

Reprinted from

Symposium on

Machine Processing of

Remotely Sensed Data

June 27 - 29, 1979

The Laboratory for Applications of
Remote Sensing

Purdue University
West Lafayette
Indiana 47907 USA

IEEE Catalog No.
79CH1430-8 MPRSD

Copyright © 1979 IEEE
The Institute of Electrical and Electronics Engineers, Inc.

Copyright © 2004 IEEE. This material is provided with permission of the IEEE. Such permission of the IEEE does not in any way imply IEEE endorsement of any of the products or services of the Purdue Research Foundation/University. Internal or personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution must be obtained from the IEEE by writing to pubs-permissions@ieee.org.

By choosing to view this document, you agree to all provisions of the copyright laws protecting it.

IDENTIFICATION OF SURFACE-DISTURBED FEATURES THROUGH ISURSL NON-PARAMETRIC ANALYSIS OF LANDSAT MSS DATA

LEONARD H. ALGER, PAUL W. MAUSEL, AND
ROBERT R. HERNER

Indiana State University

I. ABSTRACT

In response to recent state and federal legislative mandates, the Indiana State University Remote Sensing Laboratory (ISURSL) has initiated a research program applied to evaluation of coal strip mine features in Indiana, Illinois, and Ohio using machine-assisted processing of Landsat MSS data. Specifically, two large strip mines in western Indiana were analyzed implementing both supervised and unsupervised non-parametric classification algorithms which were partially or totally developed at ISURSL. Nine classes of strip mine features were identified which included bare mine spoil, revegetated mine spoil, and water features in various physical states. An estimation of accuracy was made through comparison of the Landsat classification results with 1/30,000 scale aerial photographs taken the same day as the Landsat pass. Class accuracies ranged from 73% to 96% with an overall accuracy of 85%. The non-parametric approaches to classification used at ISURSL provide coal strip mine feature information of comparable quality to that generated by commonly used parametric classification systems, but they require as little as one-fourth the computer time for analysis.

II. INTRODUCTION

The need for acquisition of information about surficial modification resulting from surface mining of coal is increasing rapidly because of: (1) expansion of mining and (2) enactment of legislation whose mandates require accurate and frequent data for their successful fulfillment (federal and state laws such as the Surface Mining and Control Reclamation Act of 1977 (P.L. 95-87); An Act Regulating Surface Mining of Coal, Clay, and Shale-As Amended by Acts of 1977 (Indiana Code 13-4-6)).^{3,7} Such legislation encourages

the use of remote sensing in providing a methodology for the data acquisition requirements mandated by the laws.

Among the categories of surficial information needed by coal companies to satisfy laws associated with strip mining are the distribution and characteristics of: (1) graded bare mine spoil; (2) ungraded bare mine spoil; (3) types of revegetated mine spoil; (4) strip mine water features; and (5) land cover features bordering surface disturbed areas.

Visual interpretation of aerial photographs and machine-assisted processing of optical-mechanical scanner data have been used to gather large quantities of useful information from surface disturbed lands. For example, the analysis of Landsat data through machine-assisted processing techniques, such as LARSYS, has proven to be an effective method for gathering many types of information from strip mine environments.⁹ Though the resolution of LANDSAT restricts the detail of surficial mapping, general categories of earth surface features can be accurately identified.

The ability to acquire data effectively through remote sensing is but one of many considerations for potential users. It is becoming apparent that users of remotely sensed data are considering cost of analysis very carefully. Obviously factors such as the quality, format, and timeliness of data, together with the speed of analysis, are among those factors which must be satisfied to meet the specifications of a project contract. But once these specifications are met then the factor of cost becomes exceedingly important. With the variety of specific remote sensing techniques now available, it is possible for potential users of remotely sensed data to consider the most economical alternative approach to data acquisition.

Several suitable analytical approaches are available for classifying MSS data collected from surface disturbed lands. This paper explores the effectiveness of a low-cost non-parametric classification approach used and partially developed at the Indiana State University Remote Sensing Laboratory (ISURSL). The methodology was applied in the analysis of southwestern Indiana coal strip mines using Landsat MSS data.^{4,6}

III. THE STUDY AREA

Landsat scene 507715362 (July 5, 1975) reveals several large surface coal mines in southwestern Indiana. These mines have features characteristic of the most recently mined areas and contain areas in various stages of reclamation.

