CHANGE DETECTION USING THE INTEGRATION OF REMOTE SENSING AND GIS: A POLYGON BASED APPROACH

Mustafa TURKER, Assistant Professor
Orta Dogu Teknik Universitesi
Fen Bilimleri Enstitusu
Jeodezi ve Cografi Bilgi Teknolojileri EABD
06531 Ankara, TURKEY
mturker@metu.edu.tr

KEYWORDS: Change Detection, RS/GIS Integration, Image Analysis, Database Updating

ABSTRACT

A per polygon image analysis scheme for detecting the changes in land cover conditions within existing land use boundaries stored as vector polygons in a Geographic Information System (GIS) is described. The scheme consists of three main steps: integration of remotely sensed and geographic data, the statistical testing of the pixels, and the assessment of the results. Change detection in the image is facilitated by a geographic database which contains the land use polygons and the attributes associated with them. The imagery is analysed in a polygon specific manner. The scheme provides a means to analyze the remotely sensed data by spatially related information stored in a database and the manipulative processing operations which a GIS supports. The concept is demonstrated by an example on the monitoring of the development of tree plantations.

1   INTRODUCTION

Keeping information up-to-date is important for management and planning of land resources. Planners and resource managers need reliable techniques for accurately detecting, monitoring and analysing land use/land cover changes quickly and efficiently. The digital nature of satellite data makes the change detection process easily amenable for computer aided analysis. Jensen (1996) provides an extensive summary of currently used change detection techniques, such as postclassification comparison, classification of multitemporal data sets, temporal image differencing, temporal image rationing and change vector analysis. All are based on pixel-by-pixel analysis of multitemporal remotely sensed images. In other words, all operate on images only. On the other hand, most maps and resource inventories are now held in GISs, and can serve as a knowledge base to facilitate the analysis of new remotely sensed images. In geographic databases, the land use boundaries, such as forest stands, orchards, agricultural fields, are represented as polygons. In addition to the graphic and topological information, GISs also contain attribute data, which are stored in a relational database and can be queried through a DBMS interfaced with the GIS.

In this study, a per polygon image analysis scheme is presented which was developed for detecting and monitoring the changes in land cover conditions within land use boundaries stored in a GIS. Change detection in the image is facilitated by a geographic database containing the land use polygons and the attribute information associated with them. The analysis of remotely sensed data is performed in a polygon specific manner in which a priori knowledge stored in the form of attributes in a database is incorporated in the decision making. It provides a means to analyse the remotely sensed data by spatially related information stored in a database and the manipulative processing operations which a GIS supports. In addition, the attribute database is updated automatically by directly entering the new polygon attribute(s) created from the image.

2   METHODOLOGY

The scheme consists of three main steps: integration of remotely sensed data and land use vector polygons, the statistical test of the pixels, and the assessment of the results (Turker, 1997). The remotely sensed imagery and the vector polygon data are integrated by geometric registration. Training samples are then selected in preparation for the statistical test of the pixels. These samples are representative of the land cover themes of interest in the study area. An advantage of operating in GIS/RS integration is that polygon attributes stored in the database could facilitate this operation and the polygons selected for training samples can easily be located by their identification codes.

A rejection procedure is used to allocate a pixel to a land cover theme considered. It is based on the value of the exponent in the likelihood. The distribution of

$$(X\text{-}m_i) \quad {}_i^{-1} (X\text{-}m_i) \tag{1}$$

where; X, the vector of a pixel, m_i, the sample mean vector for theme i, and Σ, the sample variance co-variance matrix for theme i, is χ^2_n. The estimates of the m_i and Σ_i can be used to approximate the actual values and a vector X is rejected if, for the theme to which it would be allocated, this normalised sum of squares is larger than the χ^2_n table value, where n is the degree of freedom corresponding to the dimensionality of multispectral space.

The statistical test of the pixels is performed within the entire study area, one theme at a time. The mean value and the variance co-variance matrix of a theme considered are defined from the training samples. The output for each theme is stored as a separate layer. Each theme represents a certain condition. Theme by theme statistical test of the pixels circumvents the problem of multiclassified pixels, caused by spectral similarities of certain themes, which often occur in a conventional classification operation.

