GIS-DRIVEN CLASSIFICATION OF SATELLITE IMAGERY

Radoux J.*and Defourny P.

Earth and Life Institute, Université catholique de Louvain, Croix du Sud, 2/16, Louvain-la-Neuve, B-1348 Belgium
WGIV/4

Abstract : With the increasing amount of geographic vector database available over the world, the remote sensing community has a strong potential to quickly produce up-to-date geographic products. In this paper, we propose a geographic object-based image analysis to bridge the gap between the image processing and the geographic vector database. The key component of the method is the automated sample selection refined by a statistical outlier detection algorithm. A preliminary image labelling is performed based on the intersection between a geographic vector database and the result of a multiscale image segmentation. Each class is then cleaned by a non-parametric trimming procedure that iteratively excludes outliers in order to keep only the most relevant training samples. Eventually, each image-object is classified based on a maximum likelyhood and new classes are created by decision rules. This method was succesfully applied to two study areas in a temperate and a Mediterranean region. The object-based thematic overall accuracy was close to 75%, which was as good as the accuracy of the vector data in the temperate region and a significant improvement in the Mediterranean area. The proposed framework showed very promising results, but further work could still improve the classification step in order to take the most out of the selected training samples.

1 INTRODUCTION

Ecosystem and land management increasingly rely on analysis, statistics and modelling outputs derived from GIS database for decision support and strategic planning (Hese and Schmullius, 2009; Lowell and Astroth, 1989; Aydöner and Maktav, 2009). The quality of the geospatial database is therefore of paramount importance in the decision making process. These data have to be appropriate in terms of thematic content, date and scale. For example, in an agricultural area, harvesting induces a land cover change but not a land use change. This land cover change, from crop to bare soil, has a dramatic impact on water run off that has to be accounted for in hydrological processes. As another example, urbanisation results in both a land use and a land cover change by the construction of new buildings and up to date information is here crucial for land use planning. Finally, scale issues should not be neglected in analysis. For instance, they can appear along ecotones such as forest boundaries. If the scale is too fine, it is not possible to unambiguously delineate them and if it is too coarse the precision is not sufficient to detect progressive change.

The first geospatial data were produced by ground surveys, which have integrated modern technologies such as GNSS or TPS for accurate positioning and are now collected in digital format inside geographic vector database. Remote sensing from airborne or spaceborne sensors is a possible alternative to these ground surveys. The main advantages of remote sensing are its ability to cover large remote areas and to provide updated observations on a frequent basis. Nowadays, the (semi-)automated interpretation of remotely sensed imagery to provide high quality geographic database is a main challenge of the earth observation community. Depending on the sensor type, the challenges of the workflow may consern different stages. In the case of very high spatial resolution optical images, the main issue is typically the spectral variability inside a given class, which is, for instance, particularly high in forests due to the differences in crown lighting or in urban areas due to the diversity of small spatial objects.

Since the early days of image processing, supervised classification of earth observation data relies on a highly interactive step to delineate calibration or training sample set used to transform the image radiance values into a map labelled according to the selected typology. While unsupervised method are also common, the most recent developments concern supervised classification algorithms such as support vector machine, maximum likelihood, k-Nearest Neighbours or artificial neural networks (Huang et al., 2002; Blanzieri and Melgani, 2008; Ashish et al., 2009) which all rely on quality training samples (Frery et al., 2009). On the other hand, the increasing availability of earth observation imagery calls for mass processing chains requiring the automation of the relevant training data set selection. Huang et al. (2008) proposed to automatically select the training samples in forest based on the dark object concept, that is assuming that the darkest vegetation type is always the forest. Unfortunately, this kind of *a priori* assumption is only feasible for a limited number of land cover classes. Another approach consisted in processing a chain of Landsat image one by one based on a first classification and extracting the training sample of the next image on the overlapping region with a classified image (Knorn et al., 2009). In this case, the main disadvantage was a loss of accuracy at each step.

