LARS Information Note 21567

Research Progress Report 279
March, 1967

Automatic Identification and Classification of Wheat by Remote Sensing

by David A. Landgrebe and Staff of
Laboratory for Agricultural Remote Sensing, Purdue University

Introduction

The primary purpose of this paper is to provide a brief report on early results in the automatic classification of crops from airborne multispectral data by the Purdue Laboratory for Agricultural Remote Sensing. The specific problem treated here is the automatic detection of wheat in late June.

Remote multispectral sensing research has been carried on at Purdue University for about one year. This research on remote sensing in agriculture is directed at establishing methods to determine by remote means, species identification, state of maturity, disease conditions, soil types, soil moisture conditions and many other crop and soil parameters.

Research workers from engineering and the life sciences are cooperating at Purdue to develop techniques to sense remotely (observe and record from above ground level) agricultural situations and to assess automatically (review and determine by computer) their important characteristics.

Early Results in the Automatic Classification of Wheat

In the automatic detection of wheat in late June reported here, it should be noted that wheat in the test site region (northcentral Indiana) is near maturity. Results discussed are contained on the four computer printouts which may be interpreted with the aid of the aerial photograph showing the flight line.

Figure 1 shows a portion of a computer printout of the scanner imagery in pictorial form. The computer program which produces printouts of this type was written so that data may be conveniently edited. Column numbers at the top of the printout, and line numbers at the left edge of the printout, are used to determine the address of any given data point. A comparison between the photograph and this printout reveals a number of distinctive field patterns such that orientation from the photograph to the printout is easily possible. For example, consider field A at lines 1381 to 1543 and columns 1 to 115 (Figure 5 presents this area in larger scale), which shows a wheat field with oats planted in the center of it. (Column numbers proceed from right to left due to characteristics of the airborne scanner.)

The results of a first, very unsophisticated or simple, attempt at the automatic classification of wheat is shown on Figure 2 and enlarged in Figure 6. The method used in generating these results was as follows: In one channel of data, the mean and variance of wheat samples were determined. The

The key used on this photograph to designate the primary agricultural cover is A - Alfalfa, C - Corn, O - Oats, R - Rye, S - Soybeans, W - Wheat, P - Pasture, RC - Red Clover, SC - Sweet Clover, TIM - Timothy.

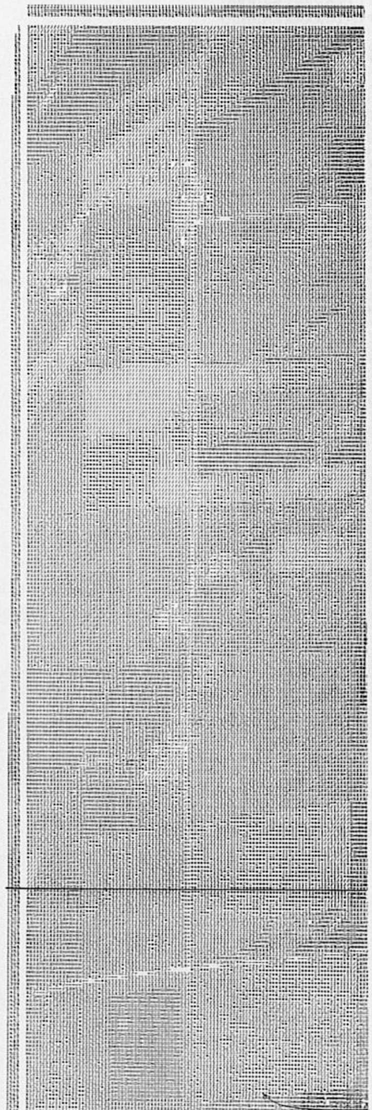

Figure 1. Computer printout of scanner imagery in pictorial form.

