SECTION 38

USER ORIENTED DATA PROCESSING AT THE UNIVERSITY OF MICHIGAN

by

Frederick J. Thomson
Willow Run Laboratories
The University of Michigan
Ann Arbor, Michigan

INTRODUCTION

Half of the effort at The University of Michigan is directed towards data processing for various user application. This work has proven valuable in a number of directions. First, it has given us an opportunity to exploit the techniques of multispectral data processing that have been developed at Michigan on a wide variety of problems. Secondly, and perhaps more important, we have learned a great deal about types of problems that users encounter in interpreting multispectral data, and the knowledge gained in collaborating with the users has been a prime importance in directing the type of multispectral research in which we engage. The range of user applications have been large and in this past year we have dealt with 11 different users on quite dissimilar problems. In this paper I will not be able to cover in detail all of these investigations but I will discuss some that illustrate the general kinds of work that we have done at Michigan in the past year and the directions that user data processing of multispectral data are taking today. The remaining investigations are of equal interest and have been presented by various users at this meeting.

I would like to cover four scientific aspects of user data processing first: Signature extension and preprocessing, change detection, area measurement and geometric fidelity, and underwater feature recognition. To conclude I would like to summarize the kinds of processing services which we have developed at The University of Michigan and present a brief summary of the quantity of processing done during this past year.

SIGNATURE EXTENSION AND PREPROCESSING

Signature extension is a term which refers to the extension of the utility of a signature derived from a single training set to classify areas of the scene far removed from the training set. Because of the assumption in most processing that the spectral radiance signature of an object is unique, changes in solar irradiance and other effects limit the scene area which can be processed with a given set of signatures. Signature extension is important to processing studies because frequently training is the only practical alternative to signature extension, the training phase of processing is time consuming, and the amount of data that can be processed with a given set of signatures is now becoming of limited value to investigators interested in broad area surveys.

One method of implementing signature extension is preprocessing. There are three basic preprocessing techniques which can be used - taking ratios of channels, dividing by or subtracting functions of scan angle, or the use of ancillary data such as the sun sensor. This is summarized in Figure 1. Various ratio preprocessing transforms which have been applied to data are shown in Figure 2. There are three basic types of ratio preprocessing transforms, the division of signals by the sum of all signals, the division of signals by adjacent signals, or dividing differences of signals by sums signals. Each of these ratio preprocessing techniques are able to correct for changes in irradiance and some are more effective than others when additive variations in path radiance are also present. In a paper by Dr. Smedes presented at this meeting, he discussed the application of ratio preprocessing transforms to a set of data collected in Yellowstone National Park in September 1967. The Yellowstone data were collected while flying in a canyon, so variations of solar irradiance occurred across the scan line. These variations prevented the signature from being extended even across the scan line , let alone along the flight line. Previous attempts to process this data ended in success only after as many as 12 training sets were used for each of 8 classes of material to be recognized to account for variation of irradiance with scan angle. The division by a function of scan angle preprocessing technique was also used for the Yellowstone data and its application is summarized in Figure 3. Figure 3 shows plots of the mean values and standard deviations of three materials as a function of scan angle for the Yellowstone data. Note that the mean value of each signature varies considerably across the scan line. Had we used only one signature from the left hand edge of the data, forest would have been misclassified as glacial till on the right hand edge of the data. By determining a function of scan angle which made the mean values of each category invariant with scan angle, and by dividing all data by this function, we arrived at the preprocessed data shown in Figure 4. In Figure 4, the preprocessed signatures are invariant with scan angle, thus suggesting that recognition using these signatures could be extended across the scan line. Preprocessing was carried out and reported by Dr. Smedes. The performance of the ratio preprocessing and the scan angle preprocessing were determined from analysis of recognition maps and found to be quite good. Most importantly, we used only 8 signatures to obtain each of these recognition maps, whereas previous work had used over 100 signatures to obtain maps of comparable quality. Digital processing time was cut by a factor of 12 through the use of only 1/12 as many signatures.

The third preprocessing technique, use of ancillary data, can be quite useful in compensating for variations of solar irradiance in long duration flights. An excellent example of the application of this technique will be discussed later in the North Dakota data processing discussion.