On the other hand, ancillary data have been used to also improve the automation of change detection workflows, for instance using machine learning to build a knowledge base (Huang and Jensen, 1997). Nevertheless, image-to-vector land cover change detection of automated classification are still rare, in spite of the great potential of conflation and integration methods. In this case, the information from two digital maps are combined to produce a third map that is better than both initial maps (Saalfeld, 1988; Cobb et al., 1998). Based on matching algorithms, these methods are used for reducing boundary conflicts (Butenuth et al., 2007) or network registration (Chen, C.-C. and Knoblock, C.A. and Shahabi, C., 2006) using, e.g., edge detection filter and snakes. Some applications used both image and vector data to detect the change (Schardt et al., 1998; Knudsen T., 2003), but image classification often remained a necessary intermediate step for conflation.

Recent studies achieved automated detection of new buildings based on colour information (chrominance) extracted from matching buildings between a vector database and an aerial photograph (Ceresola et al., 2005). A similar approach was used to classify land cover with training areas extracted from the vector database (Walter, 1998, 2004), but in this case the presence of discrepancies inside the training dataset and the spectral heterogeneity of the land cover classes reduced the accuracy of the classification. Geographic Object-Based Image Analysis (GEOBIA) addresses this issue by processing groups of pixels as single objects instead of independent individual pixels (Hay and Castilla, 2008). Based

*Corresponding author, Julien.Radoux@uclouvain.be, Fax : +32(0)10 47 88 98

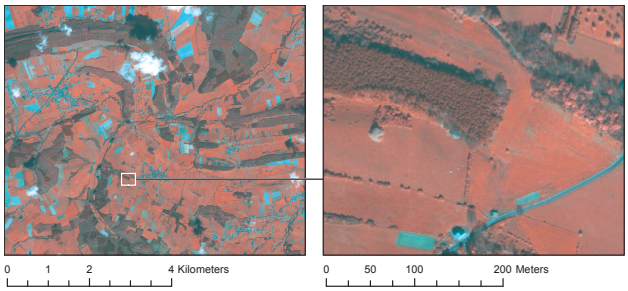

Figure 1: False color composite of the Quickbird image showing a subset of the study area and a detail of the pansharpened image.

these groups of pixels, called image-objects, more robust characteristics and faster classification can be expected.

In this study, we propose a methodological framework that combines the information of an existing geospatial vector data with remote sensing data in order to build a new updated and upgraded vector database. We aimed at selecting consistent training samples using an automated procedure. This procedure was based on a non-parametric probabilistic approach to perform an unsupervised screening of image-objects to train a supervised classification.

2 DATA AND STUDY AREA

Two study areas were considered in this paper. The first one is located in a region of Southern Belgium called "Famenne" and the second in a region of Southern France called the "Massif des Maures". In both cases, a Quickbird image was processed using a land cover GIS database. The respective type of landscapes and GIS database were however quite different.

The method was implemented on a 40 km^2 study area in Famenne. The landscape in Famenne is very fragmented. It mainly consists in small villages, crop fields, pastures and temperates forests. These forests include deciduous and coniferous stands, dominated by oaks, beeches and spruces. Figure 1 illustrates this diversity near one of the villages. We used two fine scale geographic database to test the proposed method. The image was a Quickbird scene acquired in April 2007, at the beginning of the vegetation period for deciduous trees, and is partly covered by clouds. It was orthorectified by the RPC model refined with a first order polynomial based on 5 sub-metric ground control points. The panchromatic and the multispectral images were then resampled at 70 cm with a cubic convolution and pansharpened using the Bayesian Data Fusion algorithm(Fasbender et al., 2008). The image was compared with a 1:10 000 vector database of 2005 produced by the Belgian National Geographic Institute. It included 31 land use classes that were grouped into 21 land cover classes. The main discrepancy expected between the image and the geospatial vector database is the presence of hedges and isolated trees dispersed in the landscape. These spatial objects were not included in the polygon layer of the vector database but are observed on the image. On the other hand, some land cover changes can be expected because all the forest in the area are exploited, so there are some clear cut. Furthermore, there can be small land use changes such as new roads or buildings may also be present, but but since the landscape is not quickly changing, these are expected to be very limited. At the end, two land cover classes were added to the 21 existing classes, one for forest change and one for hedges and isolated trees, as well as a "no data" class for clouds and shadows.