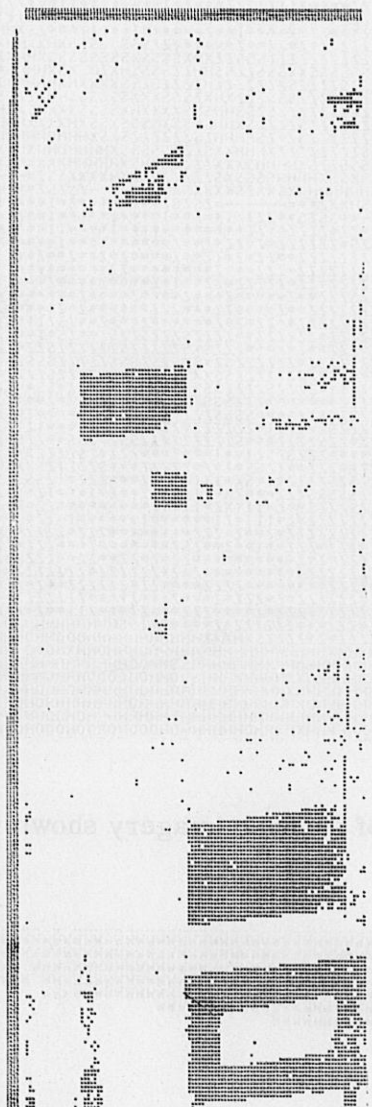

Figure 2. A first, simple attempt at automatic classification of wheat.

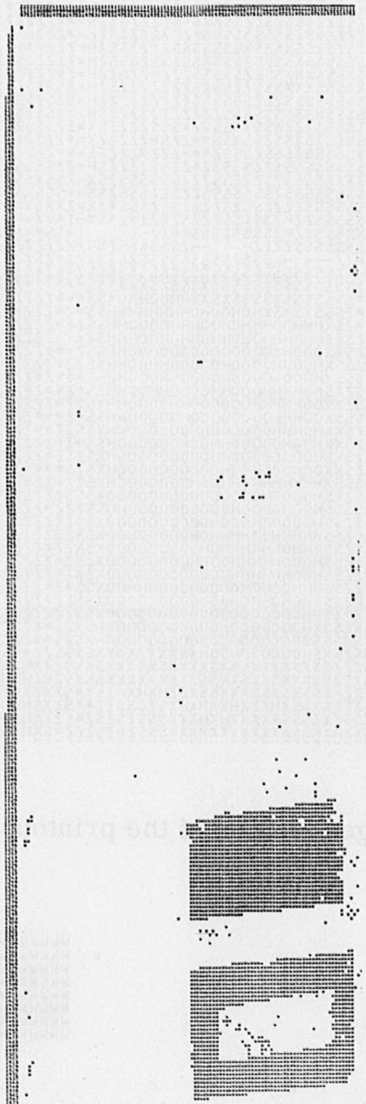

Figure 3. Similar to Figure 2 but 4 channels of information were used to simultaneously classify wheat.

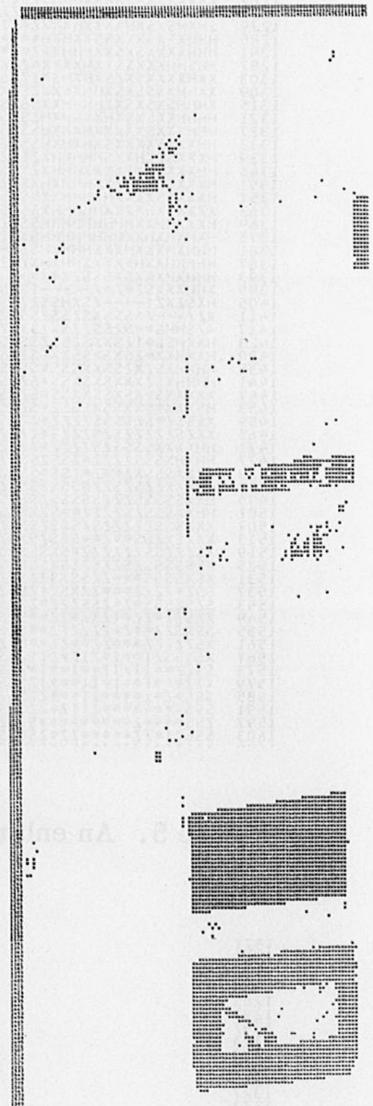

Figure 4. The first attempt to use truly automatic pattern recognition techniques.

