Rule-Based Classification Models: Flexible Integration of Satellite Imagery and Thematic Spatial Data

Paul V. Bolstad
School of Forestry and Wildlife Resources, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061
T. M. Lillesand
Environmental Remote Sensing Center, University of Wisconsin, Madison, WI 53706

ABSTRACT: A framework for automated land-cover classification based on a concept of a classification model was developed and tested. The framework employs a user-specified rule base to describe a classification model, defined as the series of spatial data operations and decisions used in landcover classification. Both evidential and hierarchical inference are supported utilizing a set of spatial data operators. The concept was tested through the development and application of a set of computer programs which support classification models. A rule base, thematic spatial data, and satellite image data were then used to define a classification model for conditions in northeastern Wisconsin. The test model incorporated Landsat Thematic Mapper data, soil texture data, and topographic position data. Classification accuracies and efficiencies using the developed system were then compared to those for supervised maximum-likelihood classifications. The classification model approach resulted in statistically significant, 15 percent improvements in classification accuracy when averaged across different analysts, geographic areas, and years.

INTRODUCTION

SUPERVISED, PER-PIXEL, MAXIMUM-LIKELIHOOD SPECTRAL classifiers are the most commonly applied automated land-cover classification techniques due in part to a well developed theoretical base, ease in automation, and proven track record (Swain and Davis, 1978; Richards, 1986; Lillesand and Kiefer, 1987). Unfortunately, when used with satellite image data, these technologies often yield unacceptable accuracies for many applications. For example forest management agencies often require landcover classification at the Anderson et al., (1976) level II/III with at least 95 percent accuracy, while automated classification of satellite data generally result in accuracies well below this level (e.g., Nelson et al., 1984; Moore and Bauer, 1990).

Spatial data in a GIS have been shown to improve classification accuracy and aid in the extraction of information from remotely sensed imagery (Strahler et al., 1978; Likens and Maw, 1981; Marble and Peuquet, 1983). Methods include incorporation before, during, or after a maximum-likelihood classification (Hutchinson, 1982; Richards et al., 1982). However, current technologies do not allow easy automated integration of non-image spatial data (such as digital thematic maps and associated attributes) in image classification. In most instances spatial data are used in manual pre- or post-classification manipulations (Gaydos and Newland, 1978; Hutchinson, 1982). The integration of non-image data during classification often involves relatively inflexible hard-coded classifiers (Fleming and Hoffer, 1979; Hoffer et al., 1978), or “logical channels” (Strahler et al., 1978; Strahler et al., 1980) which violate distributional conditions.

Artificial Intelligence (AI) and expert systems techniques have been investigated to improve land-cover classification (Ferrante et al., 1984; Wharton, 1987; Argijalas and Harlow, 1990), and show promise in the integration of non-image spatial data (Mason et al., 1989), because of their flexibility, generality, and intuitive appeal. Both evidential and hierarchical approaches have been investigated. Evidential approaches rely on obtaining measures of the relative “mass” of evidence in support of alternative hypotheses (Duda et al., 1979; Goldberg et al., 1985; Lee et al., 1987), and select the hypothesis (land-cover class assignment) with the greatest evidence mass. Hierarchical approaches, such as the “decision tree” techniques described by Swain and Hauska, (1977), eliminate competing hypotheses from consideration during inference until only one hypothesis remains. Conceptually, land-cover classes are considered leaves of bi- or multinary trees, with decision criteria applied at each node to eliminate or select the decision path (Swain et al., 1977; Duda et al., 1978; Ferrante et al., 1984). Strictly hierarchical approaches are computationally more efficient, but do not recover from decision errors; conversely, methods which strictly accumulate evidence incur high computational loads. Thus, there is a tradeoff between flexibility and run-time efficiency. Systems may be inflexible in that landcover classification can be difficult with satellite systems (e.g., SPOT versus Thematic Mapper), feature types, or areas different from those on which the systems were developed (Ferrante et al., 1984; Wang and Newkirk, 1984; Goldberg et al., 1985). Although principles apply, changed conditions entail significant recoding. Those systems that do incorporate easily modifiable rule-based strategies result in classification times on the order of magnitude or more slower than the “standard” maximum-likelihood approach (Wharton, 1987; Mason et al., 1988; Civco, 1989). This efficiency/flexibility trade-off results from the applicative approach adopted by rule-based systems (Mehldau and Schwengerd, 1990) in that software for system development is designed for quick construction and prototyping, but provides unacceptable performance with large data volumes (Jackson, 1985). Thus, a combination of hierarchical and evidential approaches is often adopted (Shortliffe, 1976; McDermott, 1982; Jackson, 1988). This paper describes a land-cover classification approach controlling RS/GIS integration through an easily modifiable rule base, and which also provides rapid throughput.