The 300 km^2 test area of the "Massif des Maures" was used to assess the method performance in another context. This landscape is mainly composed of mediterranean forest, with dispersed residential buildings along the coast and vineyards in between. The image, acquired in June 2006, had already been orthorectified by the French National Geographic Institute. This image was cloudless and the main differences expected between the image and the map were the recently burnt and the cleared forest area, which are key elements for the fire fighters. The vector database used for image processing was CORINE Land Cover 2000, a 1:100 000 European land cover map with a hierarchical legend. The five first-level classes of CORINE and an additional "degraded forest" class were used in the study.

3 METHOD

The proposed method is outlined in figure 2 and the different steps are detailed in this section. The segmentation process and the selected parameters are described in section 3.1. The resulting image-objects are then labelled based on the vector database in order to build the training samples which are cleaned by iterative trimming (section 3.2). Eventually, the process was completed by the classification of all image-objects according to the selected land cover typology (section 3.3).

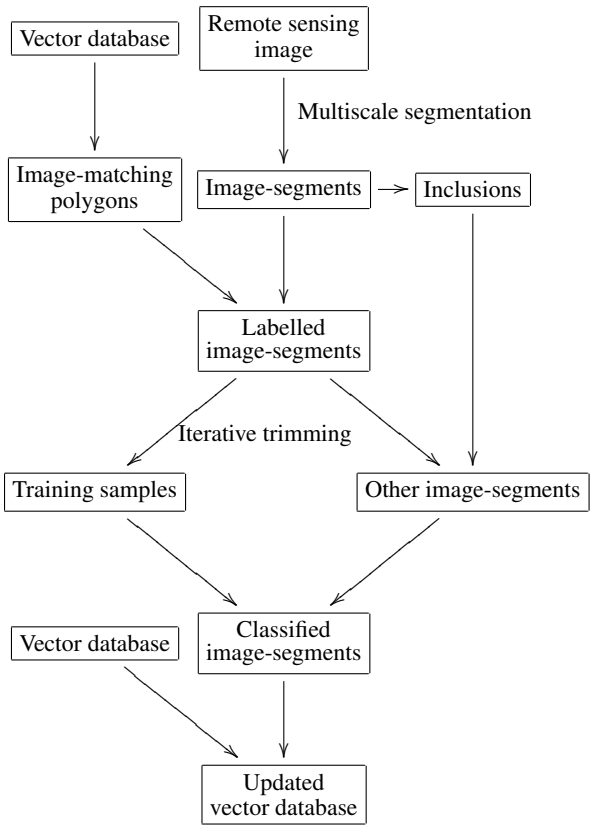

Figure 2: Outline of the method

3.1 Segmentation

The segmentation of remote sensing image is a critical step in geographic object-based image analysis. In the perspective of land cover classification, image-objects should be as close as possible to their homologous spatial objects as represented in the vector map. For the currently existing segmentation algorithms, this is difficult to achieve in a single-step segmentation because spatial objects have very different sizes and textures. When the image-objects are too large (under-segmentation), they may be overlap with different spatial objects (mixed objects) or include small artefacts (inclusions). On the other hand, if image-objects

are too small, they may only partially correspond to the spatial object (over-segmentation), which reduces the performances of object-based classification. Instead of aiming at the single best set of parameters, we took advantage of the multi-scale structure of the segmentation algorithm to adapt our segmentation result to the different land covers.

