

## ADAPTING, SPLITTING AND MERGING CADASTRAL BOUNDARIES ACCORDING TO HOMOGENOUS LULC TYPES DERIVED FROM SPOT 5 DATA

D. Tiede<sup>a\*</sup>, M. S. Moeller<sup>b</sup>, S. Lang<sup>a</sup>, D. Hoelbling<sup>a</sup>

<sup>a</sup> Z\_GIS, Center for Geoinformatics, University Salzburg, Schillerstrasse 30, A 5020 Salzburg - (dirk.tiede, stefan.lang, daniel.hoelbling)@sbg.ac.at

<sup>b</sup> Austrian Academy of Sciences, GIScience, Schillerstrasse 30, A 5020 Salzburg - matthias.moeller@oeaw.ac.at

**KEY WORDS:** LULC, adaptive per-parcel approach, object-based image analysis (OBIA), Cognition Network Language (CNL), object modelling

### ABSTRACT:

The process of satellite-based land use and land cover (LULC) mapping often needs to integrate a priori geo-spatial data. A way to consider pre-existing boundaries is per-parcel classification. In the case of crop monitoring on agricultural fields, a per-parcel approach facilitates the classification process by providing pre-defined, ready-to-use boundaries. In other cases, e.g. when using cadastral boundaries, the outlines may not necessarily coincide with LULC information. In this paper we discuss an adaptive per-parcel approach of object generation based on multi-spectral satellite data incorporating the given outlines. The approach differentiates between three cases: (1) a parcel coincides with a homogenous image object; (2) a set of parcels needs to be merged because of homogeneity of the underlying spectral information (3) a single parcel is spectrally heterogeneous, needs to be split and new boundaries are to be generated by object-specific segmentation. The study was carried out in a 3654 km<sup>2</sup> sized study area covering the Stuttgart Region in Germany. We used orthorectified, mosaicked SPOT 5 multispectral data (5 m ground sample distance; GSD), co-registered and orthorectified. The digital cadastre data were from 2005. We applied object-based image analysis (OBIA) and used cognition network language (CNL) for modelling objects individually in a semi-automated way. About one fifth of the initial cadastre units have been further subdivided due to internal heterogeneity. But the majority of the units have been merged due to redundancy of the boundaries within. By this, the initial number of units has been reduced to less than one fourth. Expert assessment revealed that more than 96 % of the boundaries dissolved were removed correctly. The accuracy of newly introduced boundaries and also the accuracy of retained boundaries was about 86 %. The result met the demand of the given task, although combining data sets of different scales implied some methodological weaknesses. Overall, high potential of the approach can be attributed to the high degree of automation ensuring cost-efficiency, transferability and compatibility of the results. With regard to related applications, further perspectives of using this approach are given in the conclusion.

### 1. INTRODUCTION

For the mapping of land use and land cover (LULC) remote sensing imagery is a useful source of information due to its synoptic capabilities, i.e. the acquisition of information for large areas at one time. In many cases digital mapping data for the area of interest is already available and might be used as an additional source of information. Especially change detection analysis at a given time (*t*) often relies on mapping results acquired at *t-1*. In those cases it is not necessary to perform a complete new LULC classification; the required update is rather limited to objects where significant changes have occurred. Objects with no change could be sorted out and then, in a second step, those objects with changes could be analyzed in detail.

A priori knowledge to be utilized in this process is available in form of pre-existing geo-spatial data representing areal objects with either spectral homogeneous appearance, usage, administration or belonging. In this research we used digital cadastral data, where outlines usually represent property. Although it is often observed, that property is somewhat linked to homogeneous land cover or land use, the property is not necessarily defined by homogeneity in terms of LULC. To tackle these observations we differentiated three cases (1) the cadastral object matches a homogenous image object; (2) several cadastral objects need to be merged because of

homogeneity in the underlying image information; (3) a single cadastral object is spectrally heterogeneous, and therefore needs to be split and new boundaries have to be generated by object-specific segmentation. One premise was that whenever possible, object boundaries should coincide with the original cadastral borders, so only if cadastral limits do indicate a change in category, the respective boundary is retained. Otherwise, if changing categories are not reflected by the cadastral data (e.g. different forest types) a new object border will be generated.

