AN INTERFACE FOR REMOTE SENSING DIGITAL IMAGE SYSTEMS AND GEOGRAPHIC INFORMATION SYSTEMS

R. R. Irish and W. L. Myers
The Pennsylvania State University
University Park, Pennsylvania 16801, U.S.A.

1. INTRODUCTION

Rapid developments in computer technology and applications programs have made possible the successful classification of a variety of regional natural resource phenomena by computer analysis of remotely sensed data. The same developments have engendered sophisticated polygon-based geographic information systems (GIS) for handling environmental and natural resources data. However, these two technologies have evolved separately, and an absence of interfacing has resulted. This is attributable, in large part, to different system concepts in representing space. Nonetheless, it is possible to establish a linkage between the two.

2. POLYGON AND RASTER DATA STRUCTURES

Polygon-based GIS's employ vector data structures in representing space. Areal entities are geocoded as an aggregation of polygons, where each polygon represents a homogeneous area of a map. An area's boundaries are commonly encoded as a circuit of X-Y coordinates. Encoded polygons are accompanied by unique reference codes that identify the polygon and serve as relational links to records in tabular files stored in the data base. These files consist of polygon descriptors--attributes that depict an aspect of the encoded map. The polygons and codes are stored in the data base as a file of polygons and, when linked to a set of attributes, constitute a layer of information.

Common borders of neighboring polygons are redigitized when polygons are encoded independently. Redundancy may be avoided by encoding boundaries delimited by nodes (points where three or more lines meet) and independent reference points, with one point in each polygon. The boundaries and points are correlated using a chaining algorithm. This results in lists of line segments composed of right and left polygon identifiers (two nodes and any number of points). Attribute values may replace identifiers for subsequent analysis and display.

Remote sensing scanners generate data in a raster format. The earth's radiant flux is recorded in two dimensions as sensing optics repeatedly scan the earth's surface in a sweeping motion perpendicular to the platform's orbit path. Telemetered data are put onto a computer-compatible tape in the format of a digital image data set. The raster-structured digital data are a matrix of spectral reflectance values, where each row represents a scan line and each cell or pixel of the matrix is composed of a series of bytes, one for each wavelength as recorded by the scanner. The data set undergoes spectral pattern analysis, in which each pixel is assigned a symbol that identifies the earth surface category of which it is a constituent. This output is called a classified image file.

3. INFORMATION TRANSFER ALTERNATIVES

The desire to make use of pixel classifications in a GIS requires a means for making vector-based and raster data structures compatible. Either raster-to-vector or vector-to-raster conversion is necessary.

Raster-to-vector conversion results in a distinct polygonal layer for loading into the GIS data base. This may be accomplished by outlining feature category boundaries on a hard-copy pixel display and encoding the graphics. However, cleaning problems associated with table digitizers become part of the process, and automatic digitizers are expensive and scarce. The human becomes a decisive ingredient in an otherwise machine-oriented environment.

Raster-to-vector algorithms have been written for converting raster-based data to polygonal layers [Morehouse and Dutton 1980, Nichols 1982]. Implementation problems arise, however, due to substantial memory requirements placed on a computing system. Peripheral storage may also be limiting when deriving additional layers from successive data, and computing costs are high.

A second interfacing possibility involves relating classified pixels to an existing GIS polygonal layer. This is accomplished by rasterizing the layer into a classified image format. Once the rasterized version is created, it can be integrated with a classified image file by means of a digital overlay. This concept was used in developing a vector-raster interface that transfers classified image information to a GIS data base.

4. ZONAL INTERFACE

4.1 Approach

The developed interface relies on existing polygonal layers from a GIS and is entitled Zonation Algorithms (ZONAL). Ownership parcels, political boundaries, administrative subdivisions, forest management compartments, and other geographic layers comprise the spatial data base of a GIS. These predefined polygon files serve as the geobase for numerically overlaying the classified pixels.

The digital overlay is accomplished by first rasterizing a file of polygons. A computer-simulated scanner generates a grid cell representation of a polygonal layer--a spatial replica of the Landsat information. However, the information content differs. Instead of a spectral reflectance value, each newly created pixel has affixed to it a code that identifies the polygon within which the center of the pixel falls. This results in an indexing scene linking each image pixel with a polygon from the GIS (Figure 1).