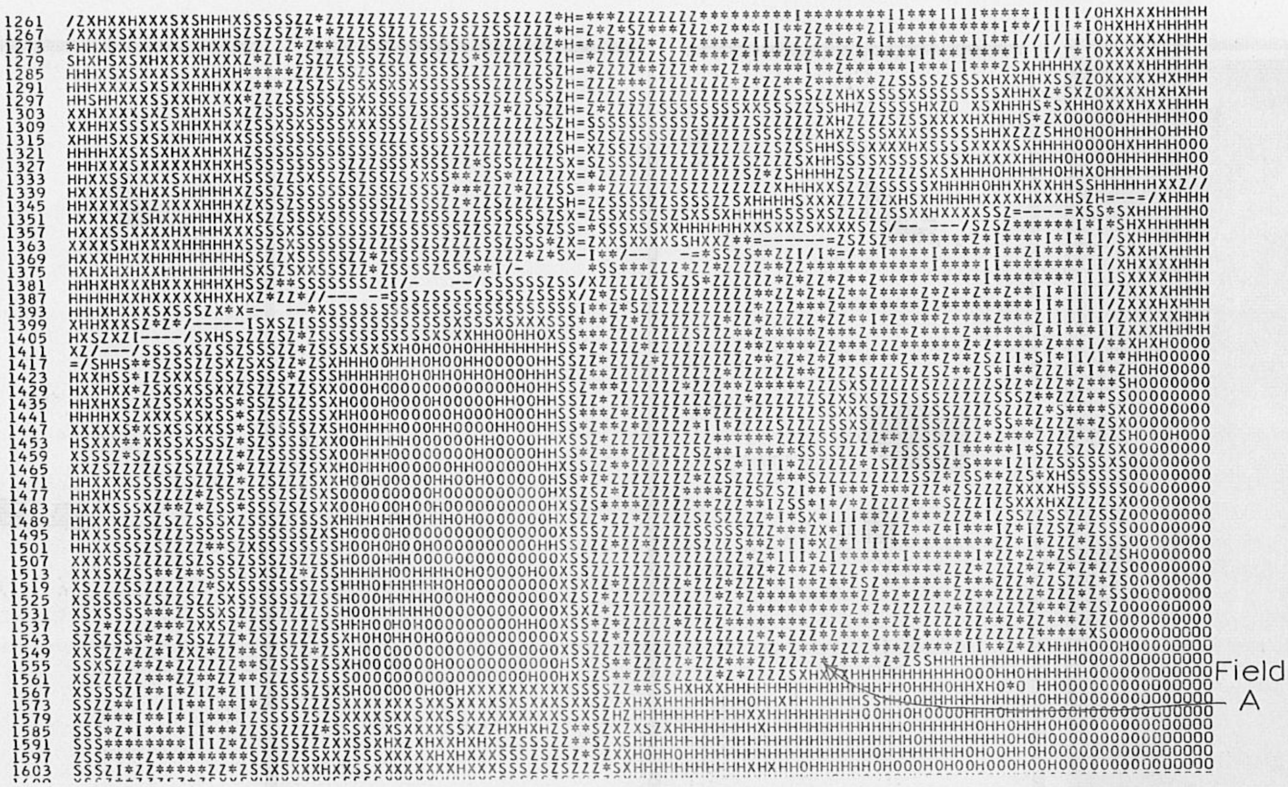

Figure 5. An enlarged section of the printout of scanner imagery shown in Figure 1.

