Identification of Agricultural and Land Cover Database Changes Using Object-oriented Classification Techniques

Luis A. Ruiz, Jorge A. Recio, Txomin Hermosilla, Alfonso Fdez-Sarría

Geo-Environmental Cartography and Remote Sensing Group (http://cgat.webs.upv.es). Department of Cartographic Engineering, Geodesy and Photogrammetry

Polytechnic University of Valencia. Camino de Vera s/n. 46022-Valencia. Spain

Email: laruiz@cgf.upv.es

Abstract – The availability of remote sensing data has increased in quantity and quality over the last several years, allowing for the creation of many different types of cartographic products and geospatial databases. An important concern is how to use the data to update these databases in a regular and efficient manner, reducing costs by increasing the automatic component of the processes. This work is focused on the development and testing of image processing methodologies applied to object oriented classification of aerial and satellite images (QuickBird) for agricultural, land use/land cover spatial database updating. After data pre-processing operations, the objects from the image are defined using cadastral databases. An exhaustive set of features is extracted from every object (parcel), from different types of information: spectral, textural, structural, shape and others. The classification of the objects is carried out by means of decision trees based on the set of features computed. The results obtained after application of these techniques to the classification of agricultural parcels show how automatic tools can be progressively incorporated into database updating processes, reducing field work and providing more efficient and dynamic spatial information management systems.

Keywords: Agricultural databases; feature extraction; object-oriented classification; high-resolution images; texture; structural features.

1. INTRODUCTION

Land use agricultural databases are an essential source of information for natural resource management and landscaping. The maintenance of this type of geodatabases is expensive and time consuming, and the updating rate must be high in order to make them useful. Currently, high-resolution aerial and satellite imagery and airborne lidar data are available on a regular basis, and their acquisition is being partially financed by public administrations, at the regional or national level, in many countries.

The development of automatic feature extraction and object-oriented classification methods can facilitate the progressive improvement of the updating tasks of spatial databases. Li et al. (2002) reported two different approaches for spatial database updating: (i) to gradually create a new database to replace the previous one. This has a high cost in terms of time and financial resources; and (ii) to indentify changes and update only these elements or objects. The second is a more efficient approach and, in the case of agricultural databases, relies on the definition of an appropriate methodology for change detection in parcel uses.

In an object-oriented classification process, image objects can be created by means of automatic segmentation algorithms, or by using available cartographic information, such as cadastral or agricultural cartography. In this case, the limits of the objects have more geographical meaning than regions created with segmentation algorithms, which produce a space division conditioned by sensor attributes instead of the territorial characteristics. The cartographic information contributes to the classification along with the geometry and semantic meaning of the objects. Image classification by parcels has been widely used for agricultural inventory updating (Berberoglu et al., 2000; Blaes et al., 2005; Ozdarici and Turker, 2006; Walter, 2006). This is due, in part, to the relative stability of the limits between adjacent agricultural plots in this type of landscapes, while crops change more frequently.

Parcel-oriented image analysis can be carried out following two different approaches: (i) prior to classification (per–parcel), and (ii) after per-pixel classification. In the per-parcel classification approach, first the descriptive features are calculated for each object, and then a classification method is applied to classify the objects: maximum likelihood algorithm (Walter, 2004; Blaes et al., 2005), neural networks (Berberoglu et al., 2000), decision trees (Hodgson et al., 2003), fuzzy logic (Hofmann et al., 2006), etc. In the per-pixel approach, the complete image is first classified at pixel level. Then, for each parcel, the frequency of the classified pixels is computed and either the majority pixel class is assigned as the label of the field (Berberoglu et al., 2000) or a mixed label describing the heterogeneity of the parcel is assigned (Aplin et al., 1999).

In this paper, we describe a methodology for updating agricultural and land cover databases from images, using cartography to define the objects. Different types of feature extraction algorithms are considered, and object-oriented classification techniques based on decision trees are applied. The main results obtained after the application of these methods over different rural areas and problems are reported and discussed.