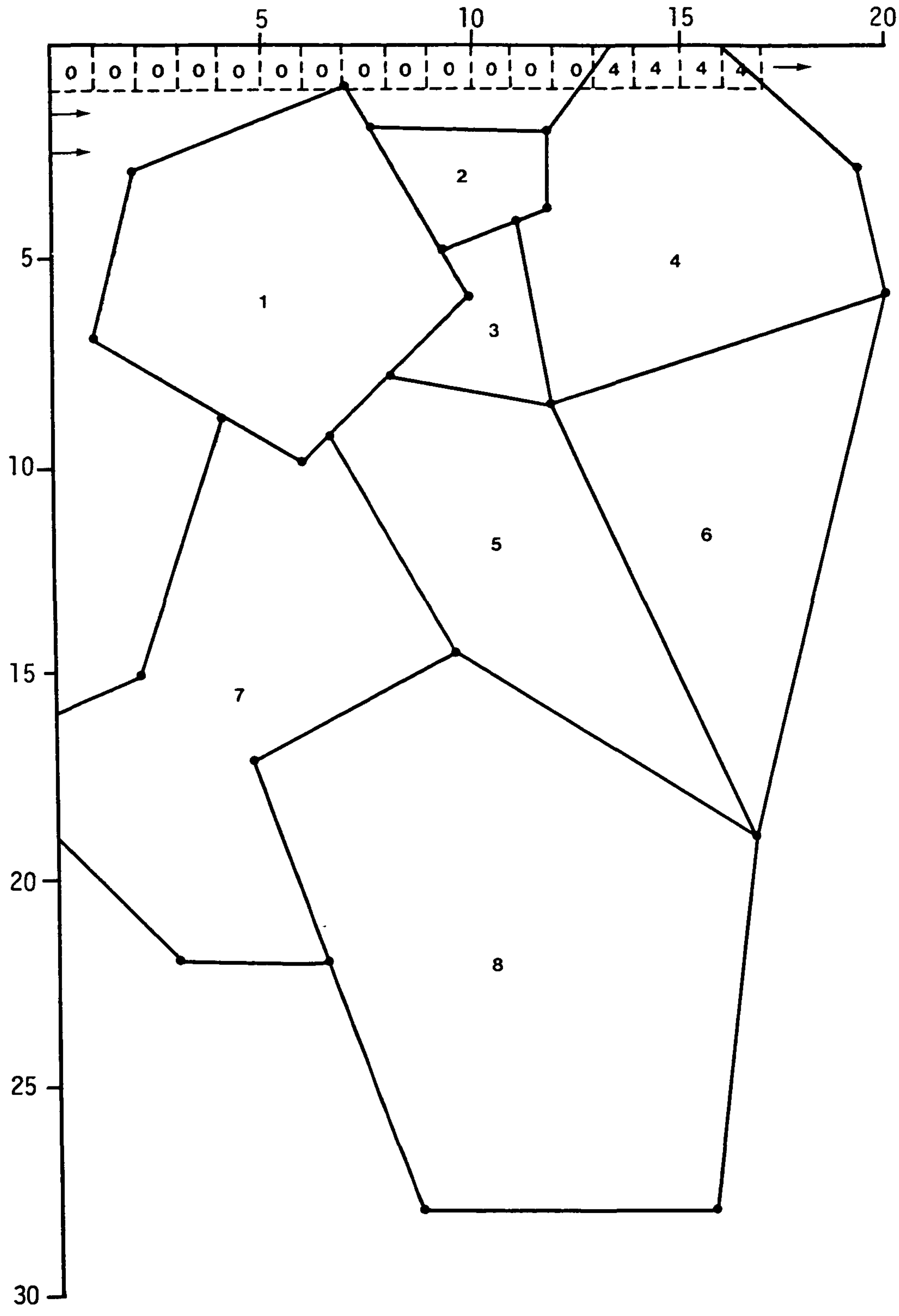

Figure 1. Illustration of how a polygonal layer is converted to an indexing scene. Polygon identifying codes are assigned to pixels on a scan line-by-scan line basis.