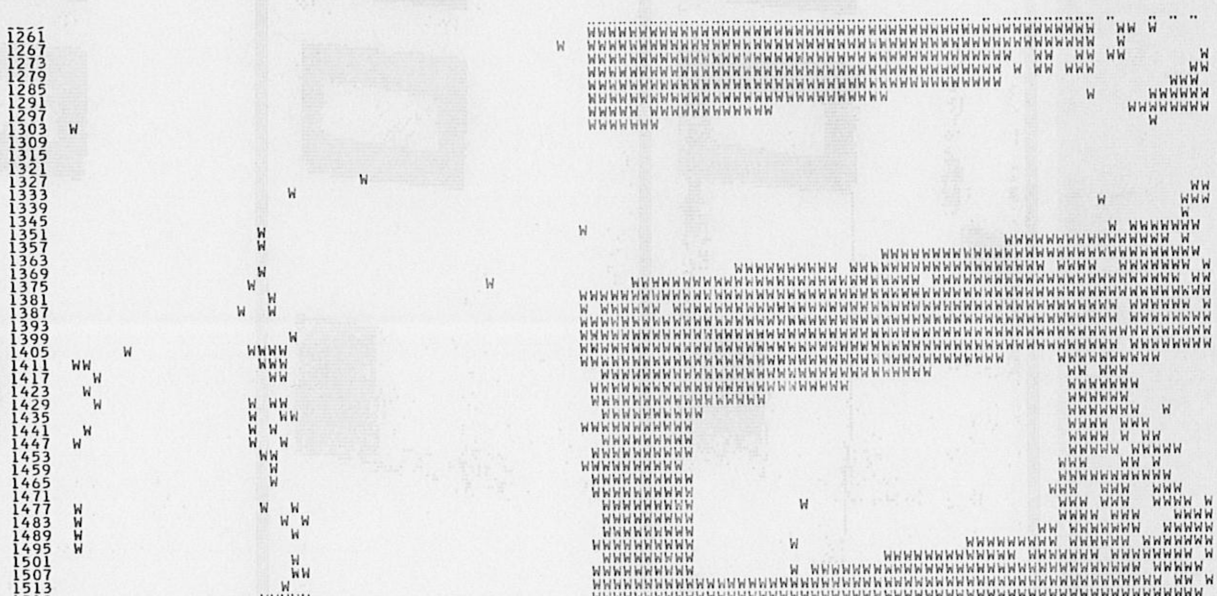

Figure 6. An enlarged section of an early attempt at automatic classification of wheat.