2. METHODS

A typical process of land cover database updating can be divided into several phases: (1) Data acquisition and preprocessing; (2) feature extraction; (3) classification; (4) checking, evaluation and edition. The efficiency of tasks 1 and 4 can be substantially improved by the generation of specific support tools. However, tasks 2 and 3 can be carried out, in many cases, using semi-automatic processes, drastically reducing the time spent to achieve reasonable results for production purposes. In the next subsections, we describe the different steps, as well as the tools and methods that can be incorporated to the overall updating process.

2.1 Input Data and Preprocessing Tasks

The type of imagery normally used for these applications are aerial or satellite images with a spatial resolution of about 0.5-1m per pixel. Considering the importance of having spectral information for the discrimination of vegetation density, at least one near infrared band is recommended, in order to be combined with the visible bands and to be able to quantify the presence of vegetation in the parcels. In our tests, DMC (*Digital Mapping Camera*) or QuickBird images have been used in different areas. In addition to the accurate geometric rectification of the images, the generation of mosaics, and the radiometric normalization, the fusion of the panchromatic and multispectral images is usually required in order to achieve a final resolution sufficient to extract proper information from the objects.

Cartographic databases (cadastre, agricultural, etc.) are used to extract the limits that define the objects or parcels to be classified. In some cases, the land use/land cover information contained in the databases can be used as a descriptive feature to include in the classification process, and may improve the global accuracy classification levels (Huang and Jensen, 1997; Rogan et al., 2003; Recio et al., 2009).

Other ancillary data may be useful in some cases, such as digital terrain models (DTM) in areas with topographic roughness, or lidar data to characterize vegetation types in forested areas or tree crops. Preprocessing of lidar data includes filtering of the original points, generation of DTM and canopy height models (CHM).

2.2 Feature Extraction

In image classification processes, and given a set of input data, the selection and adaptation of the classification algorithm is sometimes assumed as the most relevant decision to be made in order to obtain accurate results. However, the extraction of exhaustive, relevant and problem-oriented information from the data is usually the key point in order to increase the accuracy in land use classification. We propose different feature extraction techniques grouped into several categories depending on the kind of information they provide: Spectral, textural, structural, shape, and others. The computation of these features for each of the agricultural plots or parcels is obtained by means of specific software that has been developed for this purpose by the authors (*Fetex v.2.0*).

Spectral features are those that provide information about the spectral reflectance of the object, such as mean and standard deviation of the original or transformed bands (principal components, ratios, etc.), or vegetation indices (NDVI, etc.). Textural features are computed based on first order statistics, and second order statistics, such as those proposed by Haralick et al. (1973) based on the grey level co-occurrence matrix. Other texture features are obtained after applying a wavelet transform to the original images, as reported by Ruiz et al. (2004).

Structural features describe the spatial arrangement and relationships between the elements inside each object. An example is the mutual distance in different directions of the trees that are in an orchard or agricultural plot, creating a regular pattern that may be specific to a class. Extraction of features characterizing regular patterns has been studied by Ruiz et al. (2007). The two main methods we used to extract this information are: (1) Semivariogram features, based on the characterization of the semivariogram graph for each object, which describe the variance of the intensity levels as a function of the distance. An exhaustive set of indices (Durrieu et. al., 2005) has been computed, describing the variance, heterogeneity, or regularity, as shown in the examples of Figure 1. (2) Hough transform features, which are especially designed for the extraction of regular patterns. After specific processing for tree extraction based on the NDVI, this transformation allows for the definition of the principal tree alignments, and their mutual distances following the two principal directions (Figure 2).

Shape information is particularly interesting when certain land use types are related to elongated forms. This is the case of roads and paths, irrigation channels, etc. Descriptive shape indices used are: compactness, shape index and fractal dimension, as described in Recio et al. (2008).

