LARS Information Note 080372

ADP of Multispectral Scanner Data for Land Use Mapping

by
Roger M. Hoffer

The Laboratory for Applications of Remote Sensing

Purdue University, West Lafayette, Indiana

1973

ADP of Multispectral Scanner Data for Land Use Mapping[1]

by

Roger M. Hoffer[2]

Introduction

The need has long existed for a capability to obtain reliable information over large geographic areas in a timely manner. Such a need is present in many discipline areas. The past few years have seen a marked increase in the awareness of our needs for accurate land use planning, and the accompanying realization that our existing capability to obtain accurate and up-to-date land use information is not adequate to meet society's current and future needs. Because of the very rapid changes in the condition of agricultural crops and the influence of crop yield predictions on the world market, the need for accurate, timely information is particularly acute in agricultural information systems. It is for these reasons that, as remote sensing technology had developed over the past few years, the potentials for using this technology have received wide-spread attention.

In examining the possibilities for using remote sensing to help obtain accurate and timely information involving land use mapping and agricultural resources, one must review the advantages and disadvantages of the various remote sensing instrumentation and analysis techniques, so that the user needs can be met with the most efficient system possible.

Use of Multispectral Scanner Data

Land use mapping and agricultural information systems both require data from large geographic areas. It is rather obvious that high flying aircraft or satellites can collect

1) Invited paper presented at the 2nd UNESCO/IGU Symposium on Geographical Information Systems, Ottawa, Canada, August 1-9, 1972. This research was supported by NASA Grant NGL 15-005-112.

2) Associate Professor, Department of Forestry & Conservation and also Leader, Ecosystems Research Programs, Laboratory for Applications of Remote Sensing (LARS), Purdue University, West Lafayette, Indiana, 47906.

enormous quantities of data over vast geographic areas in a relatively short period of time. Such masses of data can be collected with a variety of sensor systems, each of which has its own particular advantages as well as disadvantages. However the reduction of such data into useful information appears to be a rather difficult problem, particularly if the information is required in a very short time frame. Therefore, a key factor developing a system which best meets the user needs appears to be the data reduction techniques.

One such technique that has been developed involves the use of digital computers and the application of pattern recognition theory to multispectral scanner data obtained by aircraft or satellites. This technique was conceived and has been developed by the Laboratory for Applications of Remote Sensing (LARS) at Purdue University to enable machine-assisted mapping and tabulation of earth surface features over large geographic areas in a timely manner (3, 4).

Multispectral scanner data has been primarily utilized because it can be easily quantified and subsequently processed by digital computer techniques, and also because the data format is ideally suited for pattern recognition analysis. In essence, the multispectral scanner allows one to measure the amount of energy reflected or emitted from a particular area on the ground (as "seen" by the instantaneous field of view of the scanner system), and the amplitude of the signals in the various wavelength bands can then be developed into a measurement vector, or what has frequently been referred to as a "spectral signature". (See Figures 1, 2, and 3). Such a measurement vector quantitatively defines the spectral characteristics of the material in that particular area on the ground. A key factor then, involves development of a technique whereby one can train the computer to recognize the spectral characteristics of all of the earth surface features of interest to the user (as contained in these measurement vectors) and to accurately identify each of these materials in the entire set of data obtained.

The essence of such automatic processing of remote sensor data involves a man/machine interaction, whereby the man will train the computer (utilizing data collected over a limited geographic area), and then the computer will continue to map and analyze data collected over a large geographic area at a much faster rate than would be possible for the man if he were using normal image interpretation techniques.

Considerations in Automatic Data Processing

There are a number of key elements involved in the use of ADP for analysis of multispectral scanner data. One of the first questions encountered is: "What type of categories or classes of material should the computer be trained to recognize?" Basically there are two conditions which must be met by each class involved in an analysis of remote sensor data using pattern recognition techniques:

- The class must be spectrally separable from all other classes.
- The class must be of interest to the user or have informational value.

In working with multispectral scanner data, one soon finds that often the classes of informational value cannot be spectrally separated at certain times of the year. One reason for this is that various species of green vegetation have very similar spectral characteristics even though their morphological characteristics may be quite different. The need for a class to be both separable and have informational value therefore leads to two quite different approaches in training of the computer system.

The first approach involves the designation to the computer system of locations in the data where the earth surface material is known and has informational value (i.e. a certain location in the data contains corn, another soybeans, another wheat, water etc.) This "training sample" approach has been used quite effectively for agricultural and land use mapping analysis (1, 4). One must be constantly alert, however, to recognize situations in which two classes may be of great informational value and interest, but which cannot be separated spectrally. In such situations these classes can be combined if there is a sound, logical reason for doing so (e. g. wheat and oats could be combined into a class defined as small grains, but wheat and soybeans could not be logically grouped).

A second approach to training the computer system involves "clustering" of the data, in which case one simply designates an area or a number of different areas in the data to the computer, and then designates the number of spectral classes into which this data should be divided. The computer can then be programmed to designate which areas in the data belong to which spectral classes and proceed to classify and map these results. The user must then simply relate the classification output to known surface observation data, and in essence determine what materials actually represent each of the different spectral classes (e. g. spectral Class 1 is wheat, Class 2 is bare soil, etc.) The difficulty with this system involves the determination of the number of spectral classes present and the fact that the classes of most interest often have subtle spectral differences while many of the other classes present in the data may be easily separated spectrally but be of little informational value or interest to the user. Experience at LARS has indicated that a combination of the two systems seems to be the most satisfactory and most effective procedure to follow.

Another subject that often arises in ADP of remote sensor data involves two rather different concepts used in the "training sample" approach discussed above. These two concepts could be termed the "data bank" and the "extrapolation mode" for obtaining data to be used in the training procedure. In the "data bank" approach, the concept is to have many sets of spectral signatures available in computer storage. Such a "data bank" would include signatures for all different cover types of interest. When a new set of remote sensor data is obtained, one simply selects the signatures for the cover types of interest from the existing "data bank", and, using these as training samples, classification. The advantage is a great reduction in the amount of time required to train the classification processor as to the characteristics of the spectral signatures involved in each new set of data. Such a system would be highly desirable and, in theory, could work under certain selected sets of conditions. However, in the real world situation, such a concept does not appear to be practical in most cases. The reasons for this involve the temporal effects upon the spectral characteristics of the earth surface features which are measured and which one is attempting to classify. Many studies have shown that the spectral characteristics of different types of vegetation change drastically as a function of time, both for a given growing season and from one year to the next. Plant maturity causes a nearly continuous change in spectral characteristics of agricultural crops. For instance, what is often measured early in the growing season is the amount of bare soil in the proportion to the amount of green vegetation present in the instantaneous field of view of the scanner, rather than a unique spectral signature of the different crop species. Since the farmer must depend upon weather conditions to govern the date of planting for the various crop species, the heights of the corn (and therefore the canopy coverage and the vegetation/soil relationship) will be quite different from one year to the next as of the same date on the calendar. The weather conditions governing the germination and rate of growth of the vegetation after planting are also important, and will cause considerable variation in spectral response throughout the growing season and from one growing season to the next. There are also difficulties in determining differences in atmospheric attenuation from one flight mission to another, and in being able to calibrate sensor data to a high degree of radiometric precision (2).

In view of such problems of variation from one flight mission to the next, it would seem that an extrapolation mode is much more logical for obtaining spectral data with which to train the computer for the classification task. In this case one simply uses the data collected under the existing atmospheric conditions and conditions of plant maturity, stress, growth, etc., abstracts the training samples from this data and then proceeds with the analysis task.

A Computer-Aided Analysis System for Remote Sensor Data

The following paragraphs describe the primary analysis system developed by the Laboratory for Applications of Remote Sensing, as an illustration of one technique that has been developed to allow automatic data processing of remote sensor data. Perhaps one should start by pointing out that the term "automatic data processing" is somewhat of a misnomer, in that there is a definite requirement for a very close man-machine interaction throughout the analysis sequence. Therefore, it would be much more correct to call it a "computer-aided analysis" of remote sensor data.