DESIGN PHILOSOPHY

The adopted approach is based on a concept of a classification model. Classification models are defined as an automated sequence of operations applied to image and non-image spatial data which results in a land-cover classification. Classification models may be considered analogous to cartographic models defined for geographic information systems (Tomlin and Berry,
1979; Burroughs, 1986). These models may use various decision criteria and operations to assign distinct land-cover classes. Although a broad-sense interpretation of the definition includes a "standard" supervised maximum-likelihood classifier, classification models will generally have several "levels", and use several different data types and spatial data operators. For example, in a classification model, image and non-image data may be automatically integrated to classify a geographic area into four land-cover classes: urban, forest, cropland, and water (Figure 1). Digital population density data may be used to identify urban classes, digital thematic maps to assign water, and a spectral-based maximum-likelihood classifier used to assign the crop and forest classes. Thus, in a classification model a set of spatial data analysis primitives are used to classify land cover, e.g., maximum-likelihood calculations and direct assignment based on thematic data values.

Classification models should support both evidential and hierarchical inference. As noted above, both types of inference have been successfully applied to land-cover classification, and the combination is desirable both to facilitate increases in classification accuracy and enhance run-time efficiency.

The inference strategy should be specified at a high level of abstraction, i.e., one which allows non-programmers to understand and develop classification models, and it should also be flexible and easy to modify. A higher level of abstraction is a common paradigm of much AI programming, and it allows a wider application of this approach (Jackson, 1986). Flexibility also entails an easily modifiable system which may be applied to a wide range of land-cover and data types. Classification may use only spectral data, only thematic data, continuous non-thematic and non-image spatial data, or a combination of these data types.

Classification modeling requires a defined set of spatial data operators which may be considered image classification primitives. These primitives are analogous to map algebra primitives (Tomlin and Berry, 1979), except that they are expanded to include operators specifically designed for land-cover classification. These primitives will then be called to operate on co-registered data during the inference process defined by the classification model. For example, spectral class likelihood calculation, the boolean combination of thematic spatial data, shape determination, or local texture computation may all be useful as classification primitives. The classification model is defined by the sequence and combination of these primitives.

CLASSIFICATION MODELER: PROTOTYPE

A system of programs was produced which supports the development of classification models. This system, herein called CLASMOD (classification modeler), allows the integration of thematic data, satellite imagery, and a rule-based, forward-chaining inference strategy in land-cover classification. The adopted approach uses a set of rules to define a classification model. Both evidential and hierarchical inferences are supported. Rules are used to describe feature types, data themes, the relationships among themes and feature types, and to define the inference path. Although CLASMOD is sufficient to test the concept of a classification modeler, it is an early prototype and as such is limited in many respects.

CLASMOD consists of three parts: a rule parser, "foundation code," and a compiler (Figure 2). These three are used with a rule base to define a classification model. First, the rule parser is applied to a rule base, generating computer code. For example the rule "if soil text is sand and topo position is upland then restrict by topo position" might be converted to the C code fragment "if (theme_var[i] = = theme_tab[a][b] || theme_var[i] = = theme_tab[c][d]) {restrict(theme_var[j])}.

The generated code is then compiled and linked with the foundation code to produce an executable classifier based on the classification model. The executable code then operates on the base and spatial data to classify land cover for a geographic area (Figure 2).

The rule parser recognizes a number of keywords in rules which are contained in the rule base (Table 1). The parser converts the rule base into C code and data tables which are used in the classification model. Rules are used to identify the set of

![Fig. 1. An example classification model used in feature class assignment. Note that the timing and application of restriction and evidence accumulation operations may be based on the thematic and image data values.](image-url)
Table 1. Rule Types, Examples, and Keywords Recognized by the Rule-Parse. Rules are Used to Build a Classification Model by defining Feature Types, Thematic Data Layers, Theme Classes, Evidence Values, and the Timing and Type of Decisions Used in Classification. Keywords Recognized by the Parser are italicized.