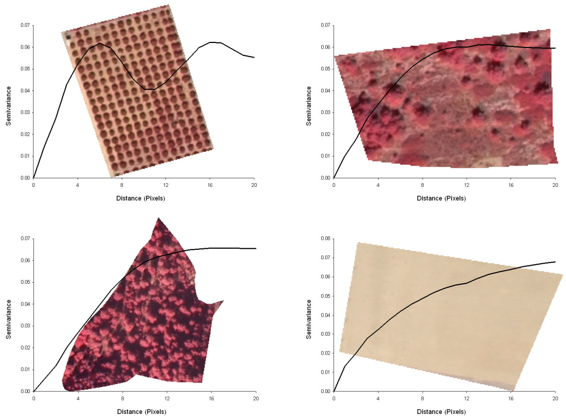

Figure 1. Example of objects with their semivariogram curve superimposed.

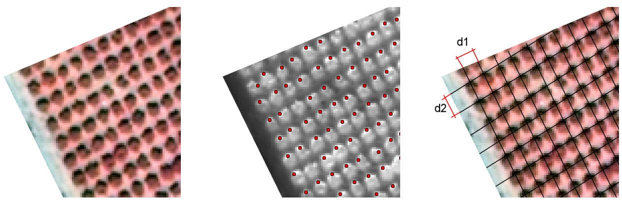

Figure 2. Location of trees on the NDVI image and extraction of distance between trees inside each object.

2.3 Classification and Change Detection

Classification rules are created by means of decision trees, a set of rules organized in a hierarchical manner. Each object is assigned to the class that fits with the rules or conditions, created as a function of the values of the features. Decision trees have been created using the C5 algorithm (Quinlan, 1993). These classification algorithms include in the rules only those features that are relevant, so they perform a feature selection process based on the training samples, avoiding the use of redundant information. Another advantage is that they allow for the inclusion of thematic variables, such as a previous land use of each object that may be available from a database. Based on the proportion of

errors detected in the classification of the training samples, this algorithm creates a confidence level per each set of objects that follow the same branch of rules in the decision tree. Once the objects are classified, the comparison of this new database with the old one allows for the detection of mismatches, that is, those parcels in which the previous land use is different from what is classified. These cases can be reviewed in order to reduce the error rate in the identification of changes. Additionally, the revision could include those groups of parcels that present a low confidence level.
In the next section, we present overall accuracy results of object-oriented classification obtained in four different land use database updating cases in rural areas of Spain.

3. RESULTS AND DISCUSSION

The methodology described has been applied to update changes of *Citrus* tree parcels in the *Citrus* inventory of Castellón, in the Spanish Mediterranean region (case 1 in Table 1). The area was stratified into three main sub-areas with substantial differences in types of crops: mountain, transition and coast. For each sub-area, a different set of training samples was collected. The original data was a mosaic of aerial visible and near-infrared images, as well as the old *Citrus* database, that was used to create the objects. Using an independent set of evaluation samples, the classification accuracy considering only the classes *Citrus* and *Non-Citrus* was 86.6%. After the revision process, the overall accuracy was over 98%, which means that less than 2% of changes remained unidentified.

In a similar agricultural area located in the region of Murcia, in the Southern Mediterranean region (case 2 in Table 1), a test was performed in order to discriminate between seven classes, including two types of tree orchards, vineyards, horticulture, etc. This time, fused satellite multispectral images (QuickBird) were used. Although the overall accuracy was 78.5%, the accuracy in the classification of tree orchards was slightly over 90%. Due to the characteristics of the images, with lower detail than the aerial images, the classification of vineyards was difficult, obtaining low accuracy rates. However, the method and data used were accurate for the discrimination of tree crops from others.

Better results were obtained in a more generic land use classification problem, in the Northwestern region of Spain (Galicia) (case 3 in Table 1), with 5 classes: *Arable lands, shrub lands, forest, buildings/urban*, and *roads*. Using aerial images and objects defined from the agricultural cartography available, the overall accuracy was 89.5%. When information of previous land use from the old agricultural database was included as a thematic feature in the classification decision tree, the overall accuracy increased to 93.7%. An example of the area is shown in Figure 3.