A digital computer system has been used at LARS because of the ease and flexibility in developing the system as a research tool and in interfacing with the multispectral scanner data. Experience has shown that a general purpose digital computer has proven very effective in these developmental phases of an analysis system. However, it appears probable that in an operational system of the future, a general purpose digital computer would not be used, but rather, a special purpose digital computer (which can obtain near real time computational speed) or possibly a hybrid analog/digital system would be more likely to be utilized.

One of the first requirements in designing an analysis system involved development of a capability for interfacing with the multispectral scanner data contained on the magnetic tapes. Since the data is originally collected by the scanner system in an analog mode, it must first be digitized for use with a digital computer system. A procedure was thus developed to assign a number of each scan line of data collected, and then to assign a sample designation to each digitized data point along the individual scan lines. Thus an X, Y coordinate system could be developed for any set of multispectral scanner data or digitized photography available for analysis.

Image Display

To most effectively interface with the digitized data, a map-like display of the scanner data in one or more wavelength bands is necessary, so that the user can work with the data when it is in an image format. At present, we accomplish this using one or two techniques. First is an alphanumeric printout in which the individual elements in the digitized data are displayed using a standard line printer, and utilizing different alphanumeric symbols to represent certain levels of radiance measurements. An example of this is shown in Figure 4. The researcher may then specify X, Y coordinates for particular areas of interest to him in his analysis task, such as training areas, test areas, portions of the data to be classified, etc. The details of the programs and how they function in order to obtain this type of printout are contained in the literature (4) and so will not be repeated in this paper.

A second method of displaying the imagery is through the use of a special purpose digital display unit as shown in Figure 5. This device presents a cathode ray tube (CRT) display of individual wavelength bands of imagery. Through the use of a light-pen the researcher may then select the areas of interest (Figure 6) and the X, Y coordinates so designated are automatically punched into a deck of data cards. It should be noted that the X, Y coordinates defined for a particular area using one wavelength band of data are equally valid for all other wavelength bands available on the data tape (perhaps as many as 24 wavelength bands with a new NASA scanner system).

Statistics Processor

The system which we developed to actually analyze multispectral scanner data is called LARSYSAA and is subdivided into four sections. These include 1) statistical analysis, 2) wavelength band or "feature" selection, 3) pattern classification, and 4) results display (Figure 7).

The LARSYSAA statistics processor can provide several types of outputs for use in the analysis of the data. These include histograms of individual fields or groups of fields which have been designated a class (Figure 8a & b), tables showing the means and standard deviations for data in individual fields or classes and correlation matrix tables (Figure 9), multispectral response graphs (Figure 10), and statistics data decks used in the classification. The multispectral response graphs represent the relative amplitude of the reflected or emitted energy in each of the wavelength bands involved in the analysis. The mean plus or minus one standard deviation is plotted in a linear format, thereby causing the length of the line to give some indication of the statistical quality of the data in that wavelength band. By comparing several different cover types in each of the various wavelength bands involved, one is able to obtain some indication of relative reflectance or emission for the different materials involved, as a function of wavelength. It should be noted that at this point in the data processing there is no correlation indicated between wavelength bands.

Wavelength Band Selection

Experience has showed that when working with many wavelength bands and a large number of categories to be classified, maximum overall accuracy in classification can be obtained by utilizing all wavelength bands available. However, this requires a marked increase in computer time. Frequently, only a slight increase (2 or 3%) in accuracy of classification is obtained when using 12 wavelength bands as compared to only 4 or 5 wavelength bands. It is therefore frequently desirable to reduce computer time by utilizing only 4 or 5 wavelength bands in the classification procedure. The question becomes: "which combination of only four or five wavelength bands would be the optimum subset to utilize in classification?" (Out of 12 wavelength bands available, there are 485 possible combinations of 4 wavelength bands.) To aid the researcher in quantitatively selecting the optimal subset, a selection processor has been devised which measures the separability between all possible pair-wise combinations of cover type materials and then assesses the average and minimum separability for all possible combination of wavelength bands (Figure 11). Thus, the researcher is able to obtain a indication of the optimum combination of wavelength bands to utilize in any particular classification task.

Classification Processor

The classification itself utilizes a pattern recognition algorithm involving a Gaussian maximum likelihood scheme. This is a relatively simple pattern recognition algorithm and has been used successfully at LARS by many different researchers. As indicated previously, there are a large number of pattern recognition algorithms which could be utilized, but this one has been found to be adequate for analyzing many different sets of data collected under a variety of conditions and being analyzed for a large number of potential user applications, including land use analysis, agricultural species and soils mapping, forest cover mapping, geological and hydrological features analysis, and others. In the classification, each resolution element or data point is assigned to one of the classes designated by the analyst, and these results are then stored on magnetic tape. The researcher can then display these results in a variety of ways.

Results Display

Two major results display formats are used. The first involves an alphanumeric map-like printout on the line printer, in which the user has selected symbols for each of the different classes of interest, such as C for corn, S for soybeans, F for forest, W for water, etc. (Figure 12). He may then compare these classification results to aerial photography or maps indicating known agricultural situations or land uses and determine in a qualitative way the accuracy of the classification results. Problem arises, however, when it becomes desirable to attempt the classification again with a different combination of wavelength bands or spectral classes. For this reason, a more quantitative evaluation procedure was devised that utilizes "test areas". In this procedure, the analyst designated (at some time before the actual classification in order to eliminate possible biases) a number of locations in the data where the cover type of the earth surface feature is known. The X, Y coordinates of these test areas are determined and one is able to obtain a table to the classification results for each of the individual test areas and for all test areas in combination. Figure 13 shows an example of such a classification table. Such tables can also be obtained for the training fields, in which case one is simply examining the results to determine whether or not there is a possibility to obtain spectral separation between two cover types of interest.

The table showing classification results for test areas is primarily an indication of the degree of reliability of the classification result. It is possible, for example, for the analyst to have specified to the computer a very atypical set of fields on which to train for the classification task. In such a case, the classification accuracy for the training fields could be quite high but the overall classification throughout the flightline would be rather poor. This would then be indicated by a gross difference in accuracy in the training class classification results and test classification results.

If the classification results are accurate, the analyst can also designate the entire flightline area as a test area and obtain a table which would give an indication of the amount of ground cover present in that flightline area for each of the different cover types classified. This could be expressed as a percentage of the total area or, by knowing the altitude of the aircraft and by applying factors to correct for scanner geometry distortion, one could obtain the table expresses as acreages of the different cover types.

Thus, the opportunity exists to obtain results from this type of classification either in a map format or in a tabular format. The need exists for both types of results formats to be available to various user groups.

Further Considerations and Future Needs in ADP of Multispectral Scanner Data

Only the spectral component of the data analysis problem has been discussed thus far. However, as previously indicated, the differences in spectral characteristics between one species of green vegetation and another have been found to be rather subtle at many times during the growing season. Of considerable importance are the changes in spectral characteristics of designated species at various times during the growing season. For example, in mid-July it is quite difficult to spectrally differentiate between corn and soybeans in that they both represent a solid canopy of green vegetation. However, two weeks later, after the corn is tasseled, enough of a difference in spectral response exists to obtain a much more reliable classification result. It therefore becomes quite important to take advantage of selected times during the growing season when different cover types of interest offer rather distinct and perhaps unique spectral characteristics. For example, in late June in central Indiana wheat is the only crop species having a rather golden brown appearance, and therefore produces a set of spectral characteristics quite different from the surrounding green vegetation. This would indicate that the time period in which remote sensor data is collected should be carefully selected to optimize the opportunity for spectral differentiation of the particular cover types of interest.

Another approach which also utilizes temporal data involves data collection from more than one time during the growing season and then combine, or overlay, one set of data on to another to obtain a set of multispectral and multitemporal data. Such an approach offers many distinct advantages. For example, even though wheat may offer a unique and characteristic spectral signature in late June, the agricultural user may desire to obtain a map of wheat acreages much earlier in the growing season. By collecting data in early fall, very little of the area overflown would be in bare soil except for fallow fields and those being prepared for planting wheat. Bare soil, however, is not positive evidence that the area will be planted in wheat. In the early spring, the wheat is lush and green, while most of the surrounding areas are in bare soil or pasture land which also would be lush and green. Therefore, wheat does not display a unique spectral signature in either the fall (when other areas also look like bare soil), or in the spring when other areas also are in a lush green vegetative condition. However, by combining data from the two seasons, if there is an area that is bare soil in the fall and bare soil again in the spring, it is not wheat; if it is green vegetation in the spring but is also green vegetation in the fall, again it would not be wheat. Only those fields displaying spectral characteristics of bare soil in the fall and green vegetation early in the spring should have a high probability of actually being wheat. From the pattern recognition standpoint, the classification algorithm is simply utilizing data in a measurement vector and it makes no difference whether such data is a vector comprised of data from different wavelength bands at one time of the year, or data from the same wavelength band at several times of the year, or a combination of data from several wavelength bands from several times of the year. By using such a multispectral and multitemporal analysis procedure, one ought to be able to obtain more accurate classification accuracies much earlier in the growing season than would be possible utilizing multispectral data alone. However, accurate overlay of different data sets is not a simple problem, particularly if different geometric configurations are involved.