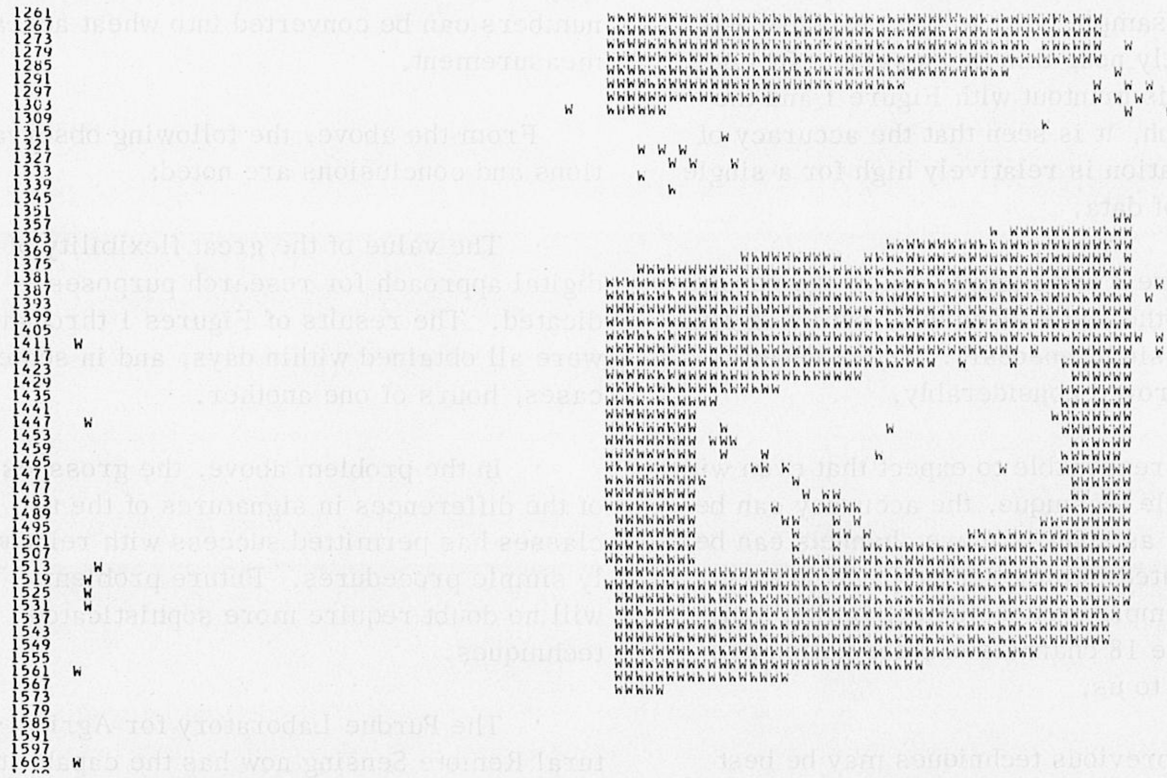

Figure 7. An enlarged section of the 4 channel automatic classification.

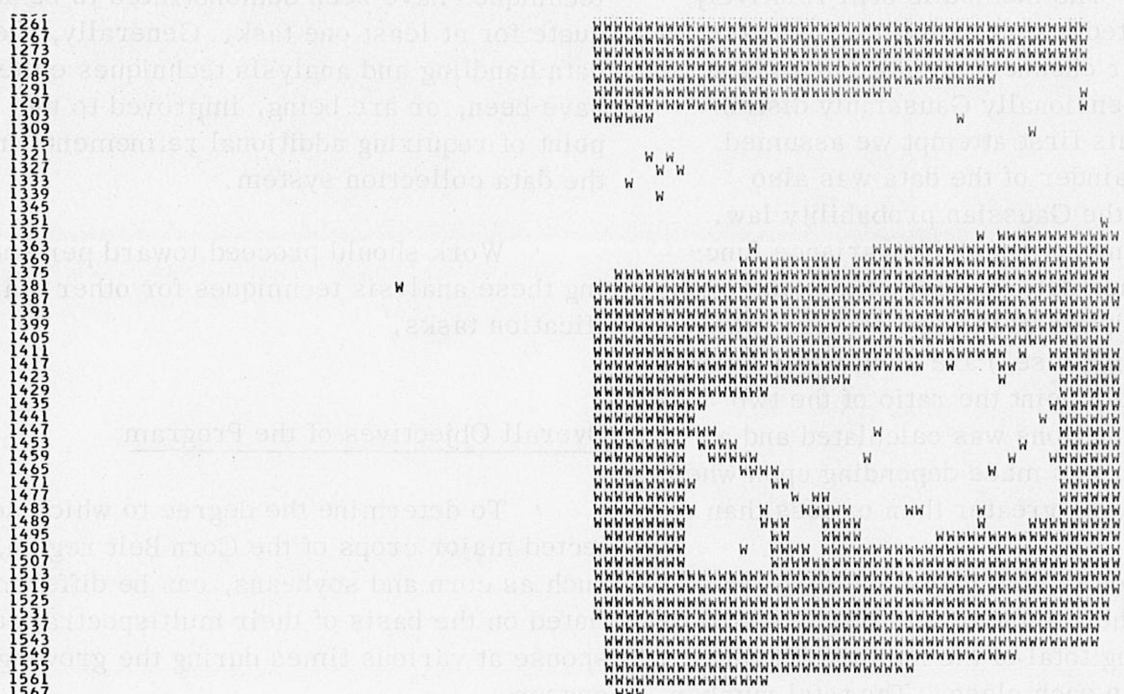

Figure 8. An enlarged section showing the truly automatic pattern recognition technique.

computer was then instructed to print a 'W' for each sample point whose response was sufficiently near this mean value. By comparing this printout with Figure 1 and the photograph, it is seen that the accuracy of classification is relatively high for a single channel of data.

Figure 3 and Figure 7 show the results of trying the same technique, but using four channels simultaneously. These results have improved considerably.