Classified pixels are tabulated within each polygon by simultaneously processing the classified image file with the indexing scene. The tabulation results take the form of frequency distributions depicting the number of classified pixels by feature category within each polygon. An analysis is performed on the frequency distributions, resulting in assignment of a feature category percentage figure or a label representing a co-occurrence of feature category percentages to each polygon. The resulting classification of polygons is formatted as either a numeric or non-numeric attribute file and, when linked to the polygonal layer, represents an additional layer of information. This file is transferred to the data base using the update facilities of the GIS.

4.2 Host System Requirements

The ZONAL interface requires that GIS spatial layers be encoded as independent polygons or chains. Independent polygons must be stored or extracted as layers without internal overlap. Chains comprising a polygonal layer must be retrieved and stored as a separate file. Also, all layers must be error free.

Several requirements are also placed on the remote sensing digital image system. Pixels generated by the image classifiers must be held in peripheral storage, because this file is used in the digital overlay. Due to variations in remote sensor altitude, attitude, and velocity, digital image data are not in positional agreement with polygon files in a GIS. Geometric correction facilities are necessary to ensure reliable indexing.

4.3 Interface Description

The ZONAL interface is composed of eight Fortran programs. Their relationships to the host systems are illustrated in Figure 2. A short description of each program follows:

POLSEG takes the polygon- or chain-based files as defined in a GIS and decomposes them into line segments.

ORDSEG orders the line segments into user-defined panels as a preprocessing step to rasterization. Ordering enhances the efficiency of pixel generation and reduces memory requirements for large polygon files.

GENPIX creates an indexing scene from the file of ordered line segments. Index pixels are generated on a panel-by-panel basis.

RIDPIX digitally overlays the classified image file with the indexing scene. RIDPIX determines the image data spatially coincident with the indexing scene and rewrites this registered subset, suppressing all non-polygon image information.

COMPIX overlays the registered image file with the indexing scene. Simultaneous processing results in frequency distributions depicting the number of classified pixels by feature category within each polygon. An output listing provides the polygon area covered by each classification category.

