POLYGON BASED ANALYSIS OF REMOTELY SENSED IMAGES IN AN INTEGRATED GEOGRAPHIC INFORMATION SYSTEM

Eugene Derényi, Hon. Research Professor
Mustafa Türker, Research Associate
Department of Geodesy and Geomatics Engineering
University of New Brunswick
Fredericton, N.B. Canada E3B 5A3

Commission IV, Working Group 6

KEY WORDS: Change_Detection, Monitoring, Updating, Database, Knowledge_Base, Image Classification, Image Statistics, Raster Vector Integration.

ABSTRACT:

Two schemes were developed and successfully implemented for monitoring changes and prevailing land cover conditions, within geographic boundaries stored in a GIS, by digital image analysis. In the first scheme, a standard pixel-by-pixel image classification is performed, one theme at a time, covering the entire area of interest. A polygon-by-polygon assessment of the results follows to separate those polygons which deviate from the apriori expectations by more than a preset threshold level. These polygons are then subjected to further examination. In the second scheme, image classification is bypassed entirely and only image statistics are generated polygon-by-polygon. A comparison of these statistics with those representing normal, expected conditions identify those polygons where significant changes occurred. Both schemes are demonstrated by an example.
The above approach to monitoring changes overcomes the problems of multi-classification and misclassification due to overlaps of spectral signatures, which often plagues traditional image classification. In addition, it provides for an immediate update of the data base in the GIS.

1. INTRODUCTION

The Earth is a planet with finite resources. As the population continues to grow, a wise and prudent management of these resources is becoming increasingly important. Such management is best achieved if accurate and up-to-date inventories are at the disposal of decision makers. More and more of the information needed for this purpose is generated by remote sensing. An image is worth a thousand words.

Remote sensing is, however, not an end in itself. Classified images displayed on a monitor or as multi-colour plots are end product of limited value. The realistic approach is to store the information extracted from images in a geographic information system (GIS) where it can be merged with other information pertinent to resource management. This is, of course, becoming more and more the norm. The closer GIS and remote sensing become integrated the greater the benefits derived by resource managers become.

The flow of information in the opposite direction is equally important and beneficial. GISs now hold vast volume of well-organized information which could serve as a knowledge base to facilitate the analysis of new remotely sensed data. This advantage has, of course, been recognized for some time and topographic, geophysical, geological, soil, etc. data are often utilized as additional raster layers in image classification. Administrative, land use, etc. boundaries are also being used for segmentation and stratification of images. In addition to the geographic features and topological information, GISs also contain attribute data which are stored in a relational data base management system (DBMS) interfaced with the GIS, and can be queried. Such descriptive information about features is, however, rarely utilized in image analysis since their alpha-numeric storage format is not compatible with the raster organization of digital images, and most digital image analysis systems are not interfaced with relational database management systems. Therefore, a universal GIS (UGIS) is needed where both vector based geographic data and raster based continuous tone image data can be processed, analyzed and manipulated and which is interfaced with a DBMS. The Computer Aided Resource Information System (CARIS) used in this project, has most features of such a facility [Derenyi and Pollck, 1990; Derenyi, 1991]. In a UGIS the per pixel image classification can be replaced by polygon specific image analysis, and apriori knowledge, stored in the DBMS, can be incorporated in the decision making. The results can directly be stored in the database.

2. THE CONCEPT

The polygon specific image analysis scheme presented here has been developed for detecting and monitoring changes within existing geographic boundaries, which are stored as polygons with associated attributes in a GIS. Final class assignment is made polygon-by-polygon, based on information extracted from the raster data layers and on the polygon attributes stored in the DBMS. The raster layers may include image and non-image data.

Two methods have been employed for analyzing the image data, and to arrive at the final class assignment of individual polygons:

1. modified per-pixel image classification, and
2. decision by polygon specific statistics generation.

In the first method, supervised classification is performed covering the entire project area. This classification is done one theme at a time, and each class is stored as a separate layer. Training samples are selected with care as usual, but multiclassification of pixels caused by spectral overlaps is of no major concern since the output of each class is clearly separated in the files. An advantage of operating in a GIS is that polygon attributes are good sources as reference data and the training samples can easily be located by the polygon identification numbers. The processing of the data one theme at a time allows the use of the parallelepiped classifier, which is the simplest algorithm.

Upon completion of the pre-pixel supervised classification, the multi-layered output is analyzed polygon-by-polygon, one layer at a time. The polygons are selected for assessment either interactively by pointing on them with the cursor on the screen, or in batch processing mode by selecting them through the polygon identification codes. A pixel count is then initiated in each polygon to determine the number of the pixels in a particular polygon that are labeled as belonging to the various classes. At this point a decision has to be made on the final class assignment of the polygon being assessed, assisted by the discipline specific expertise of the analyst, and on ancillary information available in the data base.