In addition to using spectral and temporal data, there appears to be strong evidence that a need exists for making use of the spatial data contained in scanner imagery. This is similar to using textural information in photointerpretation. For example, examination of aerial photography indicates that it is quite easy to distinguish between forested and agricultural lands because of the coarser texture of the forested areas. Such a texture is created by the individual spectral characteristics of the sun-side of tree crowns, the shaded side of tree crowns, shadow areas between the tree crowns, etc. Analysis of the spectral characteristics of leaves of various tree species has revealed that frequently they are spectrally very similar to the spectral characteristics of agricultural crop species. This similarity has been confirmed in the analysis of multispectral scanner data, laboratory DK-2 spectra, and field spectral analysis results. By utilizing the spectral characteristics not only of an individual resolution element but also the surrounding resolution elements, and taking advantage of the spatial distribution of the multispectral data, it would appear that one should considerably increase the accuracy of ADP classification results. A great deal remains to be done to develop these techniques into a usable system, but I am convinced that such development will offer a considerable increase in capability for automatic data processing of remote sensor data. Because larger resolution elements will utilize the average spectral response over larger areas, it would appear that there are some situations in which a coarser resolution in the multispectral scanner data (i.e. data obtained from the Earth Resources Technology Satellite system) would offer considerable advantage for increased accuracy in the use of ADP techniques on remote sensor data.

The primary limiting factor for ADP in the future will involve the degree of sophistication in classification required by the user, in relation to the degree of natural variability encountered over large geographic areas. Numerical analysis techniques can achieve a great deal of speed, and are therefore ideally suited to analysis problems involving large quantities of data, but will probably need.to be limited to relatively simple analysis situations. Therefore it is necessary to consider the degree of sophistication required in an analysis procedure and the amount of data requiring analysis before concluding that either automatic data processing or manual processing is the best approach for any particular information requirement.

Summary and Conclusions

When one considers current predictions concerning population growth and compares them to the existing conditions of our natural resources, it becomes obvious that a great need exists for obtaining accurate up-to-date information concerning our resource base. Only by providing our resource managers with the type of information that they require can we hope that they will be able to carry out their jobs effectively and efficiently. Remote Sensing has already been proven a useful tool in several areas of application and management of our natural resources and it would appear that the potential applications for remote sensing are nearly unlimited.

The application of ADP techniques to multispectral scanner data has been proven feasible. It seems apparent that such techniques will be necessary in the future if we were to take full advantage of our ability to collect data at frequent intervals over vast geographic areas. Future developments in the handling of temporal and spatial data as well as spectral data should bring about significant improvements in automatic data analysis techniques.

Acknowledgements

The computer printout examples illustrating the various processing sequences utilizing LARSYSAA were prepared by Mr. Michael Coggeshall, Dept. of Forestry and Conservation, Purdue University. The author also acknowledges his many collegues at LARS who were involved in the various aspects of the development of the analysis techniques described.

Literature Cited

1. Hoffer, R. M. 1968. "Biophysical Research at LARS-Purdue." Proceedings of the Earth Resources Aircraft Program Status Review, NASA Manned Spacecraft Center, Houston, Texas. pp. 31-1 to 31-28.

2. Hoffer, R. M., and F. E. Goodrick. 1971. "Variables in Automatic Classification Over Extended Remote Sensing Test Sites." Proceedings of the 7th International Symposium on Remote Sensing of Environment, The University of Michigan, Ann Arbor, Michigan. pp. 1967-1981.

3. Landgrebe, D. A. and Staff. 1967. "Automatic Identification and Classification of Wheat by Remote Sensing." Research Progress Report 279, Agricultural Experiment Station, Purdue University, Lafayette, Indiana. 7 pp.

4. LARS Staff. 1970. "Remote Multispectral Sensing in Agriculture." Laboratory for Applications of Remote Sensing, Vol. 4 (Annual Report). Research Bulletin No. 873, Agricultural Experiment Station and Engineering Experiment Station, Purdue University, Lafayette, Indiana. 113 pp.