The final result of these steps is a LULC map sharing the same object boundaries as provided by the digital cadastral map, wherever these boundaries indicate changing LULC. Such a LULC map is an accurate and compatible dataset with administrative importance and higher geometric accuracy compared to data sets already available, such as CORINE (Coordinated Information on the (European) Environment).

### 2. PROJECT SETTING, DATA AND METHODS

#### 2.1 Area of interest and aim of the study

This study has been carried out in the 3654 km<sup>2</sup> sized Stuttgart Region located in the south western part of Germany, in the federal state of Baden-Wuerttemberg (Fig. 1). In this area, as in many others of the state, property had been constantly split and

\* Corresponding author.

divided while inherited over centuries, which has led to comparatively small parcels of real estate at present.

The research reported on is embedded in a broader context and reflects the first part of a project called BIMS (Biotope Information- and Management System), which was carried out under the lead of the Stuttgart planning office GÖG (Gruppe für ökologische Gutachten). The project aims at establishing a monitoring system of biotope complexes for observing and assessing the ongoing land-use transformation of biotope complexes (Schumacher and Trautner, 2006). For the automated delineation of these complexes, the study discussed in this paper was a crucial first step.

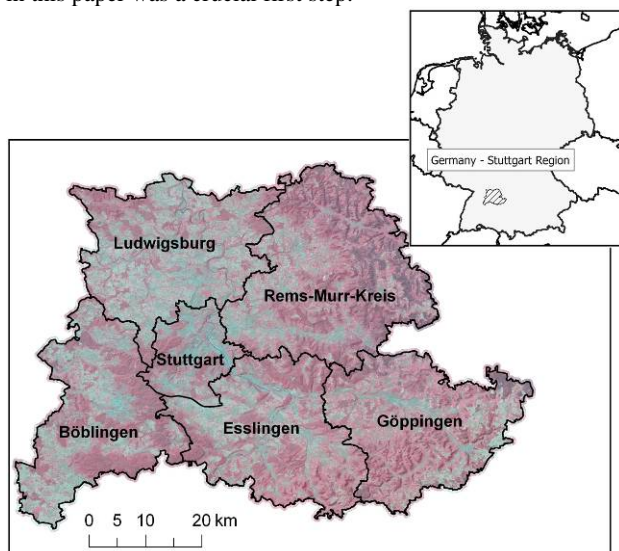


Figure 1. Stuttgart Region and administrative districts in south western Germany

## 2.2 Data Material and Preprocessing

The boundaries of the Automated Cadastral Map (ALK) represent the spatially most accurate geodata available for the entire area of federal Germany. The ALK contains the outlines of objects and the attached Automated Cadastral Book (ALB) lists the ownership information. Each object of both databases is connected via unique IDs. As a logical consequence each parcel (object) listed in the ALK belongs to one owner. In this research we focus only on parcel boundaries, however, there are a number of additional objects listed in the ALK database, e.g. buildings, power-lines, etc. The digital cadastre information used in this study dated from 2005. Some minor parts of the cadastre were not available in the revised form of 2005 and therefore had to be complemented by cadastre information from 2004 (see Fig. 2, left). Whereas using cadastre information from 2004 throughout would have been more agreeable with the date of the satellite data, working on the most recent cadastral data was required from client side. Additionally, in some parts the boundaries between the cadastral data sets of neighbouring administrative districts were inconsistent (Fig. 2, right). The workflow is illustrated in Fig. 3.