Various alternatives could now arise. For example, if the percentage of pixels belonging to a particular class is above a preset threshold, then the entire polygon area is declared as belonging to this class. This means that this area remained unchanged or the change occurred according to expectation. If the polygon contains pixels of more than one class in significant percentages, then it is declared as a dual or multiclass area. The proportion of the class distribution can be adjusted to the known area of the polygon as a constraint. Polygons with multiclass pixels, none of which add up to a minimum threshold level, or with a significant number of unclassified pixels, are flagged for further examination. Further data base query may resolve the issue or on-site inspection may be necessary.

In the second analysis method of the raster data, the per-pixel image classification is by-passed entirely, and only image statistics are generated polygon-by-polygon. Mean, standard deviation, median and mode are likely candidates. A comparison of these statistics, one theme at a time, with those obtained for the training samples, forms the basis of the class assignment of each polygon. Various alternatives can again emerge and the final decision requires project specific expertise. This method is a one step operation.

3. IMPLEMENTATION

The polygon specific classification scheme was tested on monitoring the growth of forest plantations using Landsat Thematic Mapper (TM) imagery. The 367 km^2 study area contained plantations of two coniferous species labeled as PL1 and PL2. There were a total of 225 plantation plots ranging in size from 400 ha to 16,000 ha. Planting was done between 1977 and 1993. The objective was to check whether the development of the individual plantation stands followed the expectation. Stands that diverged from the normal development curve beyond a certain tolerance were flagged for further examination. The biophysical basis of this monitoring is the assumption that the percentage of crown closure of the trees is a valid indicator of the development of the plantation [Honer, 1972; Danson, 1984], and that the spectral reflectance received by the TM sensor is a function of the crown closure [Gemmell, 1995; Leckie et al., 1992; Spanner et al., 1990; Stenback and Congalton, 1990]. The study area is located in north-central New Brunswick.

The Landsat TM image was recorded in November 1994. It was cloud free and of good quality. The boundaries of the plantation stands were shown on a 1:50,000 scale plantation map, and were manually digitized. A database was then set up in the GIS, where each plantation polygon was assigned an identification number and the year of planting was recorded as an attribute.

Geometric registration was handled in a novel way. The objective of this project was a polygon specific monitoring of plantation development. Thus it was sufficient to have an accurate geometric registration of the raster image and the vector polygon layers in relative sense, and registration to a map grid in an absolute sense, with the help of ground control points, was not necessary. Therefore, the polygon network was transformed to fit the image. This was a much faster and simpler operation then the geometric and radiometric transformation of the whole multiband image file, and was easily accomplished in the GIS [Derenyi, 1994].

Training samples were selected with the help of the polygon overlay. The TM image was displayed on the screen with the plantation boundaries superimposed, and groups of representative pixels were delineated for each theme within designated polygons. The polygons selected for training in each plantation age group were from among those where the normal expected tree growth was confirmed. The PL1 plantation species consisted of seven age groups, planted in the years: 1977, 1978, 1979, 1981, 1987, 1990, and 1993. The PL2 species were planted in the following six years: 1977, 1978, 1979, 1980, 1981, and 1992. The image statistics generated for each age class were: the minimum and maximum digital numbers (DN), the mean value, standard deviation, median and the mode. TM spectral bands 3, 4 and 5 were only used in the analysis.

The next step in the per-pixel image classification method was to create the thirteen theme layers by parallelepiped classification, performed one theme at a time. The decision region was set at three standard deviation. Thereafter, the percentages of correctly classified pixels were computed in each polygon. Correctly classified meant that the pixel in question had been placed in the growth class which corresponded to the plantation year of that polygon. In other words, the classification of this pixel agreed with the apriori expectation. Such was the case for example, if a pixel, located in the polygon with attribute PL1-77 (species PL1 planted in 1977), also resided in the PL1-77 information class (theme) layer.

The mechanics of the assessment was handled as follows: All polygons belonging to the same plantation year were retrieved one-by-one through a database query. Each vector polygon was converted into a raster object and overlaid on the corresponding information class layer obtained through the image classification. The degree of agreement of the two layers was measured by a pixel a count

The results of the assessment were grouped into four categories according to the following percentage ranges of agreement with the apriori expectation:

0-30%, labeled Very Low (VL), signaled major problems. The majority of the threes were probably lost, severely retarded in growth or perhaps there was an error in the database entry.
31-50%, labeled Low (L), indicated unsatisfactory development of the plantation. The stand may be under

stress.

51-70%, labeled Marginal (M), which warrants further investigation.

71-100%, labeled High (H), and was considered a satisfactory development.

The above categories were chosen for illustrative purposes only. The results are summarized in Table 1.

Finally, the percentage of correctly classified pixels in each polygon together with the agreement code were entered into the database.

In the polygon specific statistics generation method the image statistics were computed within every plantation polygon to be assessed, one-by-one. These statistics were then compared with those obtained from the training samples that matched the age class of the polygon being assessed. For example, the image statistics obtained within a polygon with attribute PL1-77 was compared to the training statistics obtained for the PL1-77 age class. Ranges were set in terms of the standard deviation (σ) of the training samples for the agreements between the polygon and the training values for the mean, median, and mode statistical expressions, to judge the degree of compliance of the individual polygons with the normal development of the plantation. Table 2 shows the number of polygons whose image statistics agreed with the control value at the 3σ level.

