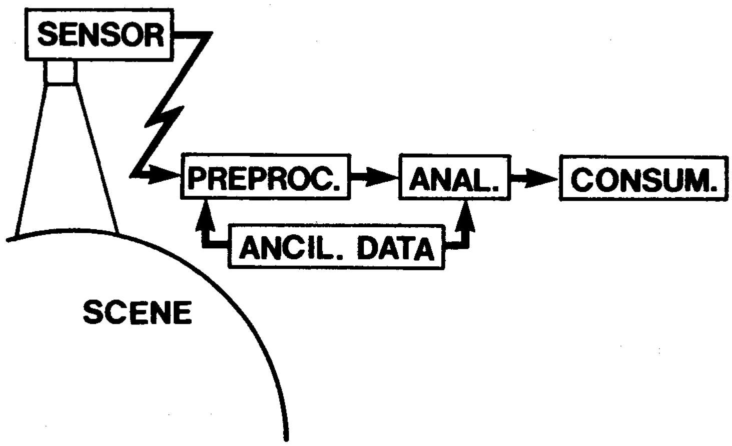
E. Classification of Real Data




| Symbol | Class |
| --- | --- |
| - | residential |
| * | commercial |
| T | trees |
| W | water |
| (blank) | agricultural |

The above is a classification of real data printed on a USGS map and the classification.

IV. Overall Perspective of the System



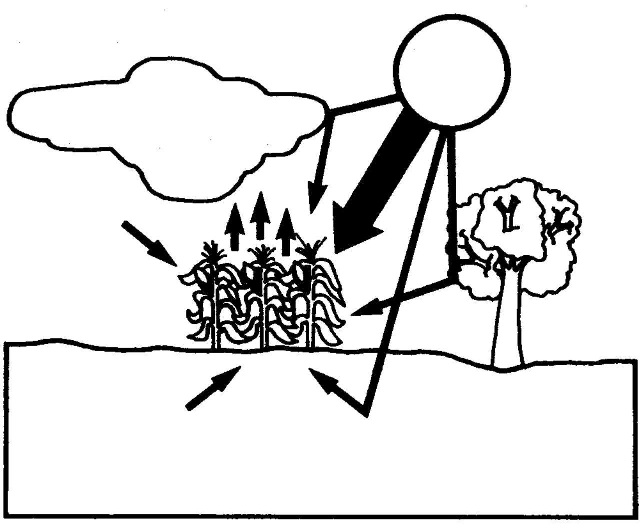

"This system can be logically divided into three distinctly different parts. The scene, the sensor system, and the processing system."

A. Scene

"The scene...consists of the sun, the earth surface and the atmosphere."

Characteristics of the scene:

- not under human control before or after sensor is built, and
- extremely complex.



The complexity of the scene leads to the statistical variability of the ground cover classes.

B. Sensor System

The function of the sensor system is to gather data about the scene.

> "You can view it as a kind of transducer which transforms the desired information from a radiometric form to the form of an electric signal."

Characteristics of the sensor system:

under human control during the <u>design</u> phase,
not under human control during operation, and
tremendous cost forces a general design so as to serve a wide range of users.

C. Processing System

> "... the processing system <u>is</u> under human control both in the design and the operations phases, and it can usually be optimized with regard to specific applications."

## V. <u>Summary</u>

We discussed

> "... how the understanding and use of spectral characteristics coupled with the preciseness of the computer now enables us to handle large amounts of data and to identify difference classes of earth surface cover...."
>
> "... why it's desirable to model or picture in our mind scene measurements in terms of a probablistic or statistical model...."
>
> "... the fundamental concept of multispectral data and computer analysis which can be used to partition an n-dimensional feature space into non-overlapping regions..."
>
> "... the total information system composed of the scene, the sensor and the processing."

SELF-CHECK QUESTIONS

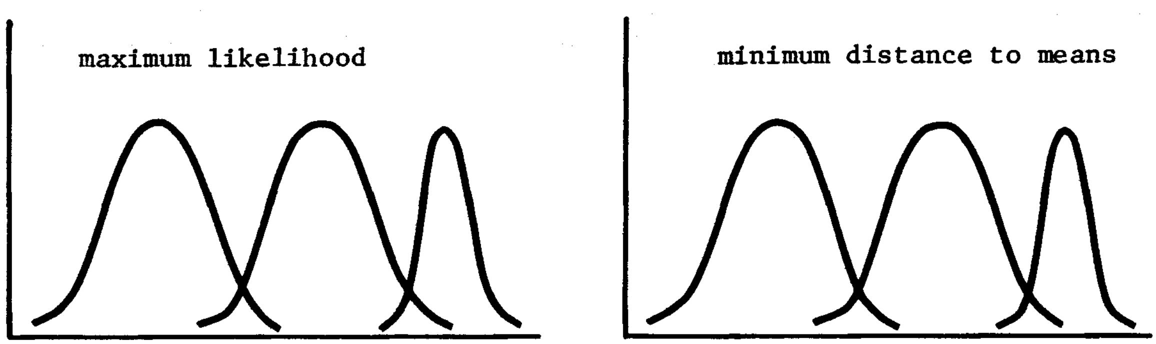
1. Give three illustrations of earth resources information which can be obtained using a combination of aerospace and computer technology.