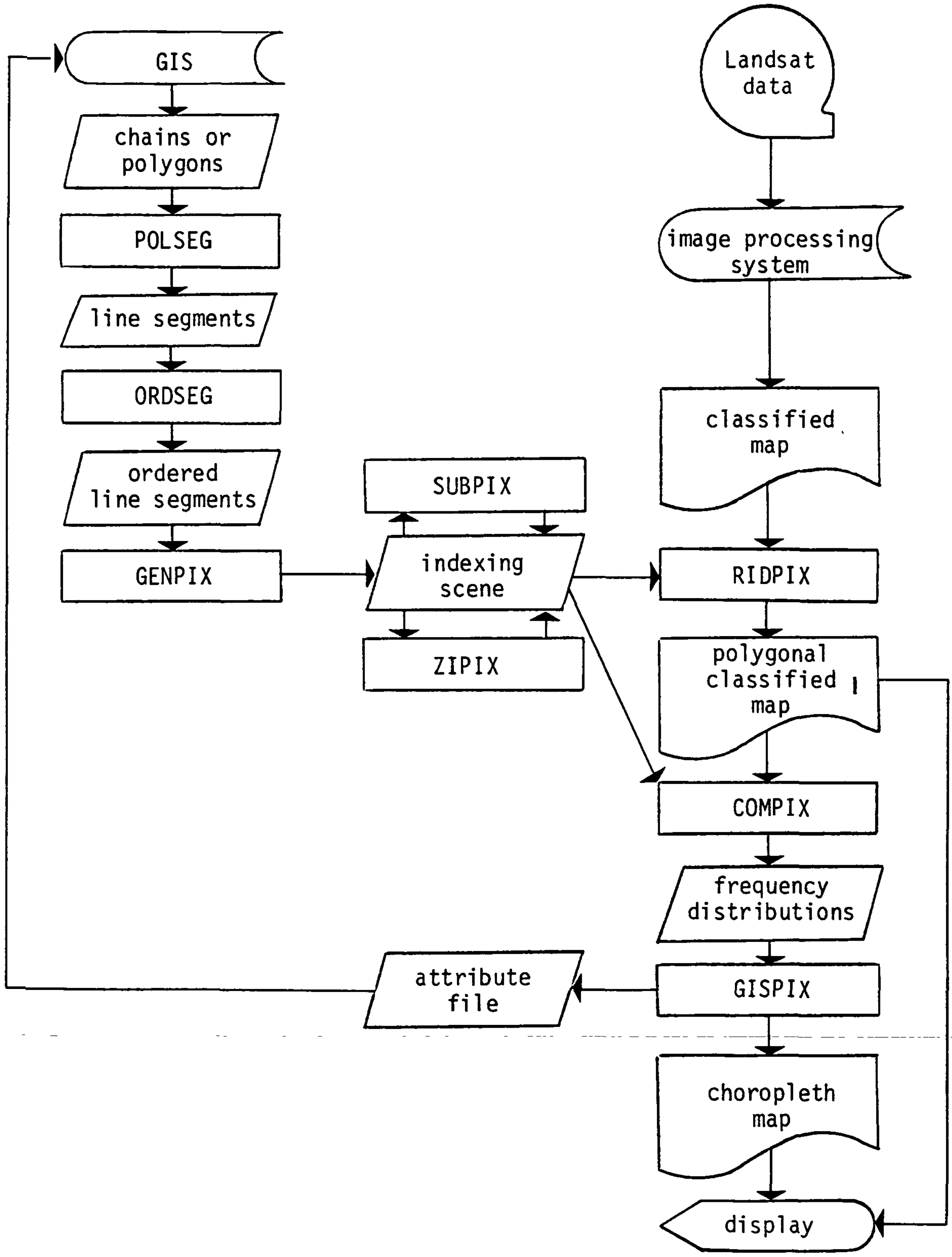

Figure 2. ZONAL flow chart.

GISPIX performs a secondary classification on the frequency distributions. Numeric or non-numeric attribute files are created by assigning a classification percentage or a label representing a co-occurrence of classification percentages to each polygon. An option generates a choropleth map.

SUBPIX extracts subareas from a larger indexing scene.

ZIPIX joins two indexing scenes along a common boundary to form one continuous indexing scene.

The ZONAL interface involves a multi-step operation. Six of the eight programs (POLSEG, ORDSEG, GENPIX, RIDPIX, COMPIX, and GISPIX) are necessary and must be used sequentially. SUBPIX and ZIPIX complement the interface by simplifying more complex indexing situations.

5. TEST CASE

To demonstrate the utility of ZONAL, a linkage was made between two host systems at The Pennsylvania State University. PSU's Experimental Forest, comprised of nine management blocks, was used as a study area.

A Task Oriented Multi-purpose Information System (TOMIS) is under development at PSU's School of Forest Resources [Myers 1982]. TOMIS, a polygon-based GIS, was designed to handle and analyze data associated with management- and research-related activities of experimental forests. Polygons are independently encoded as circuits of X-Y vertices. Attributes reside as either numeric or non-numeric descriptors and each is composed of two parts, an attribute type and value.

The digital image processing system developed by the Office for Remote Sensing of Earth Resources (ORSER) was used as the host image analyzer. A portion of a Landsat scene covering the Experimental Forest and scanned in May of 1973 was classified and geometrically corrected using the ORSER software [Turner et al. 1982]. Five land use and cover classes (water, coniferous forest, deciduous forest, senescent vegetation, and agricultural land) were defined. The senescent category comprises all areas of pre-leaf vegetation in the spring data set.

Nine polygons, representing the Experimental Forest's block boundaries, were assembled into and stored as a polygonal layer. The layer was decomposed by POLSEG into line segments. ORDSEG ordered the line segments within five panels. The indexing scene created by GENPIX involved two steps. First, the size of the raster file necessary to cover the management block layer at a given resolution (Landsat pixel) was determined. The panels were then processed sequentially, and each index pixel was assigned a code that identified the management block within which it fell.

At this point, the indexing scene and classified image file were processed simultaneously by RIDPIX. Based on a ground control point specified in terms of digitizer X-Y coordinates and image row and column positions, the subset of positionally coincident image pixels was determined from the parent file. The two were digitally overlaid, resulting in a rewritten, registered subset with all image information outside the management blocks suppressed. This file was displayed and registration accuracy visually verified.