An additional testing mechanism devised was the ratio by which the 3σ ranges of the training and the tested polygons overlapped. The count of the polygons that passed this test above the 0.6 ratio are also shown in Table 2, under the heading: Ratio.

A comparison of Tables 1 and 2 indicates a good agreement of the results obtained by the image statistics based assessment at the 3σ level, with the 70-100% class agreement in the per-pixel classification. The 3σ agreement level was, however, chosen for illustrative purposes only.

4. CONCLUSIONS

Polygon specific image analysis in a GIS is a powerful tool for monitoring changes and assessing prevailing land cover conditions. It has numerous advantages:

- The monitoring is performed within known geographic boundaries which are stored in the GIS and can be queried through the database management system.
- This monitoring scheme provides for an immediate update of the database.
- Problems caused by overlaps of spectral response patterns of the various themes, which often plague traditional image classification, are largely overcome.
- In the per-pixel classification method, the parallelepiped classification performed on a few carefully selected bands can provide satisfactory results. This classifier is computationally the least demanding.
- Satisfactory results can also be expected if image classification is bypassed entirely and change detection is based solely on comparing image statistics.

Table 1

NUMBER OF POLYGONS FALLING INTO EACH OF THE FOUR CATEGORIES OF AGREEMENT WITH APRIORI EXPECTATION IN THE MODIFIED PER-PIXEL CLASSIFICATION MONITORING METHOD.

Polygon Code	0-30%	31-50%	51-70%	71-100%	Total Polygon
PL1-77	1	2	6	9	18
PL1-78	1	1	6	22	30
PL1-79	--	1	--	9	10
PL1-81	3	2	11	22	38
PL1-87	5	--	2	9	16
PL1-90	1	--	3	15	19
PL1-93	1	1	4	--	6
PL2-77	--	--	1	7	8
PL2-78	--	--	--	15	15
PL2-79	--	1	3	20	24
PL2-80	5	2	--	3	10
PL2-81	--	--	1	23	24
PL2-92	2	--	1	4	7

Table 2

NUMBER OF POLYGONS IN AGREEMENT WITH APRIORI EXPECTATION WITH 3σ RANGE IN VARIOUS IMAGE STATISTICS COMPARISONS.

Polygon Code	Mean	Ratio	Median	Mode	Total Polygon
PL1-77	9	9	13	16	18
PL1-78	27	23	28	28	30
PL1-79	9	7	9	9	10
PL1-81	35	20	35	35	38
PL1-87	12	10	11	11	16
PL1-90	18	18	18	17	19
PL1-93	5	1	5	4	6
PL2-77	7	7	8	8	8
PL2-78	15	13	15	15	15
PL2-79	24	17	24	22	24
PL2-80	3	3	3	3	10
PL2-81	24	22	24	23	24
PL2-92	5	4	5	5	7

The objective of this paper was to introduce and explain the concept of polygon specific image analysis as applied to monitoring changes. The example of monitoring the development of forest plantations was presented only to illustrate this concept, and the quality of the results should not be judged from a forest management point of view. The real success of this method depends on the decision rules set at the assessment stage. These rules must be formulated by discipline oriented experts, on a case-by-case basis through careful consideration of the specific project. Further testing and refinement of the ideas presented here are in progress.

Polygon specific image analysis in a UGIS is an advantageous approach in a wide range of applications such as monitoring land use changes, updating forest inventories, agricultural crop classification on a yearly basis, monitoring compliances with environmental regulations, etc.

REFERENCES

Danson, F.M., 1989. Factors effecting the remotely sensed response of coniferous forest canopy. Ph.D. theses, Department of Geography, University off Sheffield, U.K.

Derenyi, E. and R. Pollock, 1990. Extending a GIS to support image-based map revision. *Photogrammetric Engineering and Remote Sensing,* 56(11): 1493-1496.

Derenyi, E., 1991. Design and development of a heterogeneous GIS. *CISM Journal ACSGC*, 45(4): 561-657.

Derenyi, E., 1994. Simple techniques for thematic mapping. *ISPRS Journal of Photogrammetry and Remote sensing*, 49(6): 27-31.

Gemmell, F.M., 1995. Effects of forest cover terrain and scale on timber volume estimation with Thematic Mapper data in Rocky Mountain Site. *Remote Sensing of Environment,* 51: 291-305.

Honer, T.G., 1972. AS high-density concept and measure. *Canadian Journal of Forest Research,* 2:441-447.

Leckie, D.J., M.D. Gillis and S.P. Joyce, 1992. A forest monitoring system based on satellite imagery. Canadian Symposium on Remote Sensing, Toronto, Ont.: 85-90.

Spanner, M.A., L.L. Pierce, D.L. Peterson and S.W. Running, 1990. Remote sensing of temperate coniferous forest leaf area index, the influence of canopy closure, understory vegetation and background reflection. International Journal of Remote Sensing, 1:95-111.

Stenback, J.M. and R.G. Congalton, 1990. Using Thematic Mapper imagery to examine understory. *Photogrammetric Engineering and Remote Sensing,* 56(9): 1285-1270.