The digital recognition maps for the study performed by Dr. Smedes were printed in color. This was done by using different colored ribbons in our IBM 1401 page printer and running the paper through four times - once for each color. The maps are impressive because of the use of red, green, and blue in addition to black for portraying scene objects. This communicates effectively to the interpreter. They are also superior to previously hand colored versions because they can be produced by the computer rather than by human effort. An example of a color recognition map of Yellowstone data is shown in Figure 5. A legend also accompanies the figure.

AREA COUNTS AND CHANGE DETECTION

Both area counts and change detection are important to multispectral processing because they attempt to extract more useable information from the recognition process. For many applications, the percentage composition of the scene or area of a particular material maybe as important or more important to a user than the recognition map itself. For example in estimating the water resources of Everglades National Park, the first step might be to make a recognition map of aquatic plant species. By measuring the area of these plant species and knowing what range of water depth they grow in, an estimate of water volume can be made.

Similarly, change detection is important for many problems. By making recognition maps from data collected at two different times and comparing the maps, changes can be detected. Area counts help quantify the changes. Both change detection and area counts play important roles in the decision making process using remote sensing data as an input.

Both concepts are well illustrated by an extremely successful study performed for the Bureau of Sport Fisheries and Wildlife in Jamestown, North Dakota. The general problem is to estimate the migratory waterfowl productivity of the Great Plains area (including Canada) so that meaningful management procedures can be used. One of the ways of managing the migratory waterfowl population is through the setting of hunting limits and season. Hunting limits are set through consultation with United States and Canadian authorities in June or July of each year.

Estimates of waterfowl productivity are currently made by noting size, shape and number of ponds in sampling areas. This year, we used remote sensing techniques to obtain this estimate for one test area. We feel that our technique yielded more accurate data, at least as fast as the previous airborne observer techniques. For this study, we used two samples of data. Data in May was used to assess water quantity important to nesting. Data in July determine those ponds productive of food. Areas of ponds and their perimeters are important measurements which allow the estimation of waterfowl. Areas are important because of the obvious relation to water quantity. Perimeter is important because nesting waterfowl require privacy. A pond of large perimeter will thus support more waterfowl than one of small perimeter, even if the pond areas are the same.

We automatically mapped ponds, their areas and their perimeters on May and July data sets collected at 4500 ft. 1.5 to 1.8 micron data were processed digitally. To obtain adequate pond recognition over the 120 mile flight line, preprocessing of the data was required. In this case, we divided all signals by the value of the sun sensor which measured ambient irradiance. This technique produced useable recognition of ponds with only one training set for the entire 120 mile run. The analysis showed that had preprocessing not been done, frequent training would have been necessary which would have slowed down the recognition process. Statistics and recognition maps were generated automatically by the computer and were in the hands of the managers within a month after the data were collected. Thus, they were well in time to influence this year's management effort.

Figure 6 and 7 show recognition maps of the same area made in May and July respectively. The two slides reveal that numerous small ponds present in May have dried up in July, and thus are not productive of food. To obtain a quantitative measure of water reduction, refer to Figures 8 and 9. These figures show areas, perimeters, and shape factors which are ratios of perimeters to square root of area normalized to one for a circle, are shown for each pond for May and July. This information was generated by the digital computer for each one mile segment of the 120 mile test area. At the end of each mile of data, a set of area and perimeter statistics were also automatically generated by the computer.

GEOMETRIC FIDELITY

There is a great potential for map generation using multispectral scanner data and automatic pattern recognition techniques. However, one important limitation is that scanner data is distorted. The scale is neither constant with the scan angle nor rectilinear. If the true potential of scanner data as a map source is to be realized, distortions must be reduced to acceptable levels. Fortunately, because the scanner data are recorded on tape, we have control over the production of pictorial output and can make some corrections electronically. Distortions in scanner data arise from two sources - scanning geometry and the variation in aircraft speed, heading, altitude and attitude. Scanning geometry distortions arise because the scanner covers the ground at a constant angular rate, while in film reproduction the intensity modulated CRT beam is swept at a constant linear rate across the film. This combination yields an image which is compressed at the edges. Instead of a linear sweep on the CRT, we need a wave form proportional to the position of the instantaneous field of view on the ground. This waveform is tan θ where θ is scan angle relative to the nadir. We generate such a waveform and use it to deflect the CRT spot thus removing scanner geometry distortion. At present, very little can be done about aircraft heading and speed variations during a run although potential corrections could be made if these parameters were known. Aircraft pitch, roll, and yaw are other quantities which can distort the scanner imagery. Corrections for roll are made on board the aircraft. Yaw corrections are also necessary to produce quality maps. A need for yaw corrections arise when aircraft is "crabbed" to maintain a given flight heading while flying in a cross-wind. Under such conditions, the scanner does not scan perpendicular to the flight direction and square corners are distorted as shown in the top photo of Figure 10. The solution is to rotate the CRT trace with respect to the film so that the geometry corresponds to the flight line-scan line geometry under which the data were collected. An example of this yaw correction is shown in the bottom half of Figure 10. Note that the square road intersections of the town at the top of the figure are distorted in the uncorrected video and correctly portrayed in the yaw corrected video. While it is not obvious from the video itself, both sets of data have had scan geometry distortions removed through the use of the tan θ sweep waveform and thus possess constant scale regardless of position in scene.