Officials from one of the larger coal companies operating the mines were contacted and permission was gained to use them (Chinook and Minnehaha Mines in Clay and Sullivan Counties, Indiana) as sites for investigating the applicability of the non-parametric approach used by ISURSL for machine-assisted processing of multispectral data to identify coal strip mine features (Figure 1). These officials showed interest in the project and provided ISURSL with high quality black and white and color photography (1/30,000) of the study areas (Figure 2). In addition access to the mining area and information from company experts were provided.

The research discussed in this paper emphasized single date Landsat data analyzed using ISURSL's machine-processing approach. The classes of features which were the focus of analysis were developed in consultation with mining experts and include: (1) unvegetated spoil-slightly graded; (2) unvegetated spoil-moderate to intensely graded; (3) grassy vegetation (25%-50%) on spoil; (4) grassy vegetation (50%-75%) on spoil; (5) grassy vegetation (>75%) on spoil; (6) forest and grass (>75%) on spoil; (7) forest (>75%) cover on spoil-dominantly coniferous; (8) slurry deposits; (9) wetlands-water; and (10) undisturbed land features.

IV. ISURSL ANALYTIC SEQUENCE

ISURSL has a remote terminal of The Laboratory for Application of Remote Sensing (LARS) and uses this facility for most of its machine-assisted analysis of spectral data. Twelve computer programs are available in the ISURSL approach to MS analysis. Seven of these programs were

used in analysis of the Chinook and Minnehaha Mines.⁶ The position in the analytical sequence, function, and characteristics of the most common programs used in one of the analysis sequences is summarized (Figure 3) to give an overview of a typical ISURSL approach to processing MSS data.

A. PICT (SPECTRAL MAPS)

The program PICT was used to establish general geographical orientation of strip mine features on the Landsat computer compatible tape (CCT). PICT used in its auto form displays relative spectral response values according to natural modes in the data (Figure 4). Similar results can be obtained by generating representative histograms of the study area through the HIST program and from that data manually set ranges (based on histogram modes) for spectral response classes. These ranges can be displayed on a computer map via PICT (manual). Either method initially attempts to identify earth surface features on a single band spectral map by the values which define modes in the representative histogram of the study area. Single band displays of relative spectral response data based on histogram analysis frequently highlight earth surface features which permit an analyst to more easily orient established ground positions with line and column coordinates designated on a CCT.

B. BIRTHA AND CCC CLUSTER ANALYSIS

The location of the study area established through the PICT program enabled further analysis using two non-parametric cluster programs designated BIRTHA and CCC (Figure 5). These cluster algorithms identify natural groupings of n-dimensional spectral data. BIRTHA is an ISURSL modification of Belur Dasarathy's HINDU algorithm and CCC is an ISURSL developed cluster classifier which uses spectral statistics generated by BIRTHA.^{1,6}

Specifically, BIRTHA analyzes large data sets in prototype format and constructs a multi-dimensional histogram. The terrain of the histogram contains every significant peak and valley in the feature space. Boundary locations in the terrain are identified. The decision process by which the boundary is located in the terrain is flexible due to the inclusion of several optional analytic techniques which include the following generally accepted decision methods:

1. Closest Cluster Centroid
2. Nearest Neighbor
3. Maximum Likelihood

4. Discriminant Hyperplane
5. Committee Approach of 2,3, and 4
6. Voter Approach, all of the above except 1.

CCC is a program which examines the cluster results generated through implementation of BIRTHA. It will, on the basis of closest centroid analysis of pixels within BIRTHA clusters, reassign pixels to different clusters to which they are more similar. CCC refines the BIRTHA clusters; generally 5-10% of the original assignment of pixels to BIRTHA clusters are reclassified through the implementation of this program.

The ten clusters delineated through BIRTHA and CCC were compared with black and white aerial photographs (1:30,000 scale and larger) to help extract earth surface feature information from the remotely sensed data. Spectral statistics in histogram and tabular form were generated and assisted in the interpretation of clusters.

C. HIST AND GRAPH PROGRAMS

The ISURSL cluster algorithms provided spectral data broadly representative of many features typically associated with strip mines. However, clustering did not identify all classes of interest, while some of the clusters with useful information required refinement before their data could be used in a supervised classification algorithm. The programs HIST and GRAPH are used in the ISURSL approach to provide smoothed or unsmoothed histograms (Figure 6) and linegraphs/columngraphs (Figure 7) of spectral responses characteristic of specific classes of earth surface features (training fields). These data are used to supplement cluster analysis or to provide spectral response characteristics of features unidentified through other methods.