In this sudy, the multi-resolution segmentation algorithm was used for the segmentation (Baatz and Schäpe, 2000). This segmentation algorithm can be tuned based on three parameters: the scale, the weight of shape constraint and the type of shape constraint (smoothness or compactness). The scale parameter is a threshold on the shape and the spectral variability of each object. Spectrally homogeneous object therefore tend to be larger than spectrally heterogenous ones. The shape constraint can help to reduce the effect of the spectral variability and is therefore useful to reduce over-segmentation of spectrally heterogenous spatial objects, such as the forest. In order to build the most reliable statistics for object-based classification, the objects need to include as many pixels of the same class as possible. The inclusions in the objects were detected based on a three level image-object hierarchy.

The lower level was built first and focused on the spectral similarity of adjacent pixels. The scale parameter was 30 and shape was not accounted for at this stage. The upper level was then built with a large scale parameter (100) and a compact shape (shape parameter 0.3). A previous study (Radoux and Defourny, 2008) on the choice of segmentation parameters indeed showed that a high compactness was particularly useful to delineate textured objects. Finally, an intermediate level was created by merging the image-objects of the lower level based on the spectral difference between them. The mean spectral values of image-objects from the intermediate level were then compared with the statistics of the corresponding image-object in the upper level in order to detect inclusions. With this goal in mind, let us assume that the pixels inside each upper-level image-object follow a Gaussian distribution, which is reasonable because, by construction, image-objects consist in a large number of pixels of similar spectral value. After estimating the parameter of this distribution from the observed pixels, a 95% confidence interval of the pixel values was built for each upper-level image-objects ($\mu+/-1.96\sigma$). If the mean value of the intermediate-object is outside of the interval, it is considered as an outlier image-object inclusion and the statistics of the upper-level objects are updated with the remaining intermediate-level objects. In the text, the term "image-object" will refer to these upper-level image-objects with statistics computed without accounting for the outlier image-objects inclusions from the intermediate level.

3.2 Automated training sample selection

The vector database explicitly contains the thematic information to be extracted in the map. The transfer of this information was performed based on a GIS-driven analysis. However, an attribute transfer assumes that there are no discrepancies between the image and the vector database, which is obviously not the case when the vector database needs to be updated. A statistical outlier detection was therefore performed in order to keep only the representative image-objects for each class.

The GIS-driven analysis assumed that the most likely land cover class for an image-object is the class of the vector database that covers the majority of its area. A spatial intersection of the database-objects and the image-objects was thus performed and the area of each class was computed for each image-object. Each image-object was then labelled according to the class with the highest proportion in terms of area. The labelled objects were considered as relevant class samples when at least 90 % of their area was covered by the same class from the vector database.

After the preliminary labelling, a non-parametric iterative trimming procedure was used to exclude outliers from the potential training samples in order to avoid training sample contamination by discrepancies. This statistical approach consists in computing the probability density function (*pdf*) of a class and excluding the outliers in an iterative way so that the estimation of the *pdf* is improved at each step. At each step, the *pdf* is thus estimated with the remaining observations and the new outliers are detected based on these improved parameters. A threshold (α) is chosen to find the minimum *pdf* values likely to belong to the main distribution and the observations with a smaller *pdf* value are excluded accordingly(equation 1).

$$\left\{ \sum_{\{x \in \mathbb{R}^d | f(x) \geq t\}} f(x) \Delta x \right\} - (1 - \alpha) \tag{1}$$

where $f(x)$ is the *pdf*. Contrary to Desclée et al. (2006) and Duveiller et al. (2008) who used Gaussian distribution in forest change detection, we used non-parametric distributions to adapt to more complex cases. These non-parametric *pdf* were computed on a regular grid by replacing each observation by a Gaussian kernel (Silverman, 1986). More details on the implementation can be found in Radoux and Defourny (2010). It is worth noting that each observation was weigthed by its area in order to avoid over-evaluating the contribution of a large group of small image-objects inside a group of a few very large objects (isolated trees in a crop field, boats on the sea).