<table>
<thead>
<tr>
<th>Rule Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Define land-cover classes and identify numeric coding</td>
<td>red_pine isa feature_type with numeric id 1</td>
</tr>
<tr>
<td>Define thematic data types</td>
<td>soil_text isa evidence_theme</td>
</tr>
<tr>
<td>Define theme classes and numeric coding</td>
<td>sand is soil_text with numeric id 7</td>
</tr>
<tr>
<td>Fix evidence values</td>
<td>if soil_text is sand and topo_position is upland then restrict_by topo position</td>
</tr>
<tr>
<td>Describe restrictions</td>
<td>level 1 if topo_position is upland then restrict_by topo position</td>
</tr>
<tr>
<td>Guide flow of control</td>
<td>level 2 accumulate_evidence_by soil_text</td>
</tr>
</tbody>
</table>

Classification must be restricted to those regions for which the classification model is valid. However, the system is sufficiently flexible to allow the generation of different classification models for different regions or conditions.

A small set of classification primitives is currently supported: thematic restriction, thematic evidence accumulation, spectral evidence accumulation, and classification quality. While these constitute only a limited subset of potential operators, they allow the flexible incorporation of spectral and non-spectral data into the classification techniques.

Restriction operations are used when the occurrence of a thematic class reduces the set of plausible feature types. For example, we may be certain that only lowland vegetation types (e.g., black spruce and bog sedge) are found in areas which have been mapped as having organic surface soils. Thus, if we have a soil map, every time we encounter an organic surface soil we can eliminate all but the lowland types from consideration. Thematic evidence accumulation occurs when thematic data support or contradict feature type occurrence with less than complete certainty. For example, areas with sandy soils may be more likely to support jack pine and red pine, while loamy soils are more likely to support sugar maple. Note that the same thematic data layer can be used for both evidence accumulation and restriction. Finally, spectral evidence accumulation adds supporting evidence based on imagery which is proportional to spectral class likelihoods. The remotely sensed data are viewed as one of many spatial data sources from which information about land-cover may be extracted.

The spectral evidence accumulation primitive is based on likelihood calculations with image data. When this primitive is invoked, the likelihoods for all plausible spectral classes are calculated and ranked. Likelihoods are scaled using a Chi-square distribution. Spectral data coincident with training set mean values yield Chi-squares of zero and correspond to the highest possible likelihood for the training set. A zero Chi-square is set to correspond to 100 "evidence points," and values above a 95 percent Chi-square threshold are assigned 0 points. Spectral evidence is linearly scaled over this range.

Finally, operations can be included which gauge the quality of the spectral evidence. Evidence provided by the imagery may be deficient in several respects. For example, the highest calculated spectral likelihood may be below an acceptable level. Alternatively, two feature types may both contribute high spe
These four categories were chosen because they were related into four nominal categories: wetland, upland, water, and road. Permanent water, road, and vegetation distribution and were well represented on available USGS 7.5-minute quadrangles. Topographic positions were categorized into four nominal categories: wetland, upland, water, and road. These four categories were chosen because they were related to vegetation distribution and were well represented on available USGS 7.5-minute quadrangles. Permanent water, road, and lowland areas were digitized directly from the quadrangles. All remaining areas were assigned to the upland topographic position class.

Landsat TM data were collected on 9 June 1988 and 23 September 1989. Radiometric quality of the 1988 scene appears good in all non-thermal bands. Degradation due to clouds and haze ranges from slight to significant over the entire TM scene, although haze and cloud effects are minimal over the study areas. The 1989 scene was also of high radiometric quality, and is nearly cloud free.

Field notes, Wisconsin Department of Natural Resources (DNR) vegetation maps, United States Forest Service (USFS) vegetation maps, and National High Altitude Aerial Photography (NHAP) color aerial photographs were assembled in support of the described research. These data were used to aid in training set development and for testing the CLASMOD. Thirteen feature types were defined in the set of rules (Table 2). These feature classes correspond approximately to the categorical detail of levels II and III of the Anderson classification system (Anderson et al., 1976). Categories were defined according to a number of considerations, including previous work on type categories for northern boreal and eastern deciduous forests using current technologies (Nelson et al., 1984; Buchheim et al., 1985; Horler and Ahern, 1986; Vogelman and Rock, 1986; Hopkins et al., 1988).

The developed rule base contained over 200 rules. Rules defined the land-cover classes, the thematic spatial data layers, thematic classes for each theme, and the sequence of spatial data operations used during classification. Examples of defined rules are included in Table 1.

The performance of the classification model was compared to traditional supervised maximum-likelihood classification methods for both study areas. This statistical comparison involved a stratified random sampling of predicted and true feature types for a number of points in both the human and rule-based classifications. Classification error rates were then calculated and compared for the two classification methods. Comparisons were made for two separate analysts on two image dates. Both the eastern and western study sites were classified.