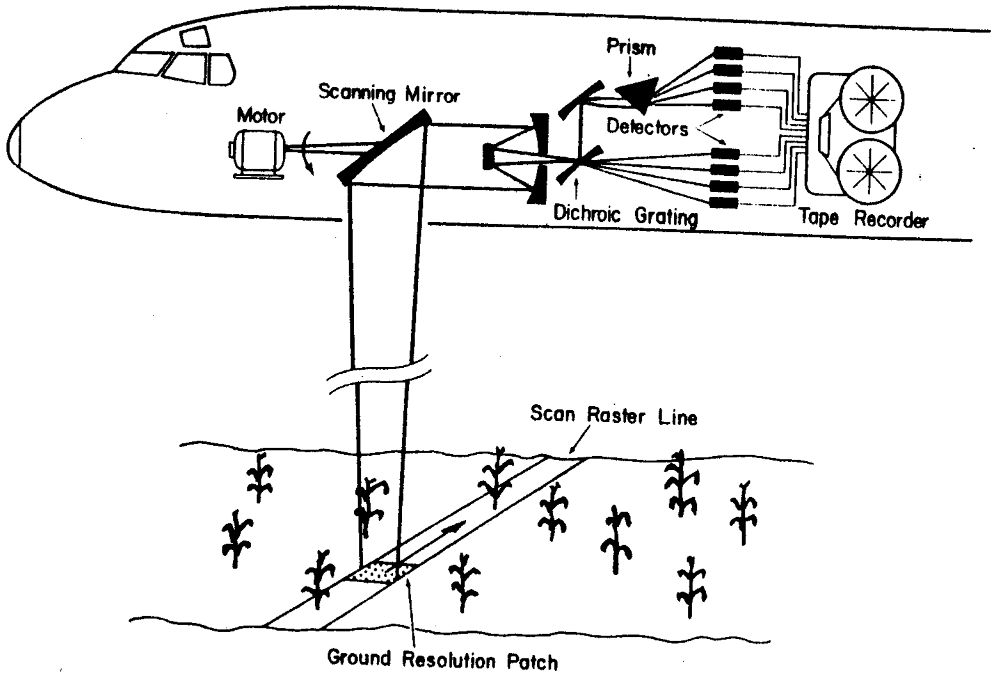


Figure 1 Schematic of a Multispectral Optical Mechanical Scanner

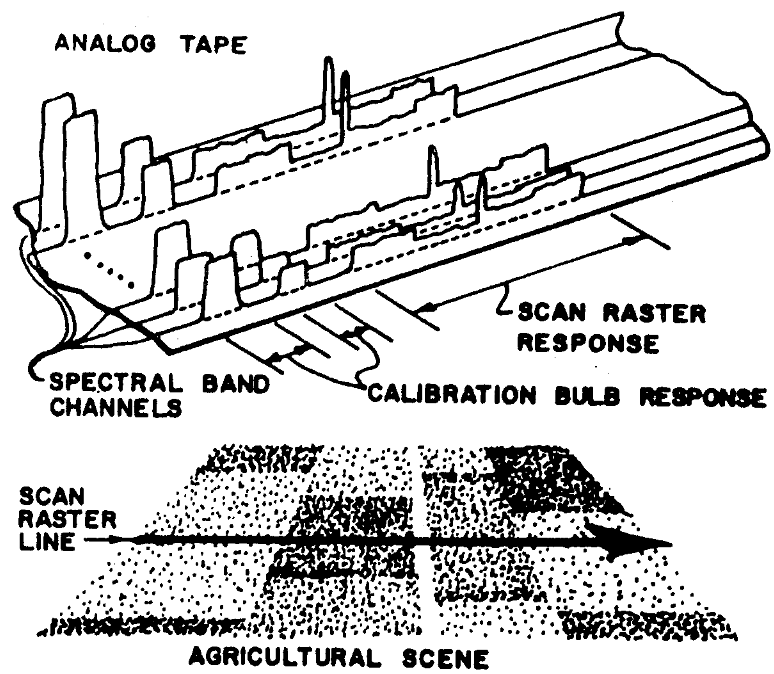
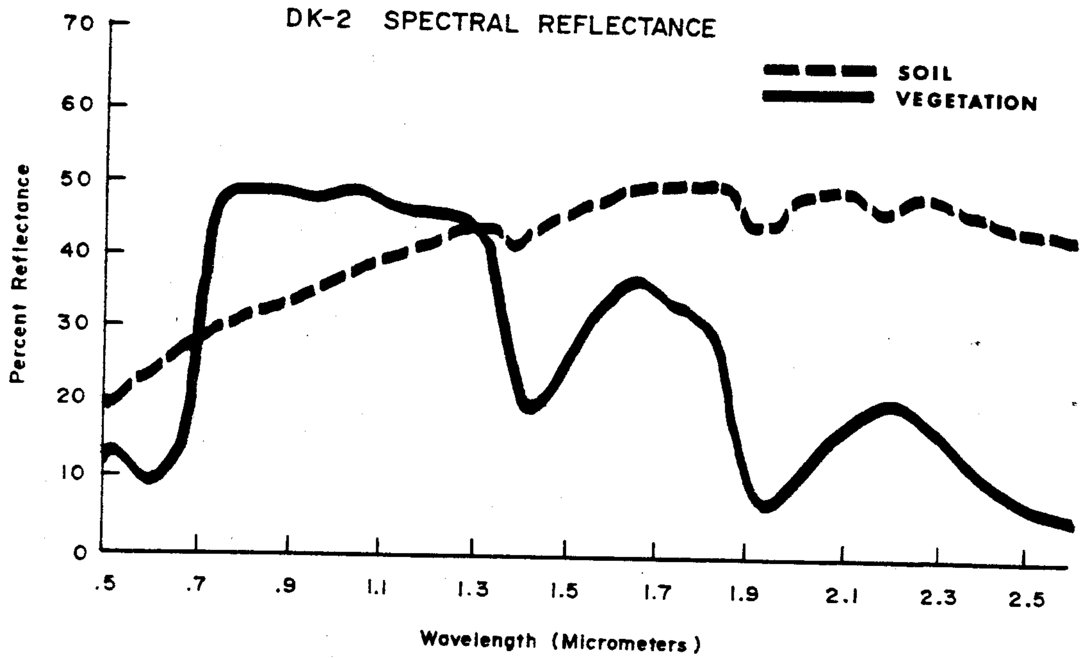


Figure 2 Data Collection Format, Using a Multispectral Scanner and Multiple-Track Recorder.



Multispectral Scanner Channels

Figure 3 Spectral Reflectance Curves for Green Turgid Vegetation and Air Dry Soils. These curves represent averages of 240 spectra from vegetation and 154 spectra from air dry soils. This type of data is frequently referred to as representing the spectral signature for the materials of interest.

LABORATORY FOR APPLICATIONS OF REMOTE SENSING
PURDUE UNIVERSITY

RUN NUMBER............ 71052501	DATE.................. 8/12/71
FLIGHT LINE.. CRN BLT LO FL218	TIME...................... 1004
DATA TAPE.................. 511	ALTITUDE............... 5000
REFORMATTING DATE.FEB 24,1972	GROUND HEADING.... 180 DEGREES

CHANNEL 7 SPECTRAL BAND 0.61 TO 0.70 MICROMETERS CALIBRATION CODE = 1 CO = 0.0

THE CHARACTER SET USED FOR DISPLAY IS

FROM 11.5 TO 27.5 DISPLAYED AS M
FROM 27.5 TO 31.5 DISPLAYED AS $
FROM 31.5 TO 35.5 DISPLAYED AS X
FROM 35.5 TO 43.5 DISPLAYED AS Z
FROM 43.5 TO 51.5 DISPLAYED AS *
FROM 51.5 TO 55.5 DISPLAYED AS I
FROM 55.5 TO 63.5 DISPLAYED AS /
FROM 63.5 TO 71.5 DISPLAYED AS =
FROM 71.5 TO 83.5 DISPLAYED AS -
FROM 83.5 TO 255.5 DISPLAYED AS

HISTOGRAM BLOCK(S)

RUN NUMBER........... 71052501
LINES........ (1300, 1600, 2)
COLUMNS..... (1, 222, 2)
CALIBRATION CODE............ 1

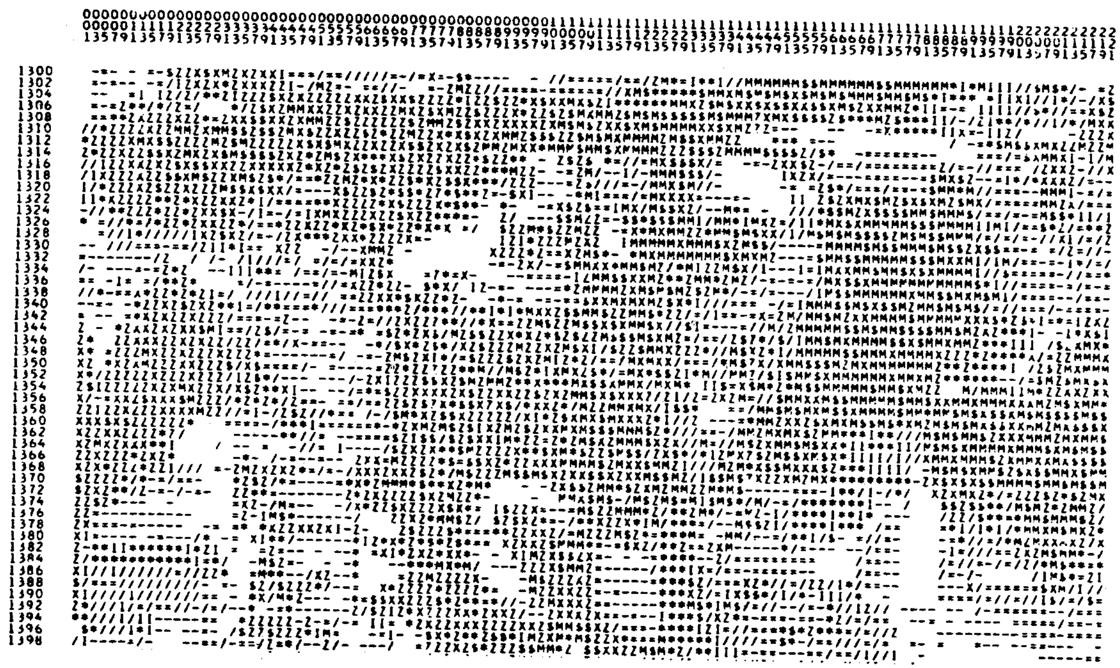


Figure 4 Alphanumeric Pictorial Printout of Scanner Data in One Wavelength Band. The different symbols represent different amplitudes of reflected or emitted energy.