UNDERWATER FEATURE RECOGNITION

One of the major advances in the application of pattern recognition techniques to user data processing was the recognition of features on the bottom of Biscayne Bay. In this study, data collected at 6500 ft on 10 March 1970 over Biscayne Bay were processed to map classes of underwater vegetation likely to be affected by a thermal outfall from a power plant. This work has already been reported by Dr. Kolipinski of U.S. Geological Survey, Miami. What makes this study significant is that not only 6500 ft of air, but 10 to 15 ft of water lay between the objects being mapped and the aircraft. Apparently successful recognition of 8 categories were achieved, but only after the reflected energy from the water surface was removed by preprocessing. The reflected energy from the scene consists of reflection from underwater vegetation and from the water surface. The water surface reflection varied with scan angle as a result of the relative position of the sun and aircraft. This variation was removed by subtracting a function of scan angle determined by digital computer analysis. The optimum spectrometer channels for classifying the underwater vegetation were also determined from digital computer analysis. These channels were located in the blue, blue-green and orange regions of the spectrum, where the water is most transparent. Near infrared channels useful for classifying terrestrial vegetation were useless for this study because the water is opaque at near infrared wavelengths. More work is required to make the recognition insensitive to water depths, or equivalently to extend vegetation signatures from shallow to deep water. This problem was averted in the present study by taking test areas in areas of uniform water depth.

QUANTITATIVE ASPECTS OF THE USER DATA PROCESSING PROGRAM

At the University of Michigan we have developed several techniques for analyzing multispectral data. We have classed these as Type I and Type II techniques, and these are summarized in Figures 11 and 12. Type I techniques are generally simple processing techniques applied to single channels of the multispectral data. The Type II processing techniques are more sophisticated processing techniques and are related to applications of pattern recognition to the data. Figures 13 and 14 illustrate the quantity of work completed during fiscal year 1970.

A list of users and applications is shown below:

User	Organization	Application
Orlo Crosby	USGS, Bismarck, North Dakota	Survey of Water Temperatures in Missouri River
A. E. Coker	USGS, Tampa, Florida	Survey of Phosphate Pollution in Tampa, Florida area
Milton Kolipinski	USGS, Miami, Florida	Survey of Underwater Vegetation Features in Biscayne Bay
Kenneth Watson	USGS, Denver, Colorado	Reflectance Measurements and Classification of Rock Types
Herbert Hamm	Bureau of Reclamation, Denver, Colorado	Survey of Saline Affected Areas in Irrigation Project in Moses Lake, Washington
William Pearcy	Oregon State U.	Studies of Ocean Temperature and Color Indicative of Nutrients and Fish.
Harry Smedes	USGS, Denver, Colorado	Classification of Glacial Terrain in Yellowstone National Park
Harvey Nelson	Bureau of Sport Fisheries & Wildlife Jamestown, N.D.	Survey of Wetlands for Waterfowl Production

Philip Weber	US Forest Service, Berkeley, Calif.	Recognition of Ponderosa Pine Beetle Attacked Trees in the Black Hills and Diseased Douglas Fir Trees in Oregon
Charles Olson	University of Mich. Ann Arbor, Mich.	Classification of Different Forest Species and Delineation of Root Rot Disease in Pine.
Robert Colwell	US Forest Service, Berkeley, Calif.	Classification of Forest and Brush Species at Bucks Lake, California

CONCLUSION

The multispectral techniques have shown themselves capable of solving problems in a large number of user areas. The results obtained are in some instances quite impressive. As we review the picture today, we can make the following observations. In many instances, the multispectral detection of various phenomena is an empirical fact for which we have no or very little physical explanation today. It is fairly clear that in many instances with just a little more work, we should be able to put the detection process on a firmer intellectual foundation than we are able to do right now. To date, most of the user applications that have been addressed are exploratory in nature. The closest approximation to an operational situation that we have encountered so far is that of the Survey of Wetlands in North Dakota reported in this paper. It is very clear that a major step is the development of operational procedures for exploiting the multispectral techniques. The best starting point for this operational exploitation is in those areas where we have, of course, some experimental verification of the technique.