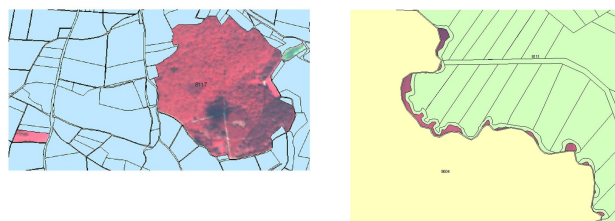


Figure 2. Problems encountered using the digital cadastre data

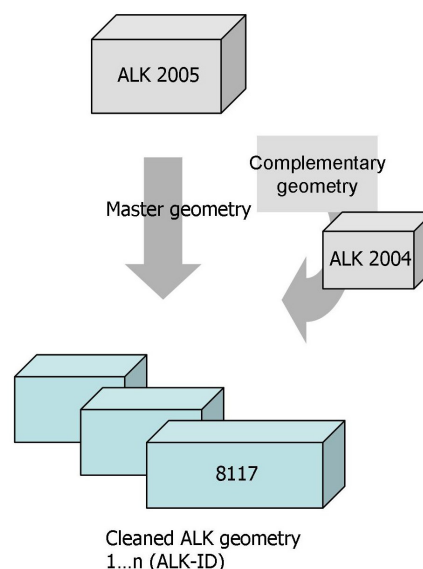


Figure 3. Workflow for preparation of ALK geometry

For determining the actual LULC type multi-spectral remote sensing imagery was used. Four SPOT 5 pan-sharpened (5 m ground sample distance; GSD) color images (scene reference: 53/251, 55/251, 53/252, 55/252), level 1A, acquired between September 2 and 6, 2004, have been obtained. For the task, SPOT 5 data provide sufficient spectral resolution, three bands ranging from the green visible light to the near infrared electromagnetic spectrum. Spatial accuracy could be improved significantly by co-registering the data to an existing orthophoto mosaic (0.25 m GSD) and orthorectifying using a DEM (5 m GSD). As the 5m-DEM did not completely cover the entire area, detected voids were filled using 30m-DEM data (see Fig. 4).

For orthorectification we have parameterised the orbital pushbroom model as implemented in the Leica Photogrammetry Suite. Mosaicking was performed using breakpoints. Finally we clipped the complete SPOT 5 mosaic to the boundaries of the respective administrative district, considering a 500 m buffer around the district outline. The workflow is outlined in Fig. 5.

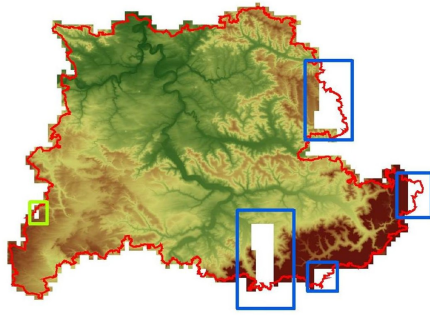


Figure 4. Mosaicked 5 m DEM of Stuttgart Region. Missing tiles (indicated by rectangles) were filled by available 30 m DEM data.

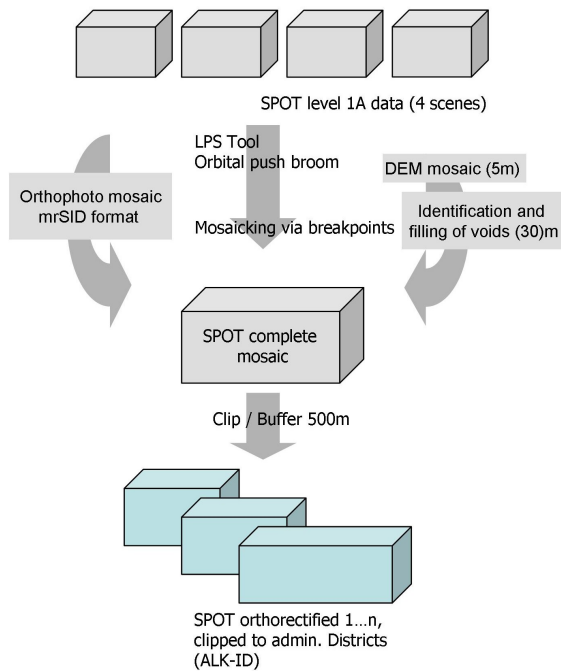


Figure 5. Workflow for orthorectification of the SPOT 5 data: (1) Co-registering to an existing orthophoto mosaic (2) Orthorectification using the DEM mosaic (3). Clipping of the orthorectified SPOT 5 mosaic to the boundaries of the respective administrative district

The resulting SPOT ortho-products showed very high spatial accuracy, the root mean square error of displacement was considerable below 1 pixel. Atmospheric correction has not been applied due to the high quality of the original data material, where no atmospheric disturbances were detectable, and the successful histogram balancing while mosaicking.

Within the framework of the BIMS project, we focused on open areas and forest. Settlements were of no concern. For masking out settlement areas, we used digital ATKIS (administrative topographical cartographical information system) data.