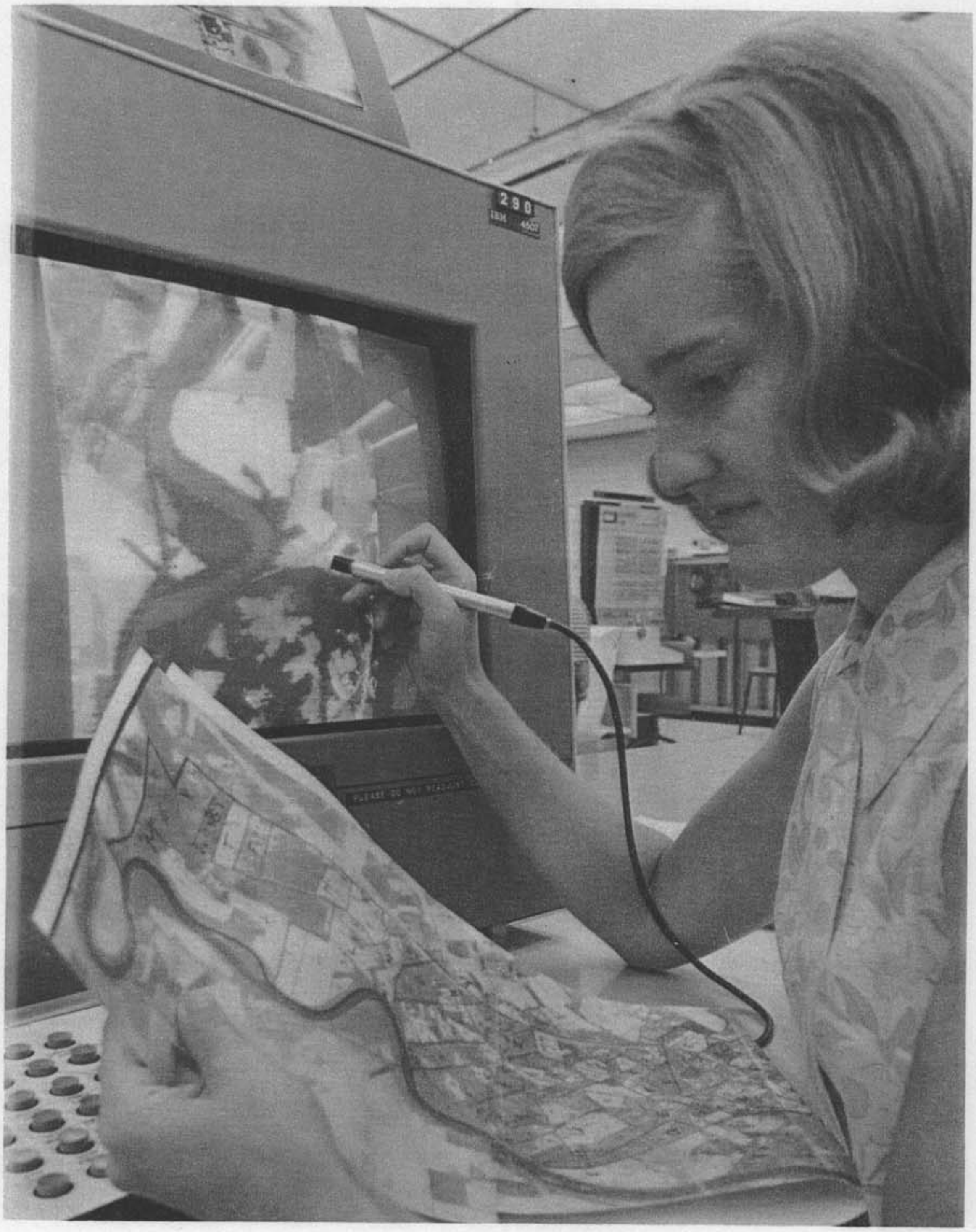


Figure 5 LARS Digital Display Unit

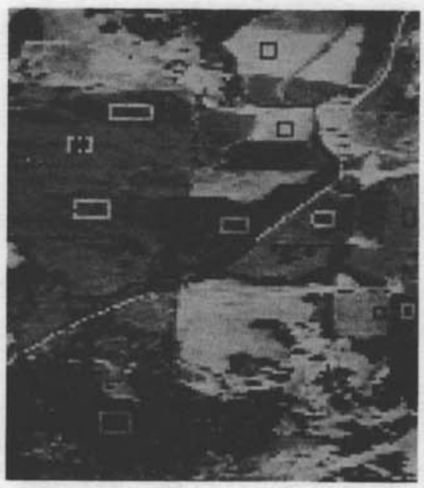


Figure 6 Illustration of Digital Scanner Data Showing Locations of Training Areas

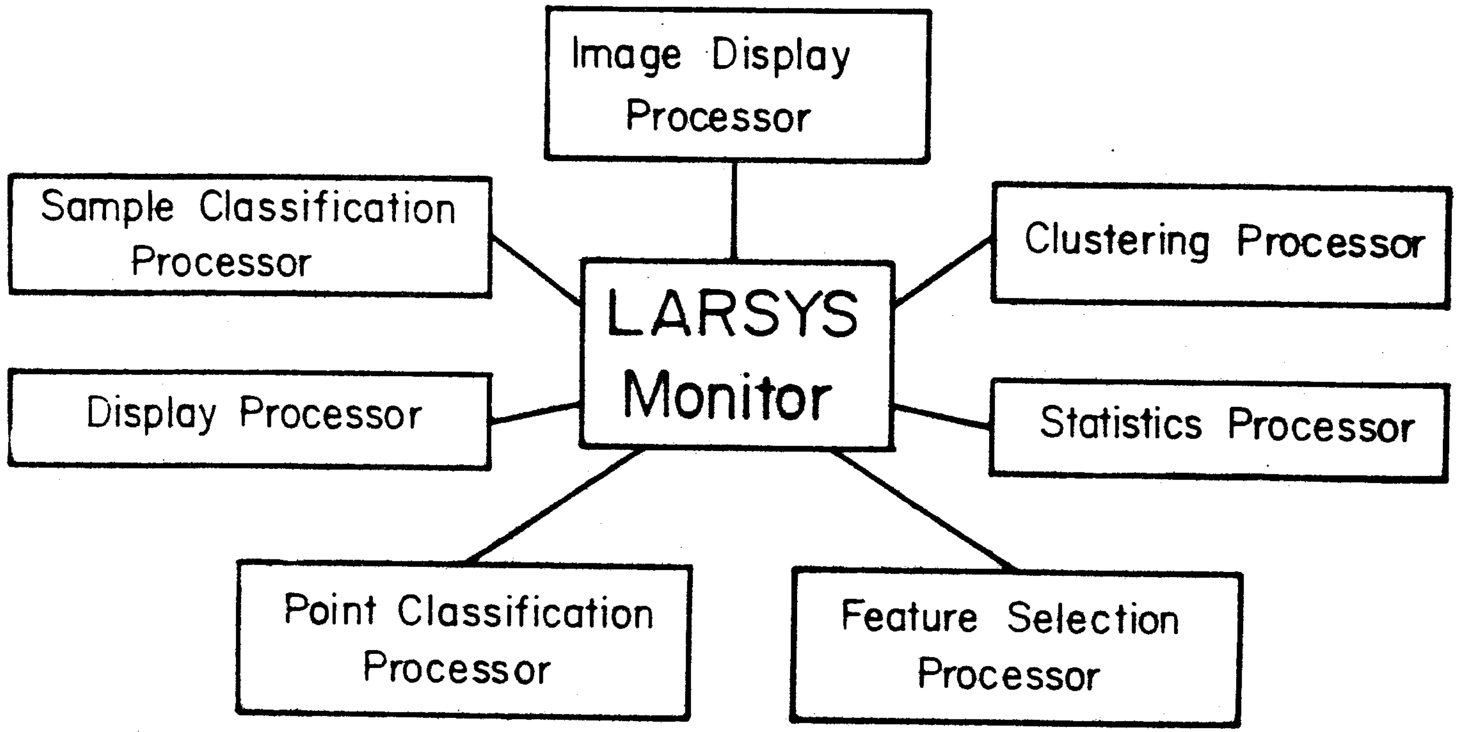
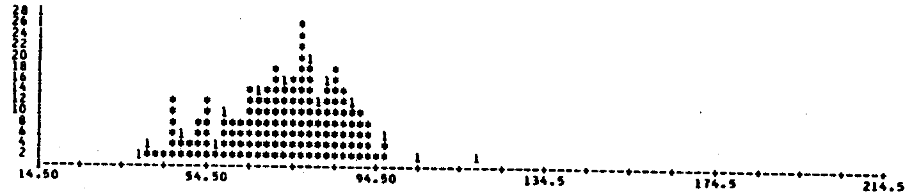