The signature extension techniques that have been developed have shown themselves to be useful in many instances. We should like to be able to report that the problem of signature extension is no longer with us, but we cannot. Progress has been made but it certainly is not complete. Future efforts should then consider unexplored areas mentioned above. First, we should begin to address very seriously using multispectral remotely sensed data for operational purposes and secondly, the problem of signature extension still remains with us, and while progress has been made, we must continue to work on this problem.

1. Channel Ratio Techniques
2. Divide by or Subtract a Function of Scan Angle
3. Use Ancillary Data (i.e. Sun Sensor)

Figure 1. Preprocessing Transforms

1. $\frac{X_1}{\sum_{j=1}^{12} X_j}, \quad \frac{X_2}{\sum_{j=1}^{12} X_j}, \quad \cdots\cdots \quad \frac{X_{11}}{\sum_{j=1}^{12} X_j}$

2. $\frac{X_2}{X_1}, \frac{X_3}{X_2}, \cdots\cdots, \frac{X_{12}}{X_{11}}$

3. $\frac{X_2 - X_1}{X_2 + X_1}, \frac{X_3 - X_2}{X_3 + X_2}, \cdots\cdots, \frac{X_{12} - X_{11}}{X_{12} + X_{11}}$

Figure 2. Ratio Preprocessing Transforms

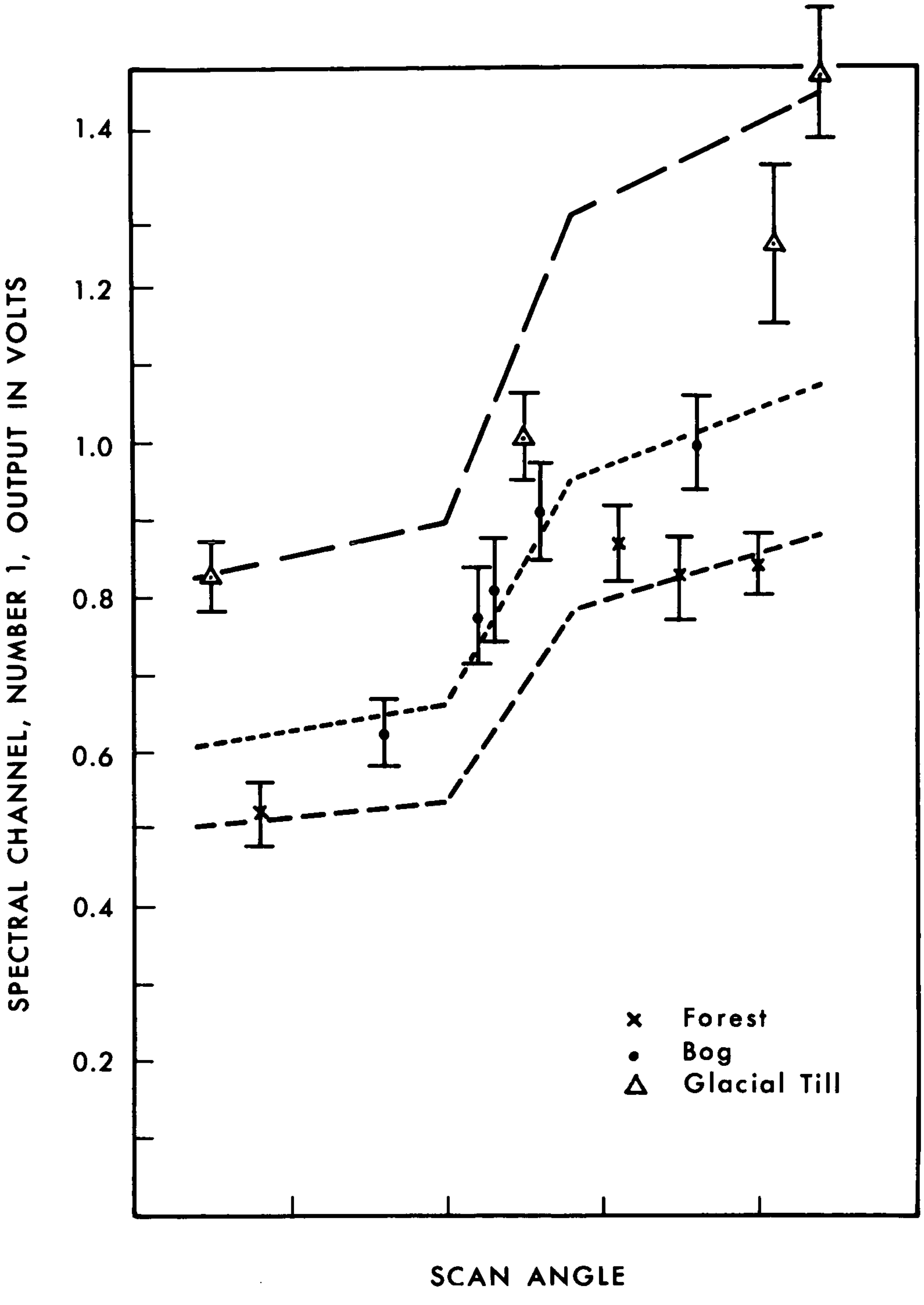


FIGURE 3. SPECTRAL CHANNEL OUTPUT VERSUS SCAN ANGLE FOR 3 MATERIALS SHOWING THE SCAN ANGLE FUNCTIONS

INFRARED AND OPTICS LABORATORY
INSTITUTE OF SCIENCE AND TECHNOLOGY
THE UNIVERSITY OF MICHIGAN