### 2.3 Methods

We used an adaptive per-parcel approach to treat cadastral parcels individually. The method differs from a ‘classical’ per-field (De Wit and Clevers, 2004) or parcel-based (Ozdarici and Turker, 2005) approaches in such a way that heuristics need to

be found to (1) tell mandatory outlines from redundant ones, and (2) introduce new boundaries wherever needed. The critical information for this task is obtained from the underlying multispectral satellite data using object-based image analysis (OBIA, Lang and Blaschke, 2006). Cognition Network Language (CNL) enables to address single objects and their specific behaviour. CNL is a modular programming language in the Definiens Software environment, which supports the development of complex, reproducible and adaptable rule sets (cf. Tiede and Hoffmann, 2006).

In a first step the digital ALK data were used as pre-defined boundaries to perform parcel-based segmentation. Resulting image objects correspond to the cadastre units (Fig. 6). Settlement areas were masked out through a spatial selection process using the respective ATKIS data layer.

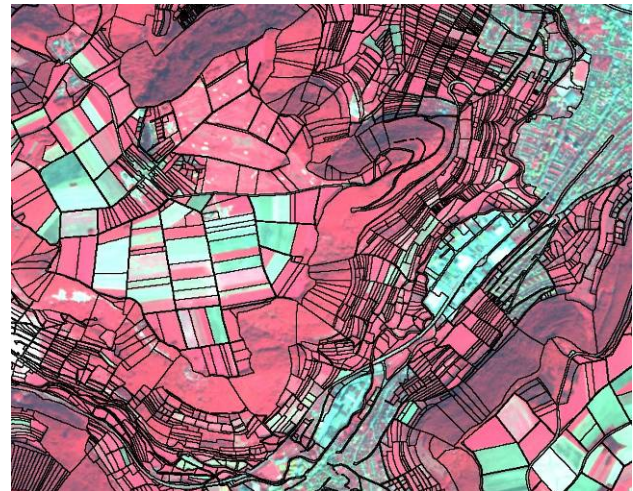


Figure 6. Subset of Spot 5 data set overlaid by ALK data. Settlements were masked out.

#### 2.3.1 Object splitting

In a cyclic object modeling process the produced image objects were checked in terms of spectral homogeneity. In this modeling process each individual object was compared with the spectral values of the underlying image data. In case the standard deviation of the spectral values exceeded a threshold ( $> 10$  in the green band) the object has been marked as a candidate for the splitting operation. Whenever splitting was required, we applied a region-based, local mutual best fitting segmentation (Baatz and Schäpe, 2000) within this object. Elongated objects from the ALK data set like streets or tracks were also selected and split since compact objects were required.

For this domain-specific segmentation (Tiede et al., 2006) we used the following parameters: Scale parameter (SP) = 50; Shape weighting (SW) = 0.5; Compactness weighting (CPW) = 0.5. New objects were generated, which were embedded in the original cadastral object. These new objects are now corresponding to the LULC information derived from the satellite data. Cadastral boundaries, not representing LULC change, are still present (cf. Fig. 7).



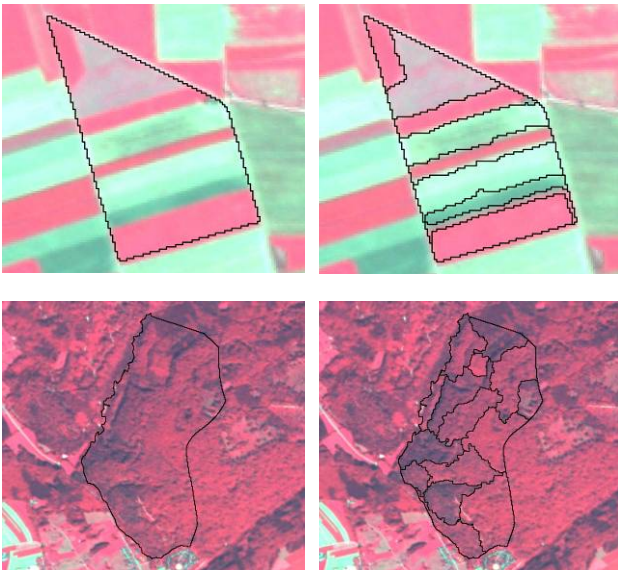


Figure 7. ALK objects before (left) and after (right) the automatic splitting process. Top row: agricultural fields; bottom row: forest.