D. REPRESENTATIVE SPECTRAL RESPONSE SELECTION

The selection of spectral responses which are representative of earth surface features of specific analytic interest can be made using vary degrees of automation. At ISURSL the various graphic outputs are interpreted through intensive examination of BIRTHA, CCC, HIST, and GRAPH data primarily using visual methods. This stage of analysis is the only one in which manual interpretation exceeds computer interpretation. The use of manual interpretation of computer generated data to develop representative relative spectral response statistics (ranges)

characteristic of classes of interest relies upon the ability of experienced analysts to make boundary decisions between classes that most benefit the research objectives. This ability to "fine-tune" a classification has proven to be effective in many instances where classification is difficult. As any option to this manual method, automated techniques can be used to help determine the spectral ranges which are submitted to LEVELSClassify, the ISURSL non-parametric classifier.

E. CLASSIFICATION USING LEVELSClassify

LEVELSClassify is related, but not identical, to the Eppler-look-up table approach.² Spectral statistics representative of each class of interest are put into LEVELSClassify in the form of relative spectral response ranges which define the dominant reflectance and emittance characteristics of a feature. Not only can the number of bands used to define a feature vary, but different band combinations (including an ISURSL ratio approach) can be used for classifying different earth surface features in the same study (Table 1). The spectral ranges developed are representative of designated earth surface features and specifically define each class, thus requiring no maximum likelihood calculations. The LEVELSClassify algorithm only requires a computer to identify those pixels which meet the spectral ranges specified for a given class.

F. DISPLAY OF CLASSIFICATION RESULTS

The classification results were originally displayed in alphanumeric form (1/24,000 scale) using the RESULT function of the ISURSL package (Figure 8). The results were examined to assess aspects of the quality and costs associated with classification of the specified surface disturbed and neighboring lands.

V. RESULTS

A. FEATURE IDENTIFICATION

Nine features associated with surface disturbed lands and a composite class of undisturbed land were identified in the south Chinook mine area using the ISURSL analytical sequence (Figure 8). In addition, subclasses of the water-wetland class were developed, but since they were not required in the project design they are not displayed or evaluated fully in this presentation. It is likely that identification of specific crops can be

made, but the reclaimed land in the study area was dominated by planted grasses and trees. Also, it is usually possible to separate the area of active strip mining from recent bare mine spoil.

B. CLASSIFICATION QUALITY

Training fields of the classes of interest were acquired from both Chinook and Minnehaha Mines. Collectively these mines cover more than 10,000 acres and represent mining activities which span more than 40 years.

Evaluation of the results was accomplished through comparison of the classification results with enlarged 1/30,000 scale aerial photographs taken the same day as the Landsat data using a zoom transfer scope. Supplementary ground information acquired from on-site inspection of parts of the study area were incorporated to improve the quality of classification evaluation.

The classification accuracy evaluation of the southern portion of Chinook Mine is presented in Table 2. These results are representative of the classification accuracy of the remainder of Chinook and the much larger Minnehaha Mine. It is evident from Table 2 that the designated surface mine features were identified at a level which assures the generation of accurate information of great value to mine operators. Although accuracy characteristics for many subclasses of the basic mine features were not intensively pursued in this research, it is worth noting with regard to future studies that preliminary examination of subclasses (e.g. water quality characteristics) indicates that additional classes frequently can be identified using Landsat data processed by non-parametric techniques.

The classification results using the ISURSL approach in this study are of the same order of magnitude as most other sophisticated techniques which focus on machine-assisted processing of MSS data. Additional research of the type presented here will be required to determine whether this non-parametric approach ultimately proves itself to be better than, equal to, or of poorer quality than other methods of analyzing MSS data for interpretation of surface disturbed land information.

C. COST OF CLASSIFICATION

The fact that non-parametric algorithms generally are less expensive to implement than those with a parametric base for machine processing of multispec-

tral data is well documented.⁵ Less certain is the relative quality of performance of parametric and non-parametric classifiers under all possible applications encountered in remote sensing. It is likely that one classification approach will have the advantage over another for a specific task, but it is unlikely that any one approach always will be best under all circumstances.

The non-parametric approach used at ISURSL certainly provided accurate identification of features associated with coal strip mining. It is very likely that implementation of traditional parametric approaches which stress point by point classification would generate similar classification results. In instances where parametric and non-parametric approaches are likely to produce very similar results, then the cost of analysis should be one of the major factors which determines the use of a particular technique.

Two of the most expensive programs commonly used in machine classification of MSS data are (1) an iterative euclidean distance cluster processor and (2) gaussian maximum likelihood point by point classifier.⁸ The high quality and effectiveness of these algorithms in application is well known, as are their relatively high computer costs for analysis. The ISURSL approach uses a low cost, basically non-parametric, cluster processor and classification program. The cost difference between the two parametric classification programs indicated above and the non-parametric BIRTHA/CCC and LEVELSClassify is assessed for this surface disturbed land study.