Upon completing the statistical test, the multilayered output is analysed polygon-by-polygon, one theme at a time, to check the degree of agreement of the content of each polygon with apriori expectation. The polygons are assessed only in the theme layer to which it is expected to belong. The procedure of the assessment is as follows:

For each theme, the polygons belonging to that theme are selected in batch mode through a database query. Then, in each polygon, the percentage of pixels is computed from the output, which belong to the theme being assessed. If the percentage of the pixels belonging to the theme being assessed is above a preset threshold then the entire polygon area is accepted as belonging to this theme. It means that this area represents the condition as expected. Polygons with a percentage of pixels below the preset threshold are declared as having failed the test and are flagged for further examination. Such polygons may be assessed in another theme layer to check if the condition matches a different expectation. As the last resort, site visit may be necessary.

3 IMPLEMENTATION

3.1 Test Area and Data

The scheme was tested on the monitoring of tree plantations using Landsat Themetic Mapper (TM). There were no clouds over the study area. The pixel size was 30m. x 30m. The 367 km^2 study area contained plantations of two coniferous species identified as PL1 and PL2. There were a total of 225 plantation plots ranging in size from 400 ha to 16,000 ha. The age of the seedlings ranged from one to seventeen. The objective was to check whether the development of plantation plots follows the expectation and detect those plots that diverged from the normal development curve beyond a certain tolerance. 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, 1989), and that the spectral reflectance received by the TM sensor is a function of the crown closure (Gemmel, 1995; Leckie et al., 1992; Spanner et al., 1990; Stenback and Congalton 1990).

3.2 Preprocessing

The boundaries of the plantation stands were manually digitized on a 1:50,000 scale map to set up a database. Each plantation polygon was assigned an identification code and the year of planting was recorded as an attribute. The polygon topology was created and the attribute database was linked to the polygons to make polygon selection in batch mode.

For this polygon based change detection operation the registration of the imagery to a map base was not necessary. An accurate geometric registration of the remotely sensed data and the vector polygon layers in relative sense was sufficient. Therefore, the vector polygon data was transformed to fit the image. Compared to the transformation of multibant images this is a much faster and simpler operation where resampling of pixel brightness values is not required.

Then, training samples were selected in the image with the help of the plantation polygon boundaries. The boundaries of plantation fields were displayed in superimposition with the imagery, and groups of representative pixels were delineated for each theme within designated polygons. The polygons selected for training samples were among those, which exhibited normal tree growth. The PL1 and PL2 plantations consisted of seven and six age groups respectively.

3.3   Statistical Test

Thirteen theme layers were created by a theme-by-theme statistical test of the pixels, each belonging to a certain species and age group. For each pixel, a statistical value was computed using equation (1) and tested against a χ^2 table value at the 2σ and 3σ confidence limits. Pixels, with a computed value falling below the table value, were labelled as a part of the theme they were evaluated for. Those pixels having a higher computed value than the table value were left unclassified. The polygon based assessment was carried out by a program that uses embedded Structured Query Language (SQL) to access information in the relational database via the INGRES DBMS.

The assessment was handled as follows: For each theme layer, the percentages of pixels that passed the test were computed in each polygon belonging to that layer. Passing a statistical test means that the pixel agrees with the apriori expectation. For example, a pixel in a polygon with attribute PL2-80 (species PL2 planted in 1980) resides in the PL2-80 theme. This also illustrates the advantage of theme-by-theme image analysis and the separation of the output into layers. There were many unclassified pixels outside the PL2-80 plantation polygons. Those pixels, however, are not included in the assessment of the PL2-80 plantation polygons. Since the PL2-80 polygons are only assessed in the PL2-80 theme layer, the overlaps of the layers do not adversely affect the results. The degree of agreement was measured by a percentage of correctly classified pixels.

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

- 0-30%, labelled Very Low (VL), signalled major problems. The majority of the trees were probably lost.
- 31-50%, labelled Low (L), signalled unsatisfactory development of the plantation. The stand may be under stress.
- 51-70%, labelled Marginal (M), requires further investigation.
- 71-100%, labelled High (H), could be considered as satisfactory development.

The above categories were chosen for illustrative purposes only. The results are provided in Table 1 and Table 2. Finally, for each polygon, the percentage value and the agreement code were directly entered into the database.

3.4   Results and Discussion

The assessment results derived from the analysis of the output of 3σ range (Table 1) show that most of the polygons have a high percentage of classified pixels. This indicates that most plantations are developing in an expected normal fashion. The polygons belonging to theme PL2-78 show the best agreement with the apriori expectations. None of the polygons belonging to theme PL1-93 fall into the highest percentage category. These are very young plantations with diverse spectral characteristics. Polygons of PL1-81 and PL2-80 also show a low agreement.