**2.3.2 Object merge**

To get rid of redundant boundaries an object merge algorithm has been developed. Objects with similar spectral information were grouped into homogeneity classes. The objects in each class were forced to be merged. In contrast to a simple dissolve operation this approach makes it possible to control the merging process: To avoid fairly elongated objects a compactness threshold (CPW = 0.9) and a maximum size (approx. 50 ha) threshold were applied. Cadastral boundaries not corresponding to LULC change were only kept in some cases to limit uncontrolled size and shape growth in the object merge process (cf. Fig. 8)

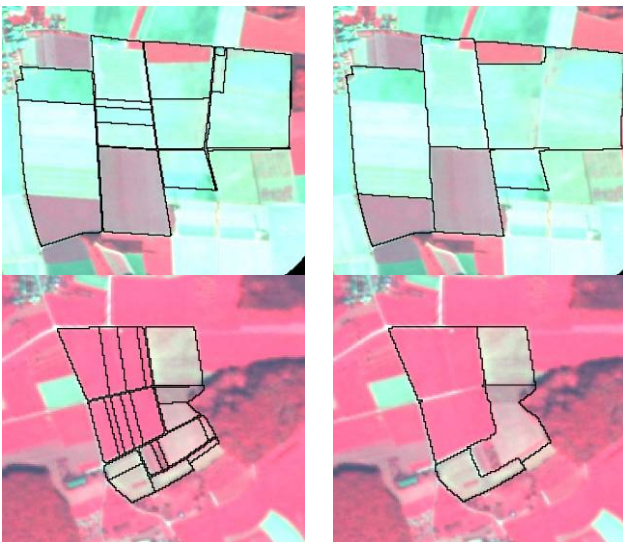


Figure 8. ALK objects before (left) and after (right) the automatic merge and cleaning process of homogeneous objects

**2.3.3 Object cleaning**

Due to the different resolutions introduced by combing SPOT and cadastral data, small objects of the cadastral map could not be incorporated in the modelling process. In addition, in some

cases sliver polygons occurred. The shift in scales was accommodated in a way that very small objects (slivers) were dismissed and a minimum size (2 ha) of generated objects was considered. Figure 9 shows an overview of the entire workflow.

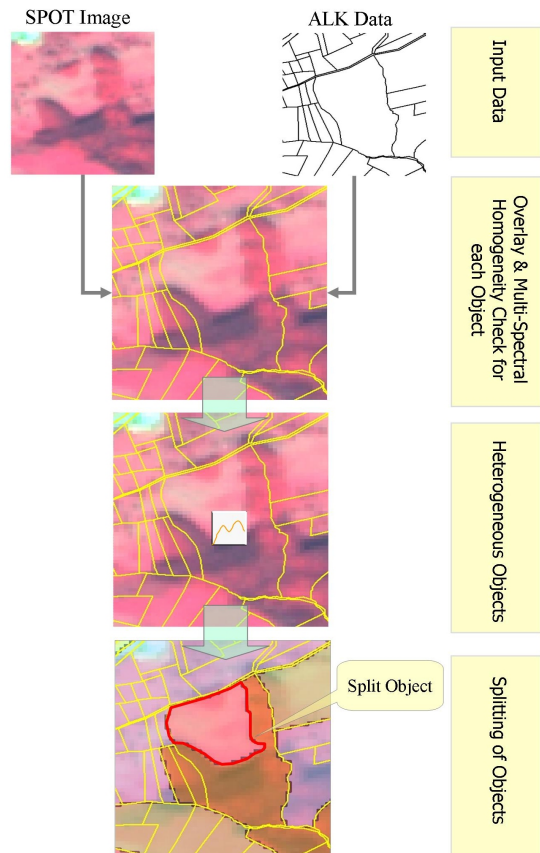


Figure 9. Processing workflow

**3. RESULTS AND DISCUSSION**

The results of the object modeling process fulfilled the given requirements of the project by representing homogenous objects in terms of LULC but sharing cadastral boundary, wherever LULC is changing. Insofar the results of our study have completely met the demands of the client or user. In terms of transferability, we applied the same approach to all administrative districts in the Stuttgart Region with only minor adaptations of parameters in the north western part of the study area.