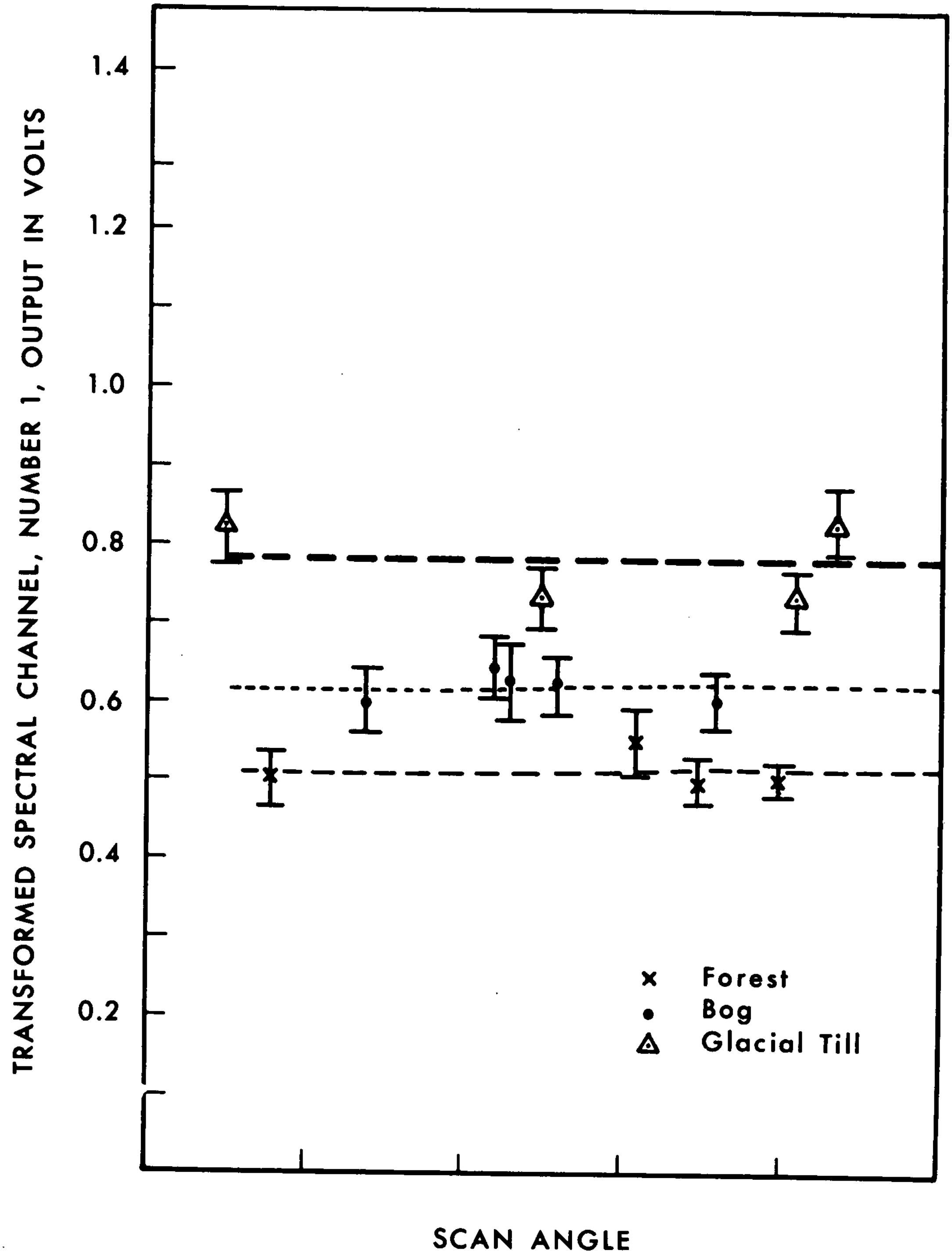
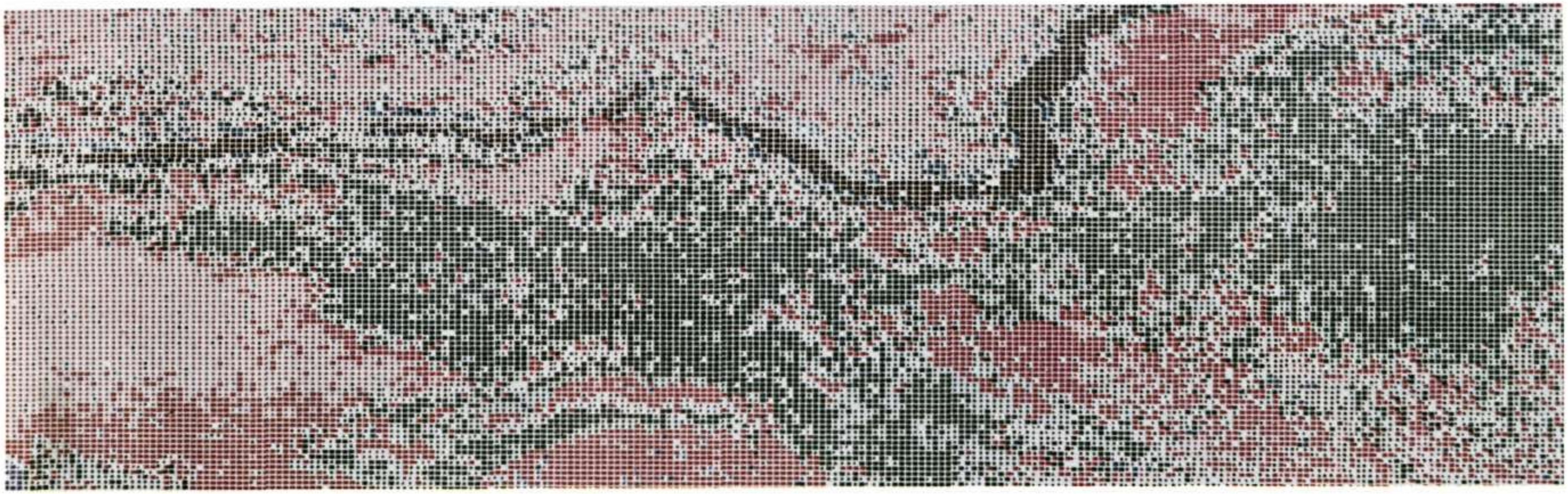


FIGURE 4. TRANSFORMED SPECTRAL CHANNEL OUTPUT VERSUS SCAN ANGLE FOR 3 MATERIALS SHOWING MEANS OF COMBINED SIGNATURES

INFRARED AND OPTICS LABORATORY
INSTITUTE OF SCIENCE AND TECHNOLOGY
THE UNIVERSITY OF MICHIGAN



COLOR CODE

Dark Blue.......Bedrock	Black..........Water
Light Blue......Vegetated Rock Rubble	Medium Grey....Talus
Dark Green......Forest	Light Grey.....Cloud
Light Green.....Bogs	Light Red......Kame
White...........Not Classified	Dark Red.......Till

Figure 5. Color Coded Digital Recognition Map of a Portion of Yellowstone National Park Data Collected 9/19/67 at 1430 hrs.