Figure 7 Organization of LARSYSAA

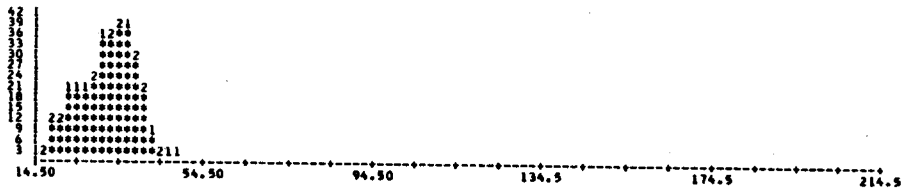
LARSYSAA
VERSION 2

LABORATORY FOR APPLICATIONS OF REMOTE SENSING
PURDUE UNIVERSITY

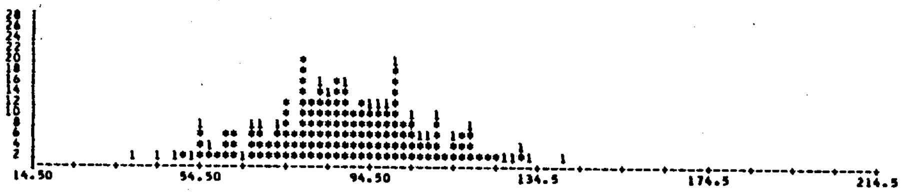
CLASS....FOREST TOTAL NUMBER OF SAMPLES... 308
HISTOGRAM(S)
CHANNEL 4 0.52 - 0.57 MICROMETERS
EACH * REPRESENTS 2 POINT(S).



CHANNEL 7 0.61 - 0.70 MICROMETERS
EACH * REPRESENTS 3 POINT(S).



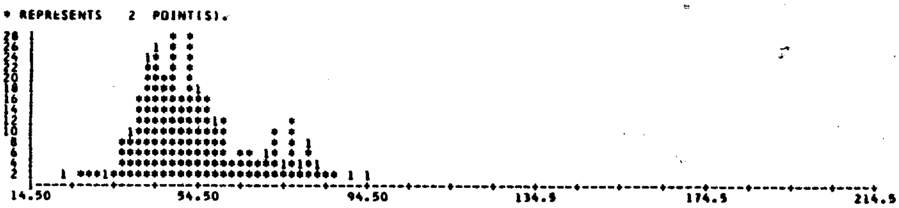
CHANNEL 8 0.72 - 0.92 MICROMETERS
EACH * REPRESENTS 2 POINT(S).



LARSYSAA
VERSION 2

LABORATORY FOR APPLICATIONS OF REMOTE SENSING
PURDUE UNIVERSITY

CLASS....FOREST TOTAL NUMBER OF SAMPLES... 308
HISTOGRAM(S)
CHANNEL 10 1.50 - 1.80 MICROMETERS
EACH * REPRESENTS 2 POINT(S).



CHANNEL 12 9.30 - 11.70 MICROMETERS
EACH * REPRESENTS 3 POINT(S).

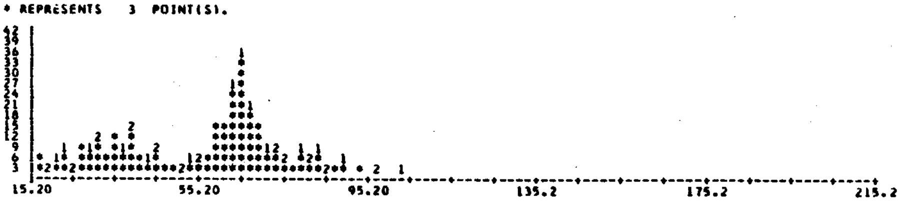
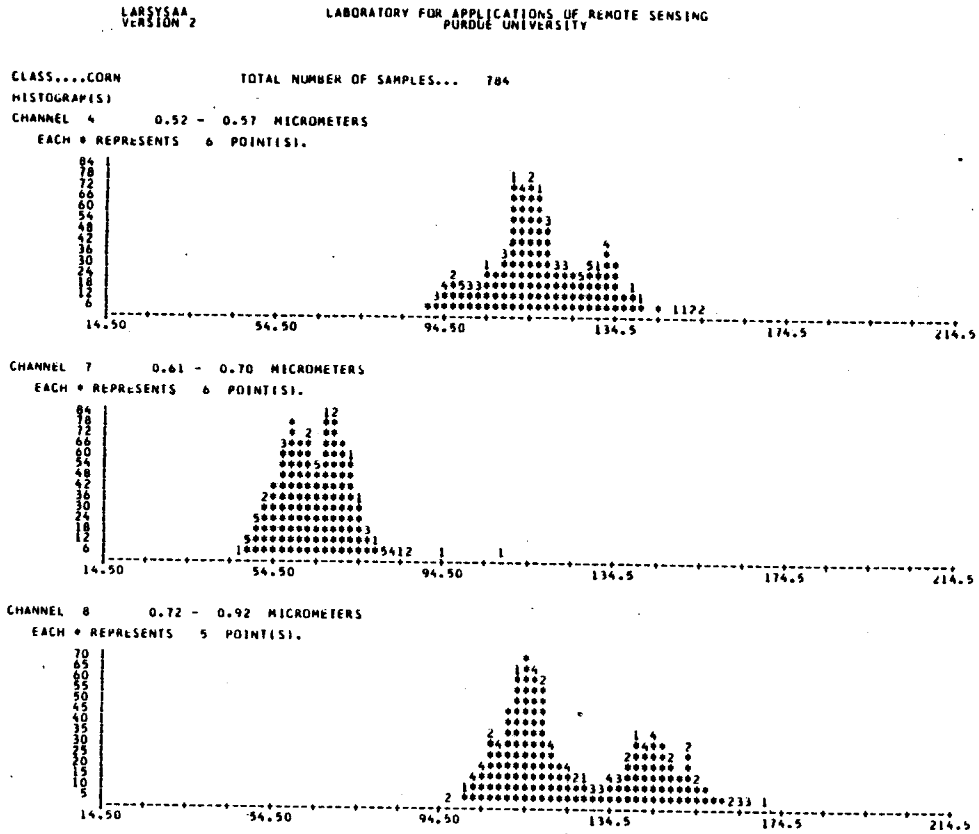


Figure 8 Histograms of Five Wavelength Bands of Data Compiled from (a) Several Forested Areas and



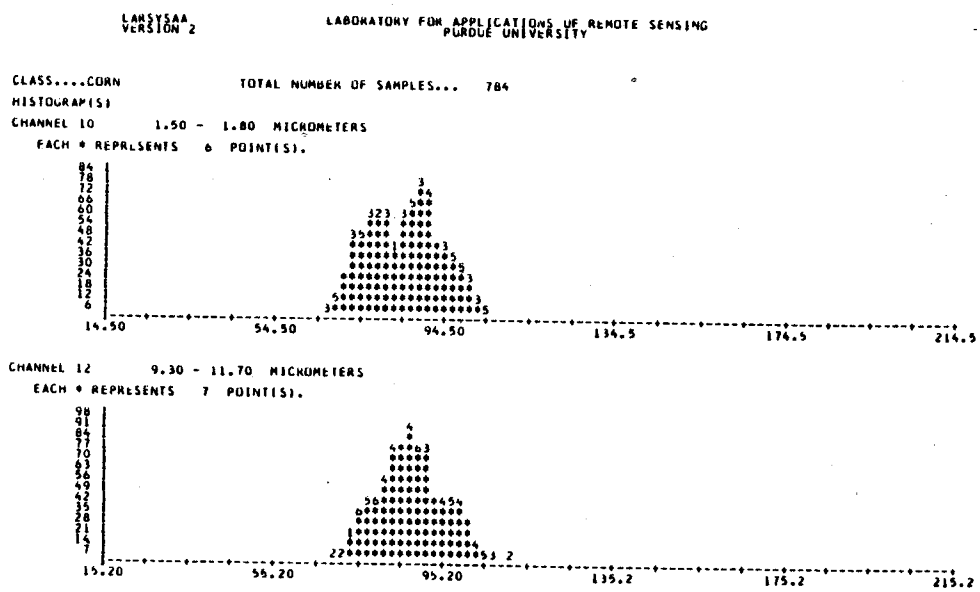


Figure 8 (b) Several Corn Fields. Note that the data can have Gaussian distribution in all wavelength bands except one, where a distinctly bi-modal distribution is found.