The following bar chart (Fig. 10) shows the decrease of the number of units for the whole study area by performing object splitting, merge, and cleaning. Whereas in about one fifth (20-25 %) of the cadastral units new boundaries have been introduced, the majority of units were merged due to spectral similarity. This accounted for an overall strong decrease in the number of units, from approximately 1 million units (without urban areas) to less than 220,000.

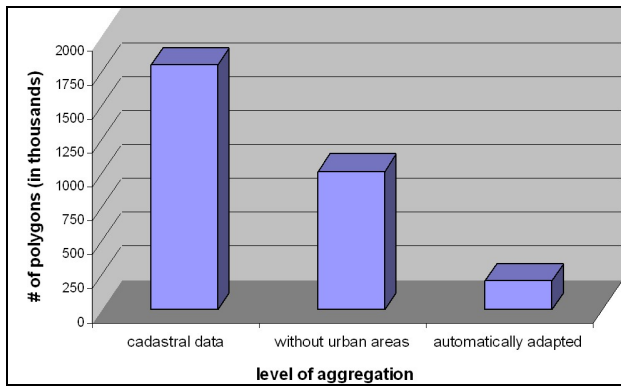


Figure 10. Decrease of the number of units by performing object splitting, merge, and cleaning.

An evaluation of the generated units was carried out by experts for habitat mapping, who used the extracted LULC types as base units for biotope complexes. In case the LULC types were not matching the given criteria for homogenous objects, the borders were changed manually. We used the expert evaluations for validating the approach and carried out an “accuracy assessment” for a subset, which covers about 583 km<sup>2</sup> (~16 % of the study area). The amount of cadastral units (without urban area) in this subset has been reduced from 87,162 to 14,291 automatically derived homogeneous LULC units. After the expert evaluation the final amount of units has again risen to 16,265. Table 1 provides an overview about the lengths of introduced, removed and retained boundaries, compared to the cadastral data and the final evaluation. More than 96 % of removed boundaries were correctly removed, which corresponds to 12,734 km. About 86 % of the newly introduced boundaries were correct (1,510 km) and nearly the same value (86.7 %) applies to the correctly retained boundaries (6,531 km). Missing boundaries, which could not be automatically derived, were digitized manually by the experts summing up to a total length of 348 km.

	Length [km]	%
<b>Introduced boundaries</b>	<b>1,745</b>	
Correctly introduced boundaries	1,510	86.5
Incorrectly introduced boundaries	235	13.5
<b>Removed boundaries</b>	<b>13,258</b>	
Correctly removed boundaries	12,734	96.0
Incorrectly removed boundaries	524	4.0
<b>Retained boundaries</b>	<b>7,531</b>	
Correctly retained boundaries	6,531	86.7
Incorrectly retained boundaries	999	13.3
<b>Missing boundaries (manually introduced)</b>	<b>348</b>	

Table 1. Evaluation of the results for a subset covering around 16 % of the study area (without urban areas).

The high amount and accuracy of correctly removed boundaries shows the potential of the approach to considerably reduce manual work and to obviously increase cost-efficiency substantially.

With regard to the integration of different data sources, there are methodological problems attached to the combination of scales. In other words, minor limitations occurred due to different spatial detail introduced by the combination of SPOT image data and cadastral vector data. The overall accuracy of the optimized vector objects and also the newly introduced boundaries correspond to the spatial image resolution of the SPOT data. Actually, when dealing with cadastral objects, i.e. features with the most detailed outlines, image data with a finer spatial resolution should be used. In our case, a fast and cost-efficient update of LULC types was requested and for this task SPOT image data were considered sufficient enough in terms of spectral and spatial information.

#### 4. CONCLUSION AND FUTURE PERSPECTIVE

The introduced method demonstrates the performance for a semi-automated optimisation of cadastre-level units and can be extended to related research problems. There is high demand of actual map information world-wide in particular in densely populated countries like Germany. The main LULC data in Germany, ATKIS, covers all digital geo-data in a representation equivalent to a scale of about 1:25.000. The DLM (digital landscape model) as part of ATKIS and equivalent to a topographic map sheet needs a frequent update, because it is used in many planning processes. In 1997 the first complete DLM25/1 was established for entire Germany. An update DLM25/3 is planned for the year 2007. At the moment this process comes very close to a halt due to lack of labour. Using very high resolution space borne image data with sufficient spectral range, an update could be performed comparatively fast and with reliable results. At the moment huge efforts are likewise made for updating the historic CORINE LULC data. To this end, the delineations from the mapping campaign carried out in 2000 (Keil et al. 2005) are compared with recently acquired SPOT and IRS image data. Adaptation of our method could lead to some remarkable time and cost savings, because only objects with a significant change will be automatically selected and further examined by an interpreter.