The comparison made between two common parametric classifiers and two non-parametric classifiers was conducted for a 7,100 pixel area centered on the Chinook Mine. The unsupervised classifier, BIRTHA and CCC were used to cluster the entire 7,100 pixel area with 10 classes specified. The clustering of BIRTHA took 54 seconds of CPU and the implementation of CCC took an additional 37 seconds. An iterative euclidean distance cluster processor was applied to the same area with a 10 class designation. This unsupervised classification program required 491 seconds of CPU. It is probable that the euclidean distance cluster program would not be used routinely to cluster 7,100 pixels in the analytical sequence of which the unsupervised classifier is a part. However using BIRTHA/CCC, a 7,100 pixel clustering (or much larger) is economically feasible even on projects with rather severe fiscal limitations.

The actual classification of 7,100 pixels in the Chinook Mine area for 10 classes of earth surface features using LEVELSClassify was 29 seconds of CPU. The gaussian maximum likelihood point by point classification program for a 10 class, 7,100 pixel classification was 86 seconds of CPU using the same computer system as used for LEVELSClassify.

This one comparison clearly indicates a great cost advantage of the ISURSL non-parametric approach over one of the more commonly used parametric approaches to machine-assisted processing of MSS data. The cost differential between the ISURSL approach and other types of parametric approaches such as LARS' ECHO is less than the example given, but nevertheless a cost advantage in favor of non-parametric approaches is evident.

VI. CONCLUSIONS

The non-parametric approach used at ISURSL proved to be effective in identifying nine basic coal strip mine features. Additional features (subclasses of the basic nine) were also identified, but they were not subjected to classification refinement or identification accuracy evaluation in this paper.

No specific comparisons of accuracy in identification of surface disturbed features were made using other parametric, non-parametric, or visual interpretation techniques, thus no claim of superior classification accuracy for the ISURSL approach can be made. However, the percent accuracy achieved through implementation of the ISURSL approach is of the same order of magnitude as that indicated in the remote sensing literature by analysts using other sophisticated machine-assisted techniques.

The non-parametric approach is characterized by frequent human interpretation during analysis, particularly in developing the spectral statistics used in LEVELSClassify. Interaction of this nature provides an analyst with good insight into the nature of the spectral data which frequently results in improved classification quality.

A distinct overall cost advantage of the non-parametric approach used at ISURSL over commonly used parametric methods was evident. This cost advantage can be particularly important for studying surface disturbed lands because of their great areal extent and their need for frequent monitoring. LEVELSClassify and

most programs ancillary to classification in the system used at ISURSL have modest computer requirements, thus making this approach feasible to use in remote sensing laboratories operating on limited funds via a mini-computer environment.

VII. REFERENCES

1. Dasarathy, B. V. 1976. HINDU: Histogram Inspired Neighborhood Discerning Unsupervised Pattern Recognition System. Computer Sciences Corporation, Huntsville, Alabama (MSFC), pp.1-42.
2. Eppler, W. B. 1974. An Improved Version of the Table Lookup Algorithm for Pattern Recognition. Proceedings of the 9th International Symposium on Remote Sensing of Environment. pp.793-812, University of Michigan, Ann Arbor.
3. Federal Register. October 25, 1978 (43FR49932), Part II, Vol. 43, No. 201. Abandoned Mine Land Reclamation Provisions. U.S. Govt. Printing Office, Washington, D.C.
4. Hyde, Richard F., Samuel N. Goward, and Paul W. Mausel. 1977. ISURSL Levels Classification: A Low Cost Approach to Multispectral Data Analysis. Symposium on Machine Processing of Remotely Sensed Data pp. 322-332, Purdue University, West Lafayette, Indiana.
5. Hung, Albert Y., 1977. A Least-Square Error Approach to Landsat Image Classification. Symposium on Machine Processing of Remotely Sensed Data. pp. 240-249, Purdue University, West Lafayette, Indiana.
6. Mausel, P. W. (ed.). 1978. ISURSL Remote Sensing Manual for Machine Processing of Multispectral Data. Indiana State University Remote Sensing Laboratory Indiana State, Terre Haute, Indiana.
7. State of Indiana (Indiana 100th General Assembly). May 3, 1977. P.L. No. 149, An Act to Amend IC 13:4-6 Concerning the Reclamation of Strip Mines. Vol. 1, pp. 676-681, Indiana Legislative Council, Central Publishing Company, Indianapolis, Indiana.
8. Swain, Philip H., 1972. Pattern Recognition: A Basis for Remote Sensing Data Analysis. Laboratory for Application of Remote Sensing, Purdue University, West Lafayette,

Indiana. LARS Information Note
111572.