LARSYSAA
VERSION 2

LABORATORY FOR APPLICATIONS OF REMOTE SENSING
PURDUE UNIVERSITY

CLASS....FOREST   TOTAL NUMBER OF SAMPLES... 308

THE MEAN AND STANDARD DEVIATION VECTORS

CHANNEL	1	2	3	4	5	6	7	8	9	10	11	12
SPECTRAL BAND	0.46 - 0.49	0.48 - 0.51	0.50 - 0.54	0.52 - 0.57	0.54 - 0.60	0.58 - 0.65	0.61 - 0.70	0.72 - 0.92	1.00 - 1.40	1.50 - 1.80	2.00 - 2.60	9.30 - 11.70
MEAN	46.26	44.44	54.29	71.16	45.08	40.73	30.57	87.76	80.68	52.67	39.49	54.69
STD. DEV.	5.05	5.45	8.87	13.79	9.16	8.48	6.46	18.15	15.54	13.13	8.03	19.35

CORRELATION MATRIX

SPECTRAL BAND	0.46 - 0.49	0.48 - 0.51	0.50 - 0.54	0.52 - 0.57	0.54 - 0.60	0.58 - 0.65	0.61 - 0.70	0.72 - 0.92	1.00 - 1.40	1.50 - 1.80	2.00 - 2.60	9.30 - 11.70
0.46-0.49	1.00											
0.48-0.51	0.72	1.00										
0.50-0.54	0.79	0.80	1.00									
0.52-0.57	0.64	0.71	0.83	1.00								
0.54-0.60	0.75	0.84	0.90	0.84	1.00							
0.58-0.65	0.57	0.76	0.72	0.79	0.80	1.00						
0.61-0.70	0.78	0.87	0.88	0.79	0.91	0.81	1.00					
0.72-0.92	0.48	0.57	0.52	0.55	0.65	0.49	0.53	1.00				
1.00-1.40	0.13	0.22	0.11	0.24	0.20	0.31	0.15	0.57	1.00			
1.50-1.80	0.30	0.32	0.26	0.31	0.31	0.34	0.29	0.63	0.81	1.00		
2.00-2.60	0.44	0.45	0.37	0.42	0.43	0.47	0.46	0.57	0.70	0.73	1.00	
9.30-11.70	0.62	0.73	0.78	0.74	0.79	0.78	0.79	0.52	0.33	0.44	0.53	1.00

Figure 9 Table Showing Statistical Data Compiled from Several Forested Areas

LARSYSAA
VERSION 2

LABORATORY FOR APPLICATIONS OF REMOTE SENSING
PURDUE UNIVERSITY

COINCIDENT SPECTRAL PLOT (MEAN PLUS AND MINUS ONE STD. DEV.) FOR CLASS(ES)

LEGEND
A = CLASS 1 DECIDUOUS
B = CLASS 2 CONIFER
C = CLASS 3 WATER
D = CLASS 4 FORAGE
E = CLASS 5 CORN
F = CLASS 6 SOYBEAN

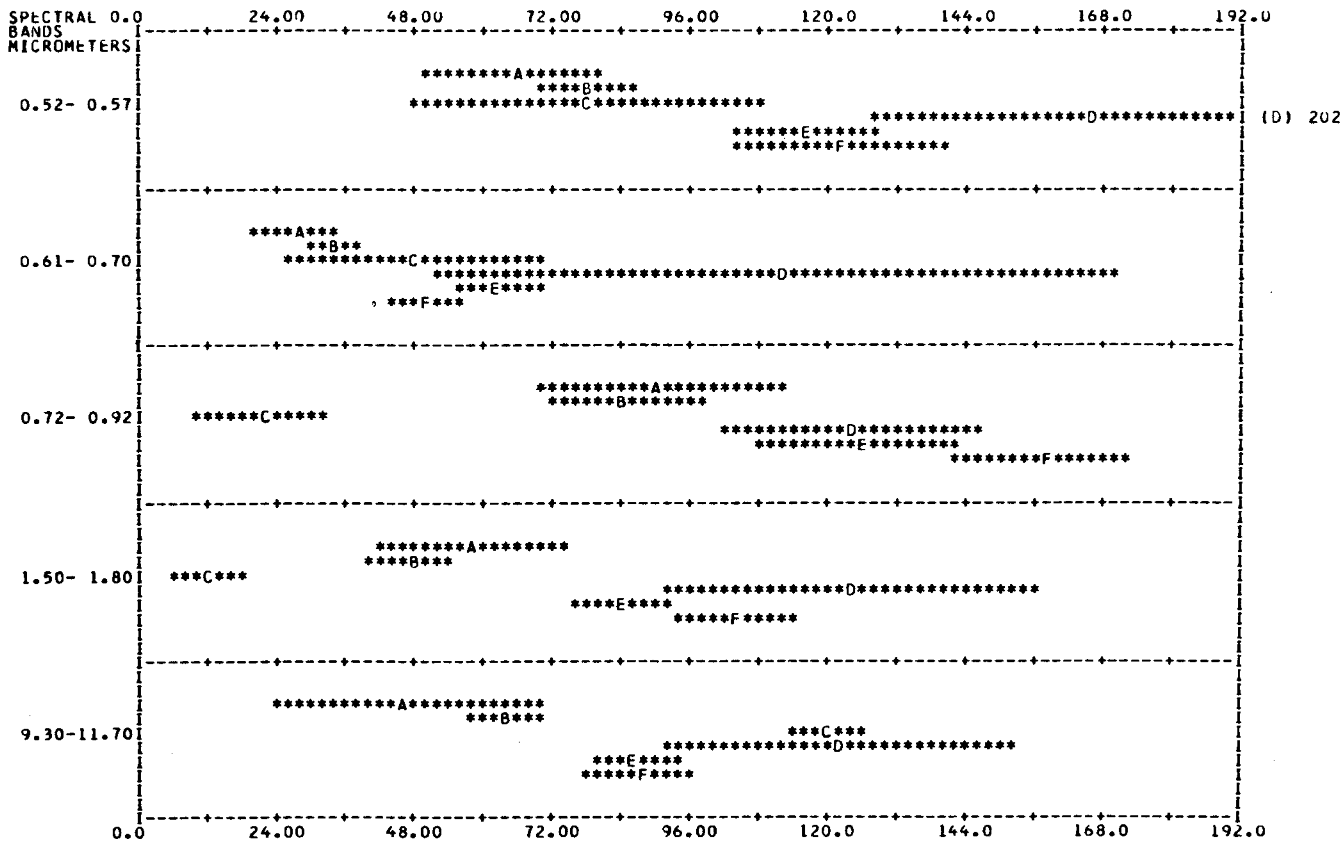


Figure 10 Multispectral Response Graph Depicting Relative Reflectance and Emittance from Five Selected Wavelength Bands for the Major Cover Types Presented in this Set of Data