2. Sketch the relationships between the four basic aspects (scene, receptor, classifier and results) of a computer pattern recognition system.

3. Describe the set of measurements used in the multispectral approach.

4. What are two important properties of the Gaussian distribution for statistical characterization of land cover classes?

5. Draw decision boundaries for the three classes shown using both the maximum likelihood and minimum distance to means procedures:




When will the two procedures give significantly different results?

CHALLENGE QUESTIONS

1. Lakes are one feature that can be identified through visual inspection of Landsat imagery. What are three other features that can be identified through visual inspection? Contrast these features with the types of information that can be extracted from Landsat data by employing computer analysis.

2. Explain why a measurement from a particular cover type at one location may differ slightly from another measurement of the same cover type at a different location in the same data set. How does this affect the way remote sensing data is analyzed?

3. What are the three main components of a remote sensing system? How do these components differ in terms of human control, complexity, and amenability to optimization?

FURTHER READING

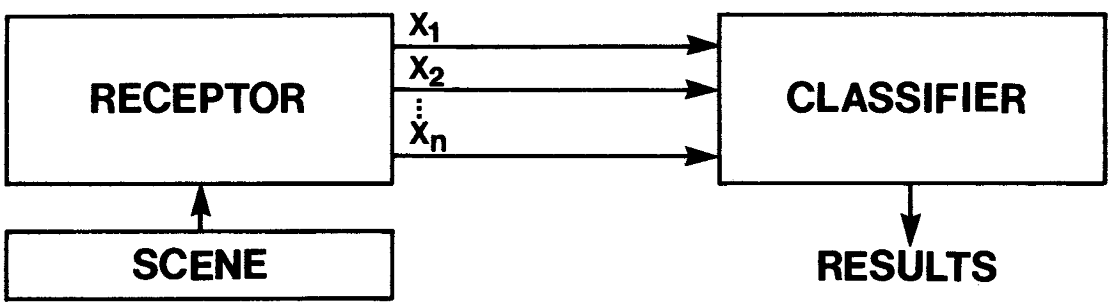
Swain, P.H. and S.M. Davis, eds. 1978. Remote Sensing: The Quantitative Approach, McGraw-Hill Book Company, New York. Pages 1-20, 136-159 and 336-374.

Lintz, J., Jr. and D.S. Simonett, eds. 1976. Remote Sensing of Environment, Addison-Wesley Publishing Company, Inc., Reading, Mass. Pages 349-370.

Landgrebe, D.A. 1981. "Analysis Technology for Land Remote Sensing." Proceedings of the IEEE, Vol. 69, No. 5, May 1981, Pages 628-642.

ANSWERS TO SELF-CHECK QUESTIONS

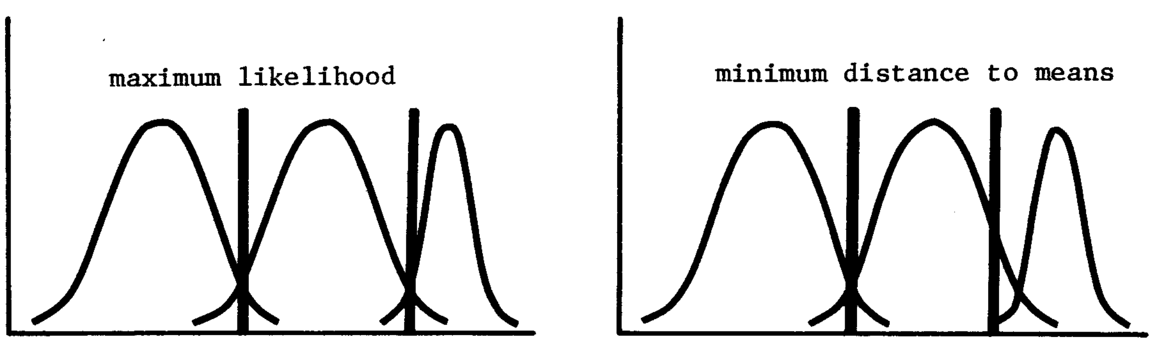
1. Information about such things as the condition of agriculture, the earth's resources, and the environment can be obtained using a combination of aerospace and computer technology.

2. 


3. The set of measurements used in the multispectral approach are measurements of the spectral response in each of several spectral bands.

4. Two important properties of the Gaussian distribution curve are:

   it is symmetrical about its mean value, and
   the standard deviation and the mean completely specify the entire distribution.

5. 


The results produced by these two procedures will be significantly different if the classes do not have similar standard deviations.