9. Weismiller, R.A. and R.P. Mrocynski.
1978. Status of Derelict Land
Associated with Coal Mining in South-
western Indiana. Laboratory for
Applications of Remote Sensing,
Purdue University, West Lafayette,
Indiana. LARS Contract Report
022378.

VIII. ACKNOWLEDGEMENT

The authors wish to thank Norman
Coopridge for his assistance in photo-
graphy and cartography.

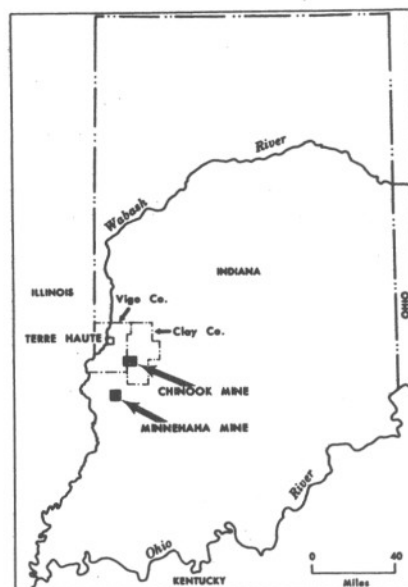


Figure 1. Study Area

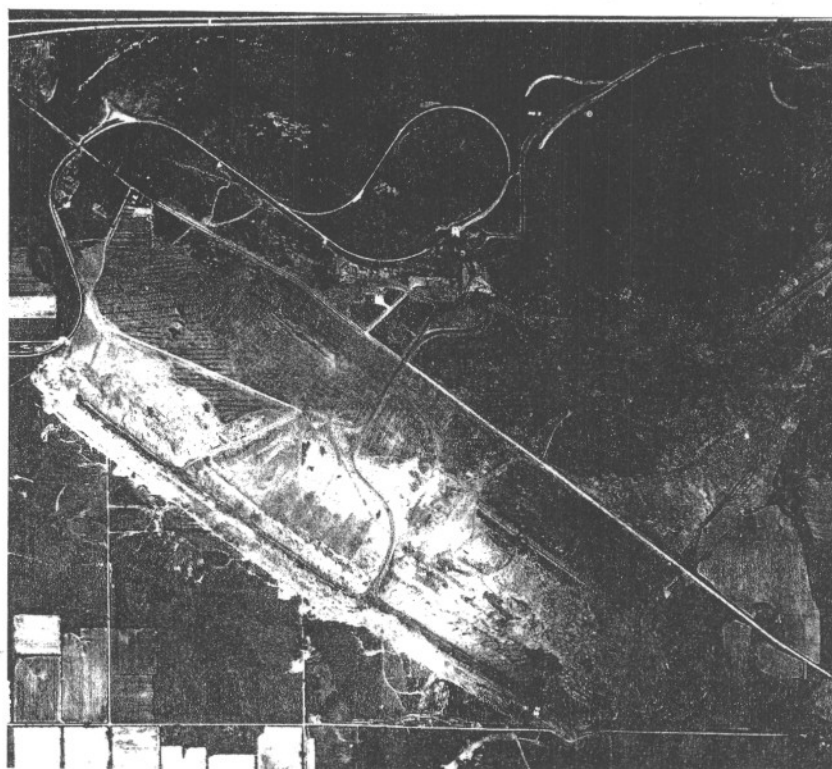


Figure 2. Black and White Aerial Photograph of the Southern
Chinook Mine Area, Clay County, Indiana. Photograph was taken on
July 5, 1975 at a scale of 1/30,000.