First, the rule-based classifier was developed using the June 1988 sub-scene for the western study area. System refinement over each iteration included classification, evaluation, evidence value modification, and rule-base modification. Accuracy of the final classification was determined by comparison of the true and classified feature type for a number of points in the study area.
area. The accuracy of a traditional maximum-likelihood classification of the same area was also determined, and a statistical comparison performed of the traditional versus rule-based classification. The traditional and rule-based classifiers were then applied to the eastern study area using the training data from the western study area. This work was performed by “Analyst 1.” A second analyst (Analyst 2) classified the western study site using the traditional and rule-based methods. No rules or evidence weightings were changed, the only difference being spectral training data developed by the second analyst. The classification involved training on the western study site, and applying the spectral training data to the western and eastern sites.

A series of classifications was performed to gauge the sensitivity of classification results to the relative weights placed on the different evidence and restriction layers. Thematic evidence values were based on long-term observations of field experts, and spectral evidence could be considered well founded, assuming representative training. However, appropriate relative weightings of spectral and thematic evidence were not clearly defined. Spectral and thematic data provide different information. A priori reason to provide greater weight to one or the other. The current system weighted them equally under the assumption that the classification model is robust with respect to relative weightings. The appropriateness of this equal weighting was tested employing a series of classifications of the eastern study site using the 1988 imagery. Before accumulation, evidence values derived from the soil theme and satellite imagery were pre-multiplied by fractional coefficients W and 1–W, respectively, where 0 ≤ W ≤ 1. The coefficients summed to 1; thus, if the soil evidence was multiplied by 0.2, spectral evidence for that classification was multiplied by 0.8. Values of W averaged 0.0, 0.2, 0.4, 0.6, 0.8, and 1.0. These sensitivity classifications were performed both with and without the topographic restriction layer.

Approximately 20 to 40 features were ground-identified or photointerpreted for most feature types. Actual sample size per category was determined after preliminary visual classification revealed approximate class distribution (Williams, 1978). Between 360 and 410 sample points were identified in each of the eastern and western study areas. These land-cover points were used as “ground truth,” that is, the true feature type for the ground sample point. Because of land-cover changes from 1988 to 1989, the same sets of points were not used for both years, although there was a high degree of overlap among sets. The ground truth data were used to calculate Cohen’s Kappa coefficients of agreement and conditional Kappa (Cohen, 1960). Pairwise significance tests between the new and traditional approaches were performed using large sample Kappa and conditional Kappa formulas developed by Fleiss et al. (1969).

RESULTS

Quantitative system assessment was based on two criteria: accuracy and efficiency. Classification accuracy comparisons are summarized below. Results from these comparisons are organized by study site, year of image acquisition, and operator. Next, the results from the sensitivity tests are provided. Finally, efficiency observations are reported.

Classification accuracy comparisons are summarized in Table 3. Improvements in classification accuracy were observed for five of the six operator, study area, and imagery combinations. Differences in Cohen’s Kappa between the new and traditional methods, based on one-tailed z-tests, were statistically significant in five cases at the 10 percent level. Overall, pooled classification accuracies ranged from a low of 54 percent to a high of 90 percent. Using the CLASMOD and the defined classification model, classification accuracies averaged 83 percent, while accuracies for the traditional method averaged 69 percent. In general, the CLASMOD method provided more consistent, accurate classifications for the different combinations of study site and analysts.

Classification improvements associated with the CLASMOD were observed for both analysts in both study sites with the 1988 and 1989 imagery. The classification model was on average approximately 15 percent more accurate than the “traditional” approach for the 1988 data. The CLASMOD approach averaged 89 percent, while the supervised maximum-likelihood approach averaged 74 percent. For the 1989 data accuracies averaged 57 percent for the traditional method and 70 percent for the rule-based method. The same general patterns of misclassification were observed for both classification methods.

Accuracies were lower for the 1989 imagery than corresponding classifications for the 1988 imagery. At least two factors may have contributed to this reduction. First, spectral training area locations were defined on 1988 imagery. Training set coordinates were then transformed to the 1989 imagery and used to extract training samples. These samples were not tailored to the spectral characteristics of the 1989 imagery. Second, class accuracies could be lower because of the spectral characteristics of the 1989 imagery. This imagery was collected during late September of a low rainfall year. Although data from earlier in the growing season were preferred, attempts at collecting imagery for the months of June, July, and August were unsuccessful due to cloud cover. Leaf senescence initiated early for some species and this may have resulted in both greater spectral variability within feature types and greater similarity among some feature types. This spectral variability was not present in the 1988 data.