Finally, selecting a particular town from the first area described above, dominated by *Citrus* orchards, arable lands and olive trees (case 4 in Table 1), a classification with a total of 11 classes oriented to agricultural land use was performed. The data used were once again visible and infrared aerial images. The overall accuracy obtained was almost 76%. Classes with high rate of mutual confusion are *olive trees* and *carob trees*, *Citrus* and *young Citrus*, or *small orchards* with *buildings*. Figure 4 shows two detail areas with the results of the classification and the legend used.

Figure 3. Details of an object-oriented classification from aerial images in a problem with 5 generic classes.

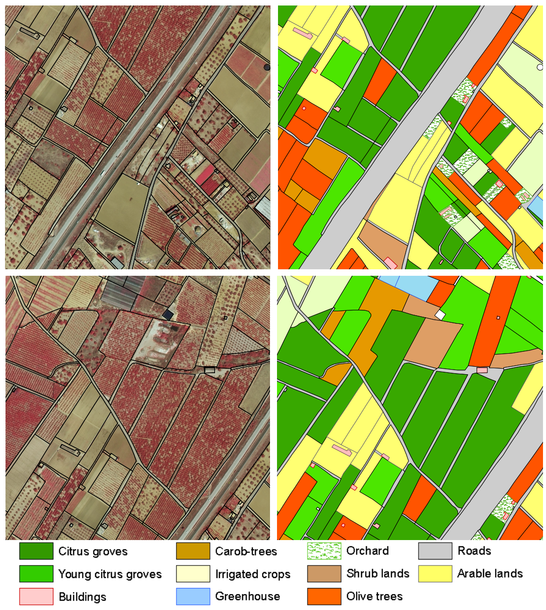

Figure 4. Details of an object-oriented classification from aerial images in a problem with 11 classes.

Case (N. of classes)	**Case 1 (2)**	**Case 2 (7)**	**Case 3 (5)**	**Case 4 (11)**
Overall accuracy	86.6-98.6%	78.5%	89.5-93.7%	75.9%

Table 1. Overall accuracies obtained after applying the methodology described for automatic land cover database change identification in four different study cases and geographic areas.

4. CONCLUSIONS

The methodology described for identification of land cover changes to update land cover/land use agricultural databases can be efficiently applied in those cases where generic classes are defined. When additional classes are introduced, the method can still be used to reduce the field work and photointerpretation tasks. The results also show that these techniques can be applied at production level in the identification of changes for updating inventories of specific crops. In particular, the use of structural features obtained for each parcel is very efficient for tree crop identification. The shape features are crucial for the classification of roads and paths, or irrigation channel networks.

The availability of digital aerial images with infrared information, obtained on a regular basis in many areas, makes the use of these techniques possible at the production level. Satellite images usually have lower spatial information, which reduce the classification accuracy of some crops, like vineyards, where the periodic patterns are better characterized using higher resolution imagery.

In general, object-oriented classification techniques can be progressively used in agricultural database updating, if combined with different degrees of manual interpretation and field work, depending on the complexity of the problem. New, additional data, such as lidar, can be introduced in the future to obtain three-dimensional structural information, increasing accuracies in the discrimination of different types of tree crops and forest areas.

REFERENCES

Aplin, P, Atkinson, P.M., Curran, P.J., 1999. Fine Spatial Resolution Satellite Sensor Imagery for Land Cover Mapping in the United Kingdom. Remote Sensing of Environment, 68 (3), 206-216.

Berberoglu, S., Lloyd, C.D., Atkinson, P.M., Curran, P.J., 2000. The integration of spectral and textural information using neural networks for land cover mapping in the Mediterranean. Computers and Geosciencies 26 (4), 385-396.

Blaes, X., Vanhalle, L., Defourny, P., 2005. Efficiency of crop identification based on optical and SAR image time series. Remote Sensing of Environment 96 (3-4), 352-365.