The registered file and indexing scene were digitally overlaid by COMPIX. Frequency distributions depicting the number of classified pixels by feature category within each management block and an acreage listing were generated. These were useful in evaluating and comparing the land use and cover classes as they occurred in the blocks. However, worthwhile feature category relationships existed among the pixel summarizations which were concealed in a tabular format. GISPIX made detection of these possible by examining the frequency distributions and extracting both numeric and non-numeric TOMIS attributes. Non-numeric attributes resulted from a second analysis of the image data in which management blocks were classified by recognizing co-occurrences of pixel percentages.

The simplest attributes consisted of percentage values of a single feature category. A request was made for coniferous cover attributes. The symbol representing conifers, the range of acceptable percentage limits (0-100 percent), and an attribute type (CONIFERS) were specified. The frequency distributions were processed and the attribute file created. Each record consisted of the block's code, the CONIFERS attribute type, and the percentage of the block covered by coniferous forest. A similar request was made for forested cover attributes. The two symbols representing the forest categories, the valid ranges for each category (0-100 percent), and an attribute type (FORESTED) were specified. In this case, each attribute record consisted of the block's code, the FORESTED attribute type, and the percentage of the block covered by forest.

A non-numerical set of attributes depicting the nature of each block's forest cover was derived by polygon classification. Criteria were established for assigning one of three attribute values to the management blocks: DECIDUOUS, CONIFEROUS, or MIXED forest. Polygons were classified on the basis of the following set of criteria:

Range Limits (%)		Threshold	Attribute	Attribute
Deciduous	Coniferous	Percentage	Type	Value
0 - 9.9	65.1-100	75	FOREST	CONIFEROUS
65.1-100	0 - 9.9	75	FOREST	DECIDUOUS
10 -100	10 -100	75	FOREST	MIXED

Any blocks with less than 75 percent forest cover were not assigned a value. All blocks containing over 65 percent coniferous or deciduous forest were assigned the appropriate value, provided the remaining forest cover was less than 10 percent and the 75 percent threshold was met. All other blocks with over 10 percent coniferous forest were assigned the MIXED forest value if the total forest cover was at least 75 percent. The decision regions and results of classification are portrayed in Figure 3.