The presented method can also be transferred to different geographical regions; huge demands for the update of map information are almost anywhere on Earth. When thinking about topo-sheet information in less developed countries, the absence of reliable LULC information becomes an issue. Usually there is some historic LULC information available, but usually several decades old and available as printed maps only. After digitizing object boundaries from these maps, an update could be performed based on recent high resolution remote sensing image data and the method described above.

Upcoming research will focus on the update of ALK cadastral data. As mentioned before, this data set is the most accurate spatial data available for Germany. But keeping the data updated is difficult and the combination of existing cadastral ALK data with recent remotely sensed imagery could help tackling this problem. Extremely high resolution airborne scanner data from airborne sensor systems like Vexcel Ultracam, ADS40, DMC and HRSC-AX will be considered for this task. This type of image data is more and more available and provides both: an extremely fine spatial resolution of up to 5 cm pixel size and reliable geometric accuracy in the range of about +/- one pixel (see Moeller, 2003). This new data material

when being used for cadastral updates, make methods as the one described nearly indispensable.

#### REFERENCES

Baatz, M. and A. Schäpe, "Multiresolution Segmentation – an optimization approach for high quality multi-scale image segmentation", in *Angewandte Geographische Informationsverarbeitung XII*, Strobl, J., Blaschke, T., Griesebner, G. (eds.), Wichmann: Heidelberg, 2000, pp. 12-23.

De Wit A. J., J. G. Clevers (2004): Efficiency and accuracy of per-field classification for operational crop mapping. *International Journal of Remote Sensing*. 68(11), pp. 1155-1161.

Keil, M., Kiefl, R., Strunz, G. (2005): CORINE Land Cover 2000 - Germany. Final Report. German Aerospace Center, German Remote Sensing Data Center Oberpfaffenhofen, July 2005.

Lang, S., T. Blaschke (2006) Bridging remote sensing and GIS - what are the most supportive pillars? In: *Proceedings of the 1st International Conference on Object-based Image Analysis*, July 4-5, 2006 in Salzburg. CD-ROM

Moeller, M. (2003): *Urbanes Umweltmonitoring mit digitalen Flugzeugs Scannerdaten*, Book with CD, Wichmann, Karlsruhe, 126 p.

Ozdarici A., M. Turker, Comparison of different spatial resolution images for parcel-based crop mapping. *ISPRS Workshop, Commission II, WG2, Spatial/Spatio-Temporal Data Mining (SDM) and Learning*, November 24-25, Ankara, Turkey, CD.

Schumacher, J. and J. Trautner (2006): Spatial Modeling for the purpose of regional planning using species related expert knowledge. The Biotope Information- and Management System of Stuttgart Region (BIMS) and its deduction from the Information System on Target Species in Baden-Württemberg In: Buhmann, E., Ervin, S., Jørgensen, I., Strobl, J. (eds.): *Trends in Knowledge-Based Landscape Modeling*. Wichmann-Verlag, Heidelberg

Tiede, D., C. Hoffmann (2006): Process oriented object-based algorithms for single tree detection using laser scanning data"; *EARSel-Proceedings of the Workshop on 3D Remote Sensing in Forestry*, 14th-15th Feb 2006, Vienna, 162-167.

Tiede, D., S. Lang, C. Hoffmann (2006): Supervised and forest type-specific multi-scale segmentation for a one-level-representation of single trees. In: *Proceedings of the 1st International Conference on Object-based Image Analysis*, July 4-5, 2006 in Salzburg. CD-ROM

#### ACKNOWLEDGEMENTS

This study has been carried out within the project Biotope Information- and Management System (BIMS), financed through the Verband Region Stuttgart (contact: Mrs. Weidenbacher) for the purpose of a regular update of the regional plan. We thank Jens Schumacher from the *Gruppe für ökologische Gutachten* for fruitful discussions during the course of the project and his effective project management.