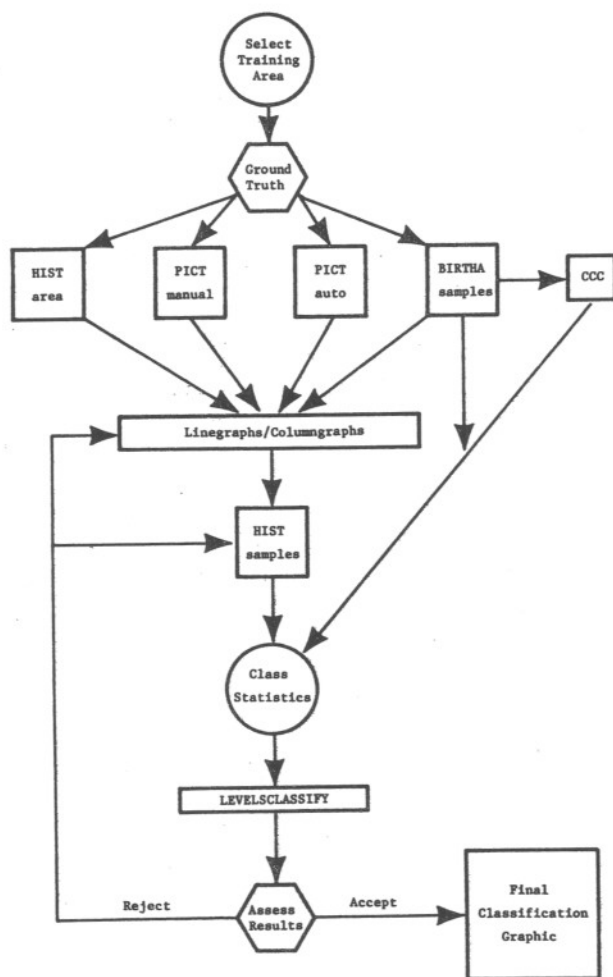
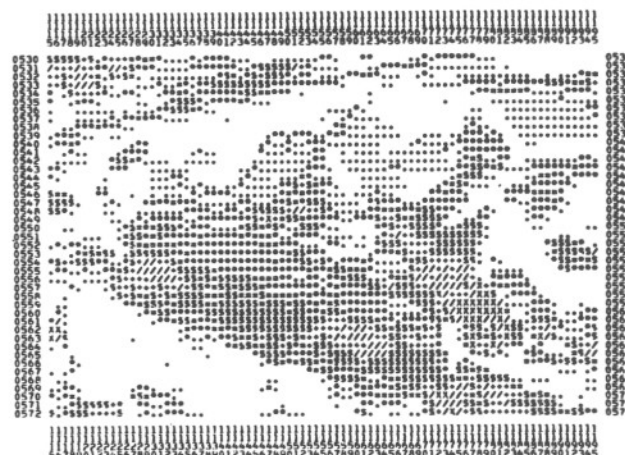


Figure 3. Flow Chart of ISURSL Procedures Used in Strip Mine Analysis.

CHANNEL 2

LOWER	UPPER	CHAN
0	27	2
28	37	3
38	41	4
42	46	5
48	53	6
54	59	7



4. Relative Spectral Response Map of the Southern Chinook Mine Area Developed by PICT. Landsat band 5(.6 - .7um) spectral response patterns are displayed.

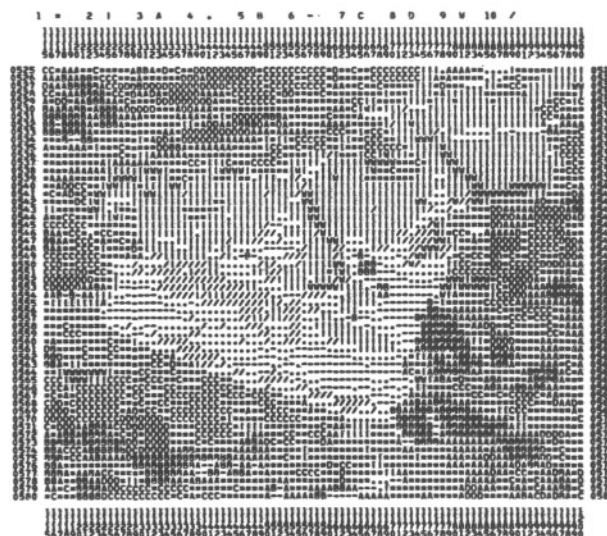


Figure 5. Cluster Map of the Southern Chinook Mine Area Developed by CCC. Four bands of Landsat MSS data were used for clustering.