Durrieu, M., Ruiz, L.A., Balaguer, A. Analysis of geostatistical parameters for texture classification of satellite images. Procs. of the 25th EARSEL Symposium: Global Developments in Environmental Earth Observation from Space. pp. 11-18, 2005.

Haralick, R.M., K Shanmugam and Dinstein, 1973. Texture features for image classification. IEEE Trans. on Systems, Man. and Cibernetics. SMC-3 : pp. 610-622.

Hofmann, P., Strobl, J., Blaschke, T., Kux, H., 2006. Detecting informal settlements from Quickbird data in Rio de Janeiro using an object based approach. 1st International Conference on Object-based Image Analysis (OBIA 2006), Salzburg University, Austria, 4-5 July 2006 (on CDROM).

Hodgson, M., Jensen, J., Tullis, J., Riordan, K., Archer, C., 2003. Synergistic use of Lidar and color aerial photography for mapping urban parcel imperviousness. Photogrammetric Engineering & Remote Sensing, 69 (9), 973-980.

Huang, X., Jensen, J.R., 1997. A Machine-Learning Approach to Automated Knowledge-Base Building for Remote Sensing Image Analysis with GIs Data. Photogrammetric Engineering & Remote Sensing. 63(10), pp. 1185-1194.

Li, D., Sui, H. and Xiao, P., 2002, Automatic Change Detection of Geo-spatial Data from Imagery, International Archives of Photogrammetry and Remote Sensing, Xi'an, China, Vol.XXXIV, Part 2, pp. 245-251.

Ozdarici, A., Turker, M., 2006. Field-based classification of agricultural crops using multi-scale images. In: Proc. First International Conference on Object-based Image Analysis (OBIA 2006), Salzburg, Austria, 4-5 July 2006 (on CDROM).

Quinlan, J.R., 1993. C4.5: Programs for machine learning. Morgan Kaufmann Publishing, San Francisco.

Recio, J.A.; Fdez. Sarriá, A.; Ruiz, L.A.; Hermosilla, T., 2008. Actualización de bases de datos de usos agrícolas mediante la integración de cartografía e imágenes de alta resolución. Congreso Internacional sobre Ingeniería Geomática y Topográfica, TOP-CART 2008. Febrero 2008, Valencia (Spain).

Recio, J.A., Hermosilla, T., Ruiz, L.A., Fdez-Sarría, A. Analysis of the addition of qualitative ancillary data on parcel-based image classification. ISPRS Workshop on High-Resolution Earth Imaging for Geospatial Information. Hannover, 2009.

Rogan, J., Miller, J., Stow, D., Franklin, J., Levien, L., Fischer, C., 2003. Land cover change monitoring in southern California using multitemporal Landsat TM and ancillary data. Photogrammetric Engineering & Remote Sensing. 69(7), pp. 793-804.

Ruiz, L.A, Fdez-Sarría, A., Recio, J.A., 2004. Texture feature extraction for classification of remote sensing data using wavelet decomposition: A comparative study. International Archives of Photogrammetry and Remote Sensing. Vol. XXXV, part B, pp. 1109-1115.

Ruiz, L., Recio, J., Hermosilla, T., 2007. Methods for automatic extraction of regularity patterns and its application to object-oriented image classification. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVI, Part 3/W 49A: 117–121.

Walter, V., 2004. Object-based classification of remote sensing data for change detection. ISPRS Journal of Photogrammetry and Remote Sensing 58 (3-4), 225-238.

ACKNOWLEDGEMENTS

This research was partially funded by the Spanish *Ministerio de Ciencia e Innovación* and the FEDER, in the framework of the projects CTM2006-11767/TECNO and CLG2006-11242-C03-03/BTE. We also thank the *Instituto Murciano de Investigación y Desarrollo Agrario*, *Banco de Terras de Galicia*, *Instituto Cartográfico Valenciano* and *Instituto Geográfico Nacional* for the support and data provided.