It is reasonable to expect that even with this simple technique, the accuracy can be improved as soon as more channels can be incorporated into the method. At present, due to (temporary) equipment limits, only four of the 18 channels of information are available to us.

The previous techniques may be best described as spectral matching techniques. Figure 4 shows the first attempt using a truly automatic pattern recognition technique. Figure 8 is a larger scale of one area shown in Figure 4. The method is still relatively unsophisticated and is as follows: We assume the four channels of data from wheat are four dimensionally Gaussianly distributed. For this first attempt we assumed that the remainder of the data was also governed by the Gaussian probability law. The two mean vectors and covariance functions were then determined. Thus, the probability description of the two classes (wheat and "everything else") are completely known. At each sample point the ratio of the two probability functions was calculated and a classification was made depending upon whether this ratio was greater than or less than one.

In addition to making a classification at each point, the computer was instructed to keep a running total of the number of points classified into each class. The total number of samples classified in this example was 64,240, and of this total 5,469 were classified as wheat. By knowing the altitude of the aircraft the area each sample point represents may be determined. Therefore, the above numbers can be converted into wheat acreage measurement.

From the above, the following observations and conclusions are noted:

- The value of the great flexibility of a digital approach for research purposes is indicated. The results of Figures 1 through 4 were all obtained within days, and in some cases, hours of one another.

- In the problem above, the grossness of the differences in signatures of the two classes has permitted success with relatively simple procedures. Future problems will no doubt require more sophisticated techniques.

- The Purdue Laboratory for Agricultural Remote Sensing now has the capability to analyze data collected on a semi-operational (large scale) nature.

- Effective data handling and analysis techniques have been demonstrated to be adequate for at least one task. Generally, the data handling and analysis techniques either have been, or are being, improved to the point of requiring additional refinements in the data collection system.

- Work should proceed toward perfecting these analysis techniques for other classification tasks.

<u>Overall Objectives of the Program</u>

- To determine the degree to which selected major crops of the Corn Belt region, such as corn and soybeans, can be differentiated on the basis of their multispectral response at various times during the growing season.

- To determine the amount of variation and to identify the major sources of variation in the multispectral response of selected agricultural features such as soil conditions and major crop species of the Corn Belt region at various times during the growing season.

· To determine and prescribe methods for gathering information from the ground that will allow prediction of multispectral response characteristics obtained by remote multispectral sensing techniques.

To achieve these objectives, the research at Purdue University involves three interrelated study areas:

1. Biophysical studies in the laboratory and field, designed to:

· Learn more about natural variations in plant and soil reflectance spectra and the factors which influence variations.

· Determine optimum portions of spectrum for remote sensoring.

· Determine the optimum time during the growing season to obtain remote multispectral data with airborne systems.

2. Remote multispectral sensing studies, using aircraft flights over selected agricultural areas near Lafayette, Indiana. Such aircraft systems can obtain simultaneous data on reflectance and emission characteristics of areas flown over in many spectral bands between 0.32 and 15 microns wavelength. Such data are currently recorded electronically on analog tapes and later converted to digital form. Data from this phase of the program are being used to determine the feasibility of optical-mechanical scanner system surveys to supplement U. S. Department of Agriculture, Statistical Reporting Service crop reporting survey system.

3. The third major study area at Purdue University is data handling and pattern recognition techniques. To study the remote sensing survey systems, enormous amounts of data must be processed and analyzed by extremely specialized techniques. Therefore, data handling and pattern recognition problems become a crucial part of a thorough investigation to determine the feasibility and practical applications of the system.

Initial efforts in this phase of the research, primarily using multispectral scanner data obtained at three different times during the growing season, are being directed toward the following classification tasks, using various pattern recognition techniques:

· Wheat vs everything else

· Oats vs everything else

· Bare soil vs vegetation vs water

· Percentage ground cover in a given bare soil/vegetated area

· Corn vs soybeans

· Soybeans vs everything else

· Alfalfa vs everything else

In conjunction with these efforts, studies will be carried out to determine which spectral bands are the most useful for each task.