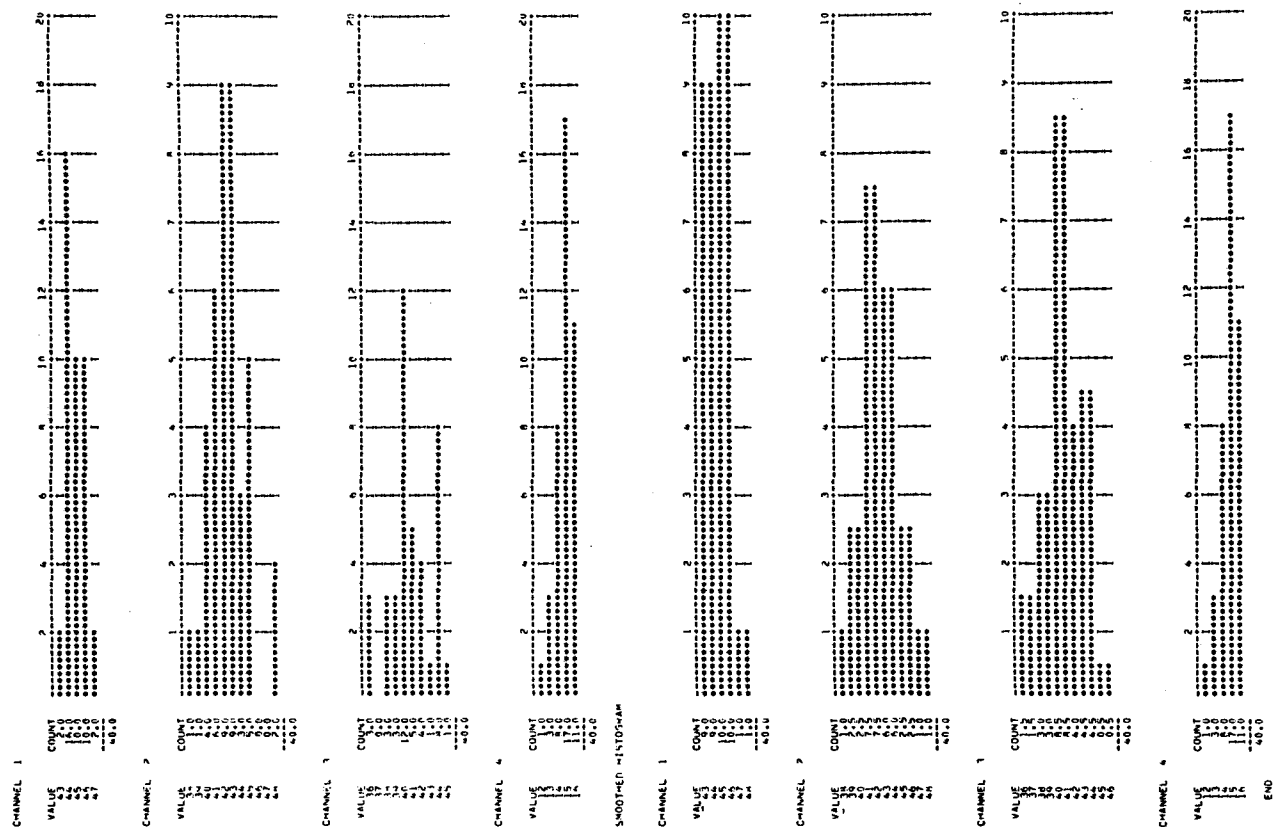


Figure 6. Unsmoothed and Smoothed Histograms of a Slightly Graded Spoil Training Sample in the Chinook Mine Developed through HIST. The first four histograms display spectral response distributions as they are stored in their decompressed form on the Landsat data tape (unsmoothed). The last set of histograms display the spectral response distributions of Landsat data in a compressed format which adjusts for the artificial values introduced into bands 4,5, and 6 on commercially available Landsat CCT's.