FIGURE 6. TYPICAL DIGITAL COMPUTER GRAYMAP OF PONDS. WOODWORTH STUDY AREA, NORTH DAKOTA. Location over Clark Lake and Big Lake. 31 JULY 1970, 1013 hrs., Flight Line 1, 2 KFT., 1.5 - 1.8 μm Data.

N

INFRARED AND OPTICS LABORATORY
INSTITUTE OF SCIENCE AND TECHNOLOGY
WRL
THE UNIVERSITY OF MICHIGAN

FIGURE 7. **TYPICAL DIGITAL COMPUTER GRAYMAP OF PONDS. WOODWORTH STUDY AREA, NORTH DAKOTA.** Location over Clark Lake and Big Lake. 22 MAY 1970, 1105 hrs., Flight Line 1, 2 KFT., 1.5 - 1.8 μm Data.

N

INFRARED AND OPTICS LABORATORY
INSTITUTE OF SCIENCE AND TECHNOLOGY
WRL
THE UNIVERSITY OF MICHIGAN

TYPICAL DIGITAL COMPUTER PRINTOUT OF POND STATISTICS. WOODWORTH STUDY AREA, NORTH DAKOTA.

ALTITUDE = 2100.FT VELOCITY = 207.FT/SEC SCAN RATE = 65.5 RPS RESOLUTION = 5.35 MRAD ONE SCANLINE IN 2. DIGITIZED NADIR AT POINT NO. 189 POINTS COUNTED IF VOLTAGE LIES BETWEEN 0 AND 1.20 VOLTS NB = 0 NE = 1 SMTH = 5.62 FEET

SCAN LINE	POINT	AREA (ACRES)	PERIMETER (FEET)	SHAPE
18414	150	.119	306.734	1.063
18461	162	.279	418.794	.950
18475	192	.052	259.766	1.362
18487	231	.116	272.230	.959
18515	150	.328	456.598	.956
18563	190	.121	340.322	1.174
18565	172	.481	603.448	1.043
18569	226	1.528	1493.685	1.447
18620	130	.050	169.165	.904
18721	162	3.418	2254.672	1.461
18733	179	.067	239.502	1.107
18768	172	.072	271.878	1.213
18774	135	.502	584.780	.988
18815	248	10.515	6193.072	2.288
18872	135	.569	620.947	.986
18875	170	.730	932.161	1.307
18914	198	.360	494.842	.988
18929	130	.757	805.891	1.109
18998	239	.496	629.921	1.071

Computer Statistics from May 1970 Data

Figure 8

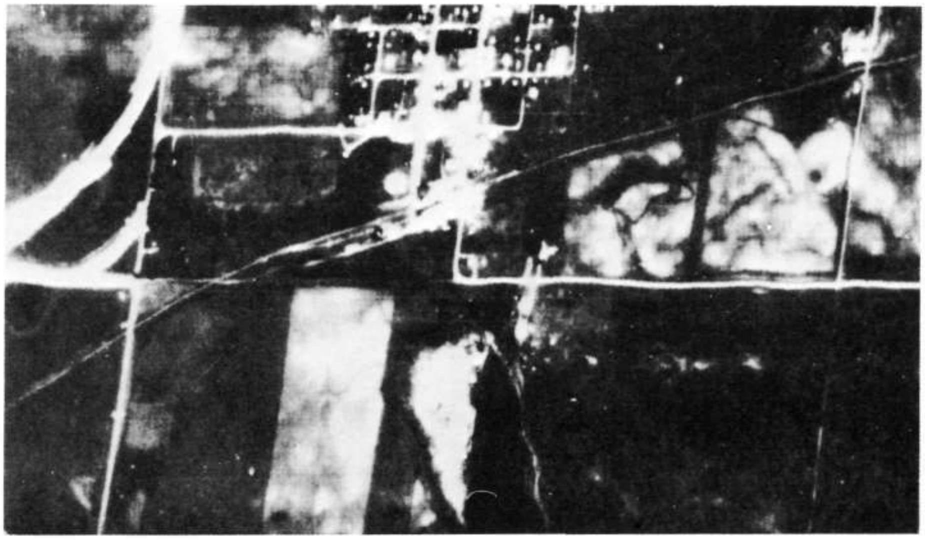
TYPICAL DIGITAL COMPUTER PRINTOUT OF POND STATISTICS. WOODWORTH STUDY AREA, NORTH DAKOTA.