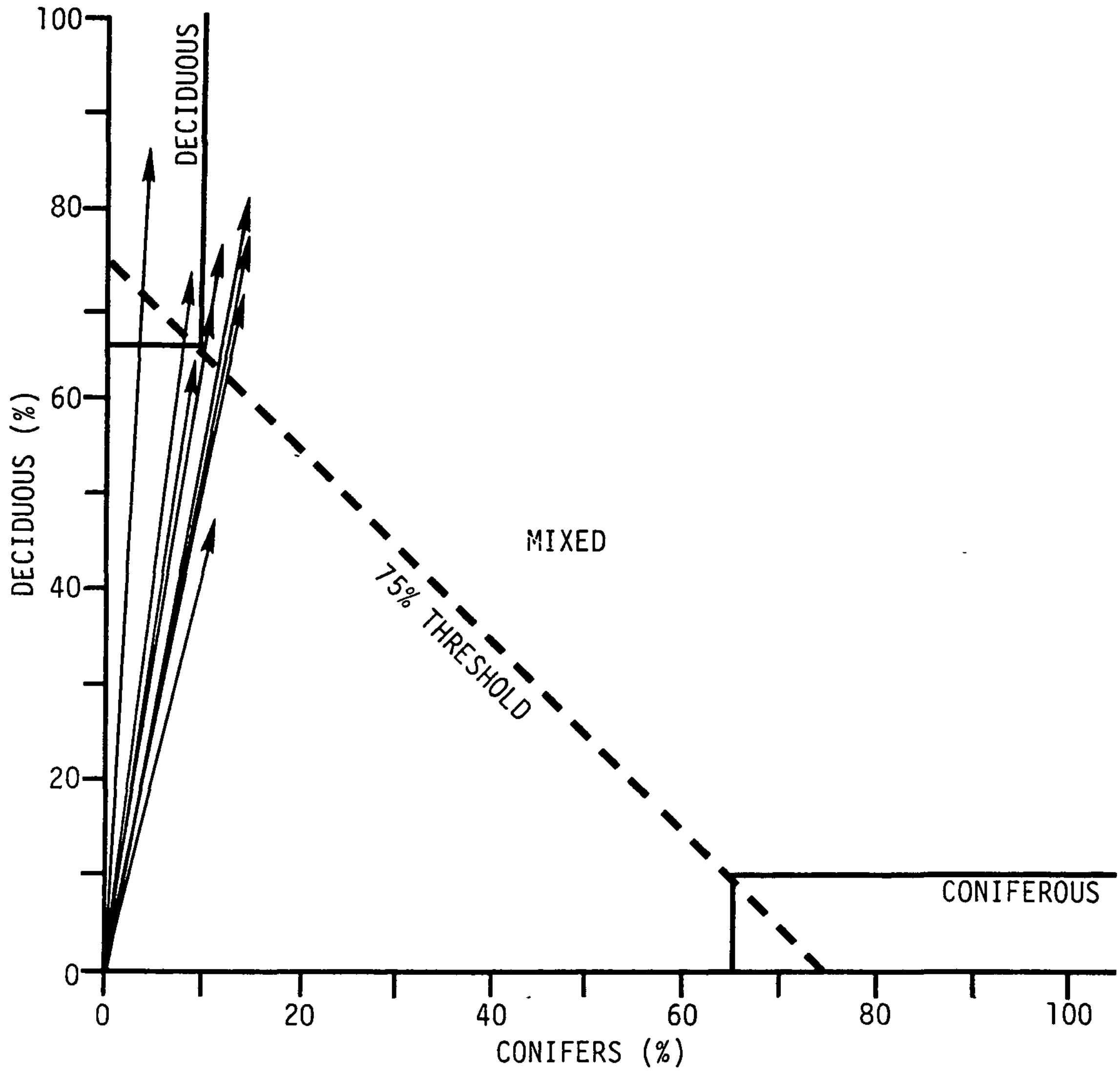

Figure 3. Decision regions and results of polygon classification. Vector endpoints portray how each management block was classified.

The preceding application serves to illustrate the ZONAL interface; however, indexing Landsat information to limited land areas is probably impractical. A unique aspect of analyzed remotely sensed data is the global perspective provided in recognizing occurrences and distributions of earth surface phenomena. Also, GIS's are typically employed to store and analyze detailed polygonal layers covering extensive land areas. More realistic applications would include inventorying and monitoring forest resources over industry-owned lands, surface mine evaluation and change detection analysis, and county-based land use and land cover inventories.

6. CONCLUSIONS

The ZONAL interface offers a reasonable means of utilizing Landsat information in a polygon-based GIS. The ZONAL mechanisms for information transfer are based on the use of existing GIS polygonal layers, thereby making the process entirely automated. The indexing scene permits location-specific inventories of terrain features within selected polygons. Through indexing, a voluminous set of image pixels is condensed to frequency distributions. By polygon classification, numbers buried in summarization tables can be extracted and analyzed. In this way, relationships are identified and polygons meaningfully characterized. GIS storage requirements are lessened by entering summations of relevant classification categories. Once an indexing scene is created, it may be used repeatedly in keeping a data base current, provided the polygonal layer remains unchanged. Additionally, ZONAL can be adapted to other processing systems and GIS's because host system modifications are not necessary.

7. LITERATURE CITED

Morehouse, S. and G. Dutton. 1980. Extraction of polygonal information from gridded data. Pages 320-327 in R. T. Aangeenbrug, Ed., AUTO-CARTO IV, Vol. 2. American Congress on Surveying and Mapping and the American Society of Photogrammetry, Falls Church, Va.

Myers, W. L. 1982. TOMIS: Task Oriented Information System Users Manual. School of Forest Resources, The Pennsylvania State Univ., University Park, Pa. (Mimeographed.)

Nichols, D. A. 1982. Conversion of raster coded images to polygonal data structures. Pages 508-515 in B. F. Richardson, Ed., Proceedings of the Pecora VII Symposium. American Society of Photogrammetry, Falls Church, Va.

Turner, B. J., G. M. Baumer, and W. L. Myers. 1982. The ORSER Remote Sensing Analysis System: A User's Manual. Res. Publ. 109/OR, Institute for Research on Land and Water Resources, The Pennsylvania State Univ., University Park, Pa.