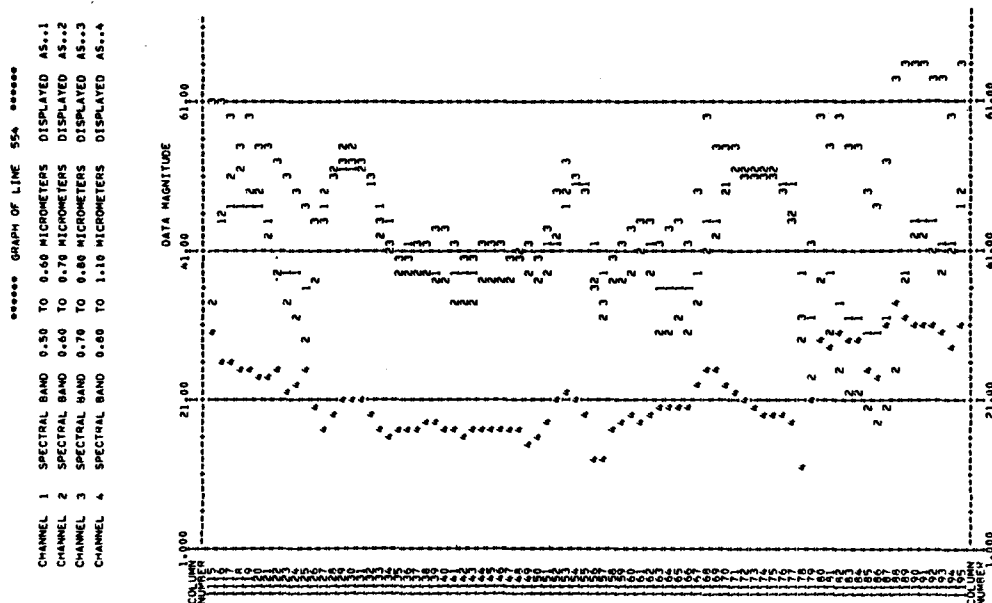


Figure 7. Four Band Relative Spectral Responses of Strip Mine Features in the Southern Chinook Mine Area Displayed by Line-graph using the GRAPH processor.

222
00011111111111222222222233333333333344444444445555555555666666666677777777778
78901234567890123456789012345678901234567890123456789012345678901234567890

Figure 8. Display of Classification Results in the Southern Chinook Mine Area. The ISURSL programs LEVELSCLASSIFY and RESULTS were implemented to provide the final classification results. Vegetated spoil may contain up to 25% vegetation cover.

Table 1. Representative Relative Spectral Response Characteristics Used in Western Indiana (Landsat Scene 507715362) Coal Strip Mine Research

Class	Landsat Band	Representative Spectral Response Ranges
Unvegetated Spoil-Intensely Graded	4	43-48
	5	38-48
	6	36-46
	7	11-16
Unvegetated Spoil-Slightly Graded	4	40-55
	5	39-57
	6	40-58
	7	15-24
Grassy Vegetation (25%-50%) on Spoil	4	34-49
	5	32-46
	6	34-48
	7	14-20
Grassy Vegetation (50%-75%) on Spoil	4	37-42
	5	23-32
	6	34-50
	7	15-23
Grassy Vegetation (>75%) on Spoil	4	29-36
	5	23-32
	6	34-50
	7	15-23
Forest and Grass (>75%) on Spoil	4	35-37
	5	17-31
	6	36-55
	7	17-33
Forest (>75% on Spoil-Dominantly Coniferous)	4	27-32
	5	17-31
	6	36-55
	7	17-33
Slurry Deposits	4	29-33
	5	21-24
	6	19-26
	7	7-9
Water-Wetlands	7	0-15

Table 2. Coal Strip Mine Feature Identification Accuracies For Southern Chinook Mine Area

Class	Area as Determined from ISURSL Analysis (Acres)	% Identification Accuracy
Unvegetated Spoil - Intensely Graded	163	78.5
Unvegetated Spoil-Slightly Graded	455	89.9
Grassy Vegetation (25%-50%) on Spoil	225	84.4
Grassy Vegetation (50%-75%) on Spoil	160	73.3
Grassy Vegetation (>75%) on Spoil	230	87.5
Forest and Grass (>75%) on Spoil	124	78.2
Forest (>75% on Spoil-Dominantly Coniferous)	107	90.7
Slurry Deposits	5	80.0
Water-Wetlands	97	95.9
	1566	
Overall Classification Accuracy		85.3

Leonard H. Alger (B.S., Northern Arizona University, M.A., Indiana State University) is a Ph.D. student in physical geography at Indiana State University. Currently he is senior analyst and operations director of the Indiana State University Remote Sensing Laboratory. Professional interests include remote sensing, technology transfer, arid lands geography, physical geography, traditional and automated cartography, and environmental analysis. He has a growing list of presentations and publications in both geography and remote sensing. Significant research includes investigating the potential for near ultraviolet remote sensing and research for contractual analysis conducted for governmental agencies.

Paul W. Mausel (Ph.D., University of North Carolina) is a professor of geography and is director of the Indiana State University Remote Sensing Laboratory (ISURSL). His research and teaching specialties are in multi-spectral remote sensing and its application to earth resources problems. Dr. Mausel has had numerous articles published in geography, remote sensing, planning, and soil science professional journals. He has directed research and educational programs for agencies which include the NSF, EPA, U.S. Forest Service, and State Planning Services Agency (Indiana).

Robert R. Herner (B.S., Indiana State University) is working toward his M.A. in geology at Indiana State University. Currently he is involved in several surface mine studies at the Indiana State University Remote Sensing Laboratory. Professional interests include remote sensing (MSS), air photo interpretation, geomorphology, and environmental analysis. He has assisted in teaching remote sensing (NSF high school honors program), photogeology, and geomorphology.