In Famenne, two different statistical thresholds were tested. The first threshold was 0.03, which has a low sensitivity to outliers but a low probability to exclude image-objects really belonging to the class. The second threshold was set to 0.15, which is likely to detect most of the outliers but with a risk of higher proportion of false positive. The sensitivity of the subsequent classification results was assessed with these two opposite parametrisations. Due to the large proportion of change expected in the Maures region, only 0.15 was used there.

3.3 Classification

The classification of the image-objects proceeded in two steps. Based on the image-object characteristics of the training sets, the first step classified the image-objects not included in the training samples. The second step refined the resulting classification based on the GIS-driven probabilities and some specific classification rules based for handling discrepancies.

The three most discriminant characteristics were selected for the classification of all objects. The maximum likelihood classification algorithm, based on the theorem of Bayes (equation 2), is then used to find the most probable class C_i conditionnally to the image-object characteristics X.

$$P(C_i|X) = P(X|C_i)P(C_i)/P(X) \tag{2}$$

where $P(C_i|X)$ is the probability to observe C_i conditionally to X and $P(X)$ is a normalising value that can be ignored in this case. Contrary to the classical approach that could be limited by its Gaussian assumption Frery et al. (2009), the non-parametric *pdf* was computed for each class based on the trimmed training

samples in order to estimate $P(X|C_i)$. Furthermore, as we assumed that the vector database is globally reliable, we estimated the prior $P(C_i)$ based on the proportion of each class in the vector database, contrary to many application of maximum likelihood classification that need to assume that all classes have the same probabiliy due to a lack of information. With a large number of classes, the use of this prior information is of paramount importance to reduce the risks of contamination of the classification by very unlikely classes.

After the maximum likelihood classification, a set of decision rules was applied to the classified image-objects in order to definitively label the image objects. These rules were of two types: refining the classification based on the geographic vector database and handling discrepancies that require the creation of new classes.

3.4 Accuracy assessment

The thematic accuracy of the method was assessed based on a probabilistic sample of image-objects selected from the object list. Based on the distribution of the size of the objects and the number of objects in the maps, the number of sampling units was set to 700. The reference dataset was created by photointerpretation after overlaying the image-object boundaries on the pansharpened image. Photo-interpretation ambiguities were resolved by a complementary field survey conduced in 2008 with a GPS. In the "Maures" region, there was no field validation but the photointerpretation of the 1200 image-objects selected on the Quickbird image was assisted by a reference vector database from the local fire fighters. The object-based classification accuracy (CA), reflecting the performance of the classifier, and the object-based overall accuracy (OA), reflecting the quality of the resulting map (Radoux et al., xxxx), were computed based on equations 3 and 4, respectively.

$$\widehat{CA} = \frac{1}{n}\sum_{i=1}^{n} C_i \tag{3}$$

$$\widehat{OA} = \frac{1}{s_T}\left(\sum_{i=1}^{n} C_i S_i + \widehat{CA}\sum_{i=n+1}^{N} S_i\right) \tag{4}$$

where n is the number of sampling units and N the total number of objects in the map; $C_i = 1$ when the i^{th} image-object is correctly labelled and 0 otherwise; s_T is the total area of the map and S_i the area of the i^{th} image-object.

4 RESULTS

The outliers detection in the multiscale segmentation scheme improved the class characteristics by reducing the internal class variance. This effect was particularly useful for the standard deviation of the pixels inside image-objects, but also for the mean values when the area of the artefacts was large compared with the upper-level image-objects. Around 15% of the intermediate level image-objects were detected as inclusion and this detection was particularly effective in homogeneous classes. Figure 3 shows some typical inclusions for different land cover classes. On the other hand, about 10% of the image-objects selected in the reference database of the Famenne site were considered as mixed. This occurred when two different land cover shared approximately the same area inside the upper segmentation level, so that lower-level image-objects could not be separated.