Training times are equal for the new approach when compared to a traditional maximum-likelihood classification (Table 4). Training times for analyst 2 using the non-traditional approach are probably over-estimates, because they include some time dedicated to familiarization with the new approach. Prior to this work, analyst 2 was not familiar with the particular programs described herein. Classification run times for the two methods are comparable, running approximately 18 percent longer for the new approach. The slight increase associated with the rule-based approach can be attributed to additional overhead such as increased file I/O and to additional operations,
such as restriction and evidence accumulation. For this specific classification model, the topographic restriction and soil evidence accumulation operators incur additional overhead relative to the traditional maximum-likelihood approach. However, they also reduced the time spent on computationally expensive likelihood calculations, in part offsetting the additional computational burden.

While these results provide a rough estimate of the relative times required for the new and traditional approaches, they do not reflect peak performance possible in a desktop environment for either approach. Although reasonable attempts were made to ensure computational efficiency for both the traditional and new approaches, neither was developed specifically to optimize speed of operation. Significant time savings could be realized in both instances through a number of techniques, including optimized I/O buffering (Knuth, 1973), table look-up classification techniques (Bolstad and Lillesand, 1990), faster clock speeds, or array processors (Westman, 1989).

Finally, results from the sensitivity tests indicate that the overall accuracies are robust relative to the relative weightings of the thematic and satellite image data during classification, at least with the adopted classification model (Figure 5). Classifications, both with and without the use of topography-based restrictions, remained high over a relative evidence weights of 0.2 to 0.8. In all cases, classification with the topographic restriction was improved over classifications using thematic soil and image data.

CONCLUSIONS

The adopted image classification approach has several advantages in comparison to standard approaches:

- The domain of discourse and control information is provided in an easily modified and understandable set of rules.
- Specific feature type, thematic variable, and image classification information can be persistent across different classifications of the same area, and can be modified for use in other regions or with different feature types.
- Computationally expensive operations can be avoided using restriction operators, without resorting to manual image recoding, masking, and image recombination.
- The modular rule-based approach allows the integration of evidential and deterministic discrimination techniques, and the incremental addition of new spatial data operators, thematic data, or knowledge, which aid land-cover classification.

The described rule-based approach illustrates a straightforward, flexible, unified means of improving classification accuracy while incorporating remote sensing, GIS, and AI techniques. Accuracy was improved in a statistically significant manner for different study areas, analysts, and image acquisition dates. This work demonstrates the potential for using AI and GIS techniques to integrate these accuracy improvements directly into the classification phase. Further, both strategic and descriptive information are explicitly represented and can be easily changed, supporting the persistence of knowledge and facilitating incremental improvement in the classifier, flexibility in the incorporation of new image processing techniques, and application of the method to new areas or feature types.

| Table 4. Total Time Requirements for Both Training and Classification Using Traditional and Non-Traditional Approaches. All Times Are in Minutes. |
| --- | --- | --- |
| Analyst No. | Traditional Method Time (min.) | Rule-Based Method Time (min.) |
| 1 | 363 | 432 |
| 2 | 402 | 492 |

![Fig. 5. Results of the sensitivity test classifications in which the relative weights of the image and thematic spatial data varied from 0 to 1. Classifications were based on a combination of TM and soils data (without topographic data) and a combination of TM, soils, and topographic position data (with topographic data). Overall classification accuracies are for 13 Anderson level II/III classes, based on 382 sample points.]

LIMITATIONS AND PROSPECTS

Conclusions for the current work are limited to the conditions of the study. Although a range of conditions was tested, comparisons of new to traditional approaches for a broader range of themes, analysts, feature types, and image types are necessary to establish broad-scale generality of the new approach. However, the success of the RS/GIS/AI approach under a broader range of conditions is probable, particularly when a broader range of spatial data primitives is supported.

ACKNOWLEDGMENTS

This work was supported by the U. W. Institute for Environmental Studies, the U. W. College of Agriculture and Life Science, School of Natural Resources, and Sigma Xi.

REFERENCES


Duda, R. O., P. E. Hart, J. Gashig, K. Konolige, R. Reboh, and J.
RULE-BASED CLASSIFICATION MODELS

RESOURCE TECHNOLOGY ’90

The latest addition to the remote sensing section of ASPRS publications. These proceedings are from the Second International Symposium on Advanced Technology in Natural Resource Management, held in November, 1990. Topics include spatial data systems, remote sensing, decision support, global climate change, ecological modeling, and other integrated technologies for natural resource management.

1991. 830 pp. $70 (softcover); ASPRS Members and Students $45. Stock # 4525. Please see the ASPRS Store in this issue for ordering information.