ALTITUDE = 2015.FT VELOCITY = 204.FT/SEC SCAN RATE = 65.6 RPS RESOLUTION = 5.15 MRAD ONE SCANLINE IN 2. DIGITIZED NADIR AT POINT NO. 136 POINTS COUNTED IF VOLTAGE LIES BETWEEN 0 AND 2.00 VOLTS NB = 1 NE = 1 SMTH = 53.84 FEET

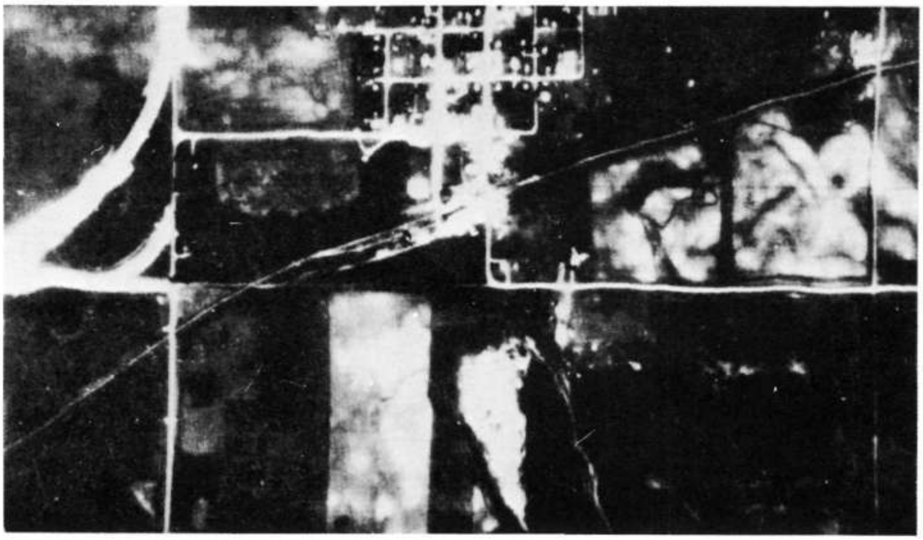
SCAN LINE	POINT	AREA (ACRES)	PERIMETER (FEET)	SHAPE
16256	203	.108	258.012	.941
16338	161	.055	237.148	1.215
16343	144	.269	443.305	1.024
16454	140	1.226	1055.163	1.142
16554	221	10.295	6011.876	2.241
16655	103	.356	557.124	1.118
16705	103	.510	604.893	1.015
16881	61	.210	403.672	1.053
17020	127	41.618	9186.915	1.706

Computer Statistics from July 1970 Data

Figure 9



(a) Video with Yawed Geometry



(b) Video with Yaw Corrected Geometry

COMPARISON OF TYPICAL YAWED AND YAW CORRECTED DATA

N

INFRARED AND OPTICS LABORATORY
INSTITUTE OF SCIENCE AND TECHNOLOGY
WRL
THE UNIVERSITY OF MICHIGAN

FIGURE 10

Imagery

Contouring and Quantization

False Color Films

Digital Tape

Canvas Panel Reflectance Measurements

Figure 11. TYPE I Processing Services

Optimum Channel Analysis

Preprocessing Analysis

Digital Recognition Map

SPARC Recognition Map

Signature Extraction

Figure 12. TYPE II Processing Services

Imagery ——— 2500 mi. Data

Thermal Contouring and Quantization ——— 10 jobs

Digital Tapes ——— 1 job

Canvas Panel Reflectance Measurements ———2 jobs

Figure 13. TYPE I Work Completed - FY70

Optimum Channel Analysis——— 11 jobs

Preprocessing Analysis ——— 6 jobs

Digital Recognition Map ——— 8 jobs

SPARC Recognition Map ——— 11 jobs

Signature Extraction ——— 2 jobs

Figure 14. TYPE II Work Completed - FY70