Polygon Code	0-30%	31-50%	51-70%	71-100%	Total Polygon
PL1-77	1	2	6	9	18
78	1	2	9	18	30
79	1	-	2	7	10
81	3	2	14	19	38
87	5	-	2	9	16
90	1	1	5	12	19
93	1	1	4	-	6
PL2-77	-	-	2	6	8
78	-	-	-	15	15
79	-	1	6	17	24
80	5	2	-	3	10
81	1	1	3	19	24
92	2	-	3	2	7

Table 1. Number of polygons falling into each of the four categories of agreement with
apriori expectation at 3σ range.

The assessment results derived from the analysis of the output of 2σ range (Table 2) show significant reduction in the number of the polygons falling into the highest category (71-100%) compared to the output of 3σ range. This is not surprising, however, since the number of pixels falling within 2σ decision range would be expected to be significantly less compared to the 3σ decision range. The highest decrease is observed in the theme PL1-81, which has only 2 polygons in the highest percentage category, compared to 19 polygons in the output of 3σ range. A high amount of decrease is also observed in the theme PL1-78. The other themes show various degrees of reduction. In general, the number of polygons falling into lower percentage intervals of 0-30% and 31-50% increased compared with the output of 3σ decision range. Statistical testing performed using 2σ decision range significantly limits the number of pixels falling within the decision boundary. It should be emphasised, however, that 3σ is a generally accepted decision range in the analysis and thus the results shown in Table 1 are the more relevant ones.

Polygon Code	0-30%	31-50%	51-70%	71-100%	Total Polygon
PL1-77	2	7	2	7	18
78	5	10	11	4	30
79	1	1	6	2	10
81	6	13	17	2	38
87	5	2	6	3	16
90	2	4	5	8	19
93	1	5	-	-	6
PL2-77	1	-	4	3	8
78	-	-	5	10	15
79	1	7	8	8	24
80	7	1	-	2	10
81	2	4	5	13	24
92	2	3	-	2	7

Table 2. Number of polygons falling into the four categories of agreement with apriori expectation at 2σ range.

In each theme, the polygons falling into the lowest category (0-30%) were reassessed based on the output of the succeeding years. It was observed that most of these polygons were in closer agreement with the themes of one year younger plantations. Based on this reassessment it appears that some of the plantations were underdeveloped. Better decision about the assessment of these plantation fields can, however, be made by discipline oriented experts.

4 CONCLUSIONS

Of the many benefits of remote sensing and GIS integration, one distinct application is polygon based image analysis. In this study, a per-polygon image analysis scheme for detecting the changes in land cover conditions within existing land use boundaries stored in a GIS was described. Existing land use boundaries stored in a GIS divide the image into homogeneous units, which may not be evident to the analyst and can assist in the selection of training samples and reference data. Theme-by-theme image analysis overcomes the problems caused by multiclassified and misclassified pixels due to overlaps of spectral signatures. Image analysis is carried out in a polygon specific manner. Specific polygons can be selected and the assessments can be performed within those designated polygons only. The scheme does not require multi date imagery. In addition, it provides for an automated updating of the database.

REFERENCES

Danson, F.M., 1989. Factors Effecting the Remotely Sensed Response of Coniferous Forest Canopy. Ph.D. thesis, Department of Geography, University of Sheffield, U.K.

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, Vol. 51, pp. 291-305.

Honer, T.G., 1972. As High-Density Concept and Measure. Canadian Journal of Forest Research, Vol. 2, pp. 441-447.

Jensen, J. R., 1996. Introductory Digital Image Processing: A Remote Sensing Perspective, Prentice Hall, Englewood Cliff, NJ, pp. 257-279.

Leckie, D.J., M.D. Gillis and S.P. Joyce, 1992. A Forest Monitoring System Based on Satellite Imagery. Proceedings of Canadian Symposium on Remote Sensing, Toronto, Ont., pp. 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, Vol. 1, pp. 95-111.

Stenback, J.M. and R.G. Congalton, 1990. Using Thematic Mapper Imagery to Examine Understory. Photogrammetric Engineering and Remote Sensing, 56(9), pp. 1285-1290.

Turker, M., 1997. Polygon Based Image Analysis within an Integrated GIS/RS Environment. Ph.D. thesis, Department of Geodesy and Geomatics Engineering, University of New Brunswick, Canada, 137p.