RESTORED DATA
RETENTION LEVEL .. 495 MINIMUM 0
MAXIMUM 30000

DIVERGENCE **WITH** SATURATING TRANSFORM

*** RESULTS ORDERED ACCORDING TO DIJ(MIN) ***

	FEATURES	DIJ(MIN)	D(AVE)	WEIGHTED INTERCLASS DIVERGENCE (DIJ) TW (10)	TF (10)	TC (10)	TS (10)	WF (10)	WC (10)	WS (10)	FC (10)	FS (10)	CS (10)
1.	4 7 10 12	1975.	1997.	2000	2000	1998	1997	2000	2000	2000	2000	2000	1975
2.	4 7 11 12	1971.	1997.	2000	2000	1997	1998	2000	2000	2000	2000	2000	1971
3.	5 7 10 12	1969.	1996.	2000	2000	1998	1997	2000	2000	2000	2000	2000	1969
4.	5 7 11 12	1963.	1996.	2000	2000	1996	1998	2000	2000	2000	2000	2000	1963
5.	5 7 8 10	1959.	1990.	2000	2000	1960	1978	2000	2000	2000	2000	2000	1959
6.	5 7 9 10	1957.	1991.	1999	2000	1967	1983	2000	2000	2000	2000	2000	1957
7.	4 6 10 12	1957.	1995.	2000	2000	1993	1997	2000	2000	2000	2000	2000	1957
8.	5 7 10 11	1956.	1987.	1996	2000	1960	1956	2000	2000	2000	2000	2000	1959
9.	3 7 10 12	1955.	1995.	2000	2000	1998	1997	2000	2000	2000	2000	2000	1955
10.	2 4 7 10	1954.	1987.	1995	2000	1955	1967	2000	2000	2000	2000	2000	1954
11.	1 4 7 10	1952.	1986.	1994	2000	1954	1958	2000	2000	2000	2000	2000	1952
12.	4 7 9 10	1951.	1991.	1999	2000	1951	1987	2000	2000	2000	2000	2000	1970
13.	3 7 11 12	1951.	1995.	2000	2000	1997	1998	2000	2000	2000	2000	2000	1951
14.	4 7 8 10	1950.	1991.	2000	2000	1950	1986	2000	2000	2000	2000	2000	1972
15.	5 7 9 11	1948.	1989.	1999	2000	1962	1985	2000	2000	2000	2000	2000	1948
16.	6 7 10 12	1947.	1994.	2000	2000	1997	1996	2000	2000	2000	2000	2000	1947
17.	5 7 8 11	1946.	1988.	2000	2000	1946	1983	2000	2000	2000	2000	2000	1953
18.	5 6 7 11	1945.	1983.	1987	2000	1946	1949	2000	2000	2000	2000	2000	1945
19.	4 5 7 11	1945.	1983.	1985	2000	1946	1953	2000	2000	2000	2000	2000	1945
20.	2 5 7 10	1945.	1988.	1995	2000	1975	1961	2000	2000	2000	2000	2000	1945
21.	4 7 9 11	1944.	1990.	1993	2000	1944	1988	2000	2000	2000	2000	2000	1965
22.	4 6 11 12	1943.	1993.	2000	2000	1991	1998	2000	2000	2000	2000	2000	1943
23.	1 5 7 10	1943.	1986.	1994	2000	1972	1952	2000	2000	2000	2000	2000	1943
24.	4 5 7 10	1942.	1985.	1994	2000	1963	1942	2000	2000	2000	2000	2000	1952
25.	3 4 7 10	1942.	1983.	1995	2000	1945	1942	2000	2000	2000	2000	2000	1950
26.	2 4 6 10	1942.	1985.	1997	2000	1945	1967	2000	2000	2000	1999	2000	1942
27.	4 7 10 11	1940.	1986.	1996	2000	1940	1959	2000	2000	2000	2000	2000	1968
28.	4 6 7 10	1940.	1984.	1997	2000	1940	1943	2000	2000	2000	2000	2000	1955
29.	2 5 7 11	1940.	1986.	1998	2000	1966	1969	2000	2000	2000	2000	2000	1940
30.	2 4 7 11	1940.	1985.	1987	2000	1940	1974	2000	2000	2000	2000	2000	1950
*** 106.	2 6 10 12	1883.	1988.	2000	2000	1997	1998	2000	2000	2000	2000	2000	1883

Figure 11 Table Showing an Example of the Feature Selection/ Separability Portion of the LARSYS Program

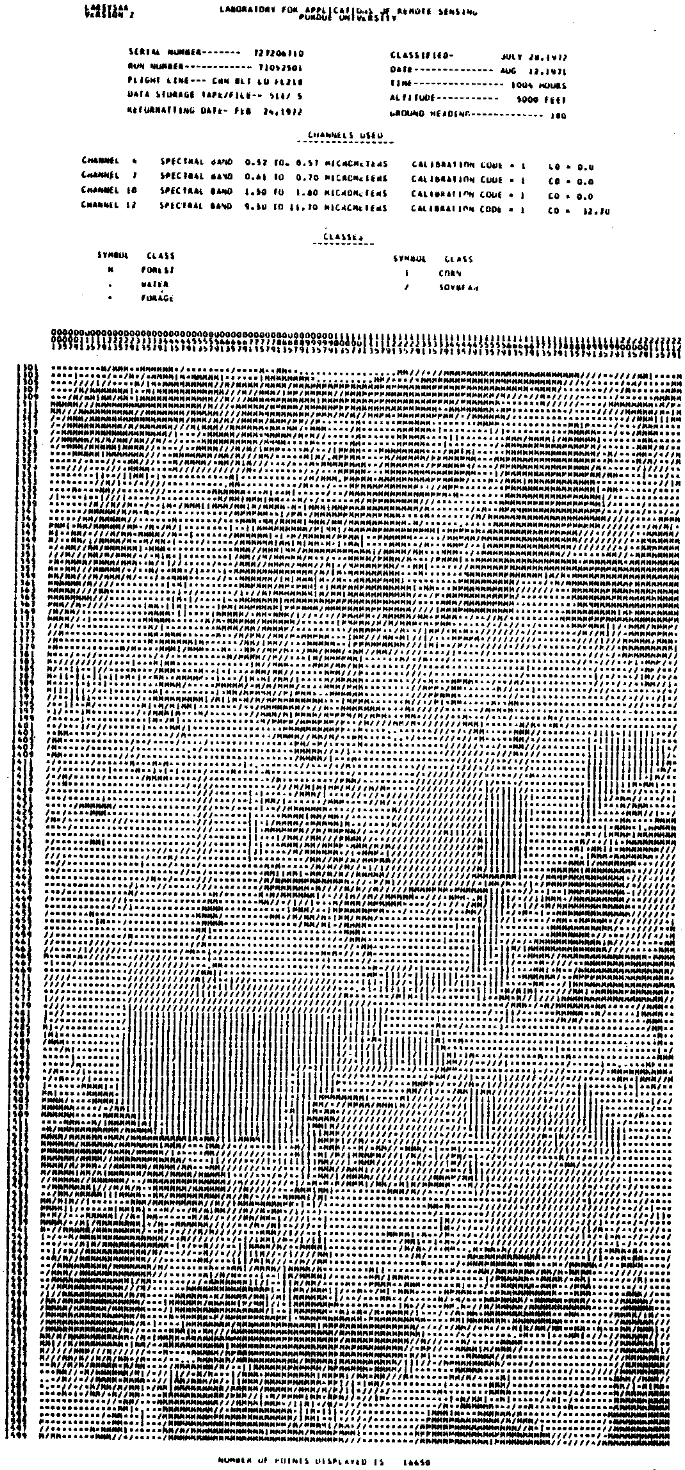


Figure 12 Alphanumeric Printout "Map" Showing the Classification Results Using Four Wavelength Bands and the Following Cover Types: Forest (M), Water (1), Forage (=), Corn (I), Soybeans (/)

LARSYSAA
VERSION 2

LABORATORY FOR APPLICATIONS OF REMOTE SENSING
PURDUE UNIVERSITY

SERIAL NUMBER------- 727206710 CLASSIFIED- JULY 28,1972

CHANNELS USED

CHANNEL 4	SPECTRAL BAND 0.52 TO 0.57 MICROMETERS	CALIBRATION CODE = 1	CO = 0.0
CHANNEL 7	SPECTRAL BAND 0.61 TO 0.70 MICROMETERS	CALIBRATION CODE = 1	CO = 0.0
CHANNEL 10	SPECTRAL BAND 1.50 TO 1.80 MICROMETERS	CALIBRATION CODE = 1	CO = 0.0
CHANNEL 12	SPECTRAL BAND 9.30 TO 11.70 MICROMETERS	CALIBRATION CODE = 1	CO = 32.70

CLASSES

	CLASS		CLASS
1	FOREST	4	CORN
2	WATER	5	SOYBEAN
3	FORAGE		

TEST CLASS PERFORMANCE

				NUMBER OF SAMPLES CLASSIFIED INTO				
	GROUP	NO OF SAMPS	PCT. CORCT	FOREST	WATER	FORAGE	CORN	SOYBEAN
1	FOREST	8101	93.8	7602	0	70	60	369
2	WATER	90	96.7	3	87	0	0	0
3	FORAGE	2953	85.6	5	1	2528	97	322
4	CORN	644	96.0	0	0	19	618	7
5	SOYBEAN	660	95.6	0	0	17	12	631
	TOTAL	12448		7610	88	2634	787	1329

OVERALL PERFORMANCE(11466/ 12448) = 92.1

AVERAGE PERFORMANCE BY CLASS(467.7/ 5) = 93.5

Figure 13 Table Showing Classification Results for the 179 Test Area Used to Sample the Accuracy of the Classification Results