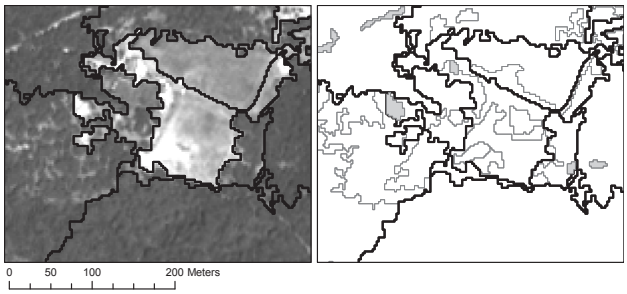

Figure 3: Inclusions (in grey) extracted from the upper-level image-objects (bold black) illustrated on a subset in the Maures region: road segments (top left), buildings (centre) and forest gaps (bottom right).

Figure 4: Preliminary labelling illustrated on a subset of the Famenne study area. From left to right, Quickbird image, labelled image-objects and vector database from the Belgian NGI.

About 60% of the image-objects had an overlapping proportion above 90% and all classes were represented in the training set except railways and rivers in Famenne that were always hidden by tree alignments. However, the class "reed-bed" was remove from analysis because it only corresponded to three objects. Figure 4 shows the candidate training samples resulting from the intersection of the images-objects and the vector database in Famenne.

In Famenne, the iterative trimming excluded respectively 6.3 and 30.3% of the potential training samples with 0.03 and 0.15 thresholds, respectively. In case of unfrequent land cover classes (less than 30 image-objects), the trimming did not modify the class distribution because their variability was too high to decide either or not a single observation should be part of the distribution. In the "Maures" region, the only statistical threshold tested was 0.15, resulting in screening out 25% of the potential training image-objects. Qualitatively, the correspondance between the image-object outliers and the spatial object outliers was better in the Maures than in Famenne, where the clear cut detection in deciduous forest was very poor due to the absence of leaves on deciduous trees at this period.

The results of the classification workflow were satisfactory, but the thematic overall accuracy was improved thanks to *a priori* probabilities derived from the vector database. In Famenne, the values of the overall and classification accuracy were higher for $\alpha = 0.03$ (see table 1). As the vector database was very good at the beginning, it can be seen that it is difficult to improve its overall accuracy when adding new classesas they are additional sources of errors. On the other hand, it is worth noting that the overall accuracy increased from 75 to 77% in the best case when the mixed objects were not accounted for in the reference data set. In the Maures region, the object-based overall accuracy was of 74%, which significantly improved the overal accuracy of the original vector database ($\approx 60\%$ if we consider that the CORINE land cover had to include a "degraded forest" class). Table 2 shows the confusion matrix per image-objects.

The geolocation quality was not quantitatively assessed, but the

Table 1: Overal and classification accuracy (n=694) in the Famenne study area for different scenarios

Process	OA (%)	CA (%)
$\alpha = 0.03$ without decision rules	71	71
$\alpha = 0.15$ without decision rules	70	69
$\alpha = 0.03$ with decision rules	75	75
$\alpha = 0.15$ with decision rules	72	72
Original vector database	75	75

Table 2: Confusion matrix (n = 1200) for the classification of the multispectral Quickbird image in the Maures region

	Urb.	Agri.	For.	Water	Degr.F.
Urban	259	10	25	4	28
Agriculture	27	150	49	0	36
Forest	2	0	261	0	42
Water	1	0	0	24	0
Degraded forest	39	33	57	0	153

results were consistent with the photo-interpretation. In the Famenne area, the high quality of the geographic vector database was maintained in the results. It was therefore preferable to keep the generalized lines. In the Maures, however, there was an obvious improvement of the object delineation, as illustrated on figure 5.

5 DISCUSSION

The proposed workflow showed that geographic object-based image analysis could help to bridge the gap between remote sensing image interpretation and GIS database. For both case studies, the object-based overall accuracy was close to 75 % without any interactive training data set selection. Most of the errors concerned the additional classes (degraded forest, hedges, clear cuts...) not initially included in the existing database and artefacts such as the clouds and their shadows. It is also worth noting that the confusion with "water" in the Maures region were due to shallow water along the coast. These results are thus promising towards an operational image processing, yet different parts of the process could still be improved.

- Segmentation improved the labelling automation and the class discrimination. Despite the performance of the segmentation algorithm, it remained a critical step due to the necessity to adjust several parameters. The main issue is the creation of mixed image-objects. The use of a multi-scale approach was able to strongly reduce the number of these mixed image-objects, but was not able to remove all of them.

- The second issue is the selection of image-object characteristics for trimming and classification. More than 10 different variables were used as first and second best choices for the trimming of the 22 different classes (the 21 existing land cover classes and the expected shadows). However, only three variables were included for the classification because of the selected algorithm performance. Other feature selection method or classification algorithms better adapted to large dimensional spaces (kNN, SVM) could possibly provide better results.

Nevertheless, the proposed sample selection method proved to be robust and can benefit from the numerous existing vector database

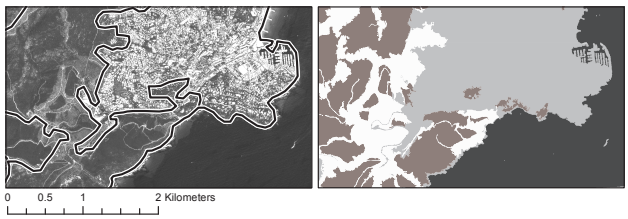

Figure 5: New delineation from the GEOBIA (right) compared with the original data (left)

to train any classification algorithm. It is then of paramount importance to properly define the legend, and fully documented land cover typology supported by classification systems such as LCCS (Di Gregorio and Jansen, 2000) should help to better incorporate to information for a direct image-to-map comparison. Otherwise, it is necessary to adapt the land use class of the geographic vector database into land cover classes before performing the analysis. This was done for the NGI vector database, which was primarily a land use geographic database. Furthermore, it is worth noting the ambiguity of the forest class, which are not necessarily covered by tree in their land use definition.

The proposed method could also be used to facilitate map update by detecting the discrepancies between the image and the map for each class. As the values of the trimmed *pdf* are computed for every image-object and every class, these can be stored in the database to document the likelihood of the discrepancies.

6 CONCLUSION

The proposed method provided a GEOBIA framework to combine the information of a remote sensing image with a vector database. Thanks to its flexible statistical approach, the method could be used on different study areas without changing its implementation. After building appropriate image-objects and optional decision rules for the labelling of new classes, the all process did not require any manual labelling and relied on a single tuning parameter. Further work is necessary to select the best parameters for the classification or to use classification algorithms designed for large dimensionnality. As the proposed method provides large and clean training samples, these methods would be used in optimal conditions.

ACKNOWLEDGEMENTS

The authors thank the Belgian Scientific Policy (ORFEO/ASSIMIV project) and the Université catholique de Louvain for funding. We also thank the Belgian NGI for providing the reference data in the Famenne and Mj Poppi for his precious help on the Maures region.

References

Ashish, D., McClendon, R. W. and Hoogenboom, G., 2009. Land-use classification of multispectral aerial images using artificial neural networks. *International journal of remote sensing* 30(8), pp. 1989–2004.

Aydöner, C. and Maktav, D., 2009. The role of the integration of remote sensing and gis in land use/land cover analysis after an earthquake. *International Journal of Remote Sensing* 30(7), pp. 1697–1717.

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

Blanzieri, E. and Melgani, F., 2008. Nearest neighbor classification of remote sensing images with the maximal margin principle. *IEEE Transactions on geoscience and remote sensing* 46(6), pp. 1804–1811.

Butenuth, M., Gösseln, G., Tiedge, M., Heipke, C., Lipeck, U. and Sester, M., 2007. Integration of heterogeneous geospatial data in a federated database. *Journal of photogrammetry and remote sensing* 62(1), pp. 328–346.

Ceresola, S., Fusiello, A., Bicego, M., Belussi, A. and Murino, V., 2005. Automatic updating of urban vector maps. *In: F. Roli and S. Vitulano (eds), ICIAP 2005, LNCS 3617*, pp. 1133–1139.

Chen, C.-C. and Knoblock, C.A. and Shahabi, C., 2006. Automatically conflating road vector data with orthoimagery. *Geoinformatica* 10(4), pp. 495–530.

Cobb, M., Chung, M., Foley, H., Petry, F., Shaw, K. and Miller, V., 1998. A rule-based approach for the conflation of attributed vector data. *GeoInformatics* 2(1), pp. 7–35.

Desclée, B., Bogaert, P. and Defourny, P., 2006. Forest change detection by statistical object-based method. *Remote Sensing of Environment 102(1-2)*, pp. 1–11.

Duveiller, G., Defourny, P., Desclée, B. and Mayaux, P., 2008. Tropical deforestation in the congo basin: national and ecosystem-specific estimates by advanced processing of systematically-distributed landsat extracts. *Remote Sensing of Environment.*

Fasbender, D., Radoux, J. and Bogaert, P., 2008. Bayesian data fusion for adaptable image pansharpening. *IEEE transaction on geoscience and remote sensing* 46(6), pp. 1847–1857.

Frery, A., Ferrero, S. and Bustos, O., 2009. The influence of training errors, context and number of bands in the accuracy of image classification. *International journal of remote sensing* 30(6), pp. 1425–1440.

Hay, G. and Castilla, G., 2008. Object-based image analysis. *Springer, Verlag Berlin Heidelberg, chapter Geographic object-based image analysis(GEOBIA): an new name for a new discipline*, pp. 76–89.

Hese, S. and Schmullius, C., 2009. High spatial resolution image object classification for terrestrial oil spill contamination mapping in West Siberia. *International journal of applied earth observation and geoinformation* 11(2), pp. 130–141.

Huang, C., Davis, L. and Townshend, J., 2002. An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing* 23(4), pp. 725–749.

Huang, C., Song, K., Kim, S., Townshend, J., Davis, P., Masek, J. and Goward, S., 2008. Use of a dark object concept and support vector machines to automate forest cover change analysis. *Remote Sensing of Environment* 112, pp. 970–985.

Huang, X. and Jensen, J., 1997. A machine-learning approach to automated knowledge-base building for remote sensing image analysis with gis data. *Photogrammetric engineering and remote sensing* 63(10), pp. 1185–1194.

Knorn, J., Rabe, A., Radeloff, V., Kuemmerle, T., Kozak, J. and Hostert, P., 2009. Land cover mapping of large areas using chain classification of neighboring landsat satellite images. *Remote Sensing of Environment* 113(5), pp. 957 – 964.

Knudsen T., O. B., 2003. Automated change detection for updates of digital map databases. *Photogrammetric engineering and remote sensing* 69(11), pp. 1289–1296.

Lowell, K. and Astroth, J., 1989. Vegetative succession and controlled fire in a glades ecosystem: geographic information system approach. *International Journal of Geographic Information Systems* 3(1), pp. 69–81.

Di Gregorio, A. and Jansen, L., 2000. Land cover classification system (lccs): Classification concepts and user manual. *GCP/RAF/287/ITA Africover-East Africa Project and Soil Resources, Management and Conservation Service, Food and Agriculture Organization.*

Radoux, J. and Defourny, P., 2008. Object-based image analysis. *Springer, Verlag Berlin Heidelberg, chapter Quality assessment of segmentation devoted to object-based classification*, pp. 257–272.

Radoux, J. and Defourny, P., 2010. Image-to-map conflict detection using iterative trimming. *Photogrammetric Engineering and Remote Sensing* 176, pp. 173–181.

Radoux, J., Bogaert, P., Fasbender, D. and Defourny, P., xxxx. Thematic object-based overall accuracy assessment. *International Journal of Geographical Information Sciences* xx((in revision)), pp. xxxx.

Saalfeld, A., 1988. Conflation, automated map compilation. *International Journal of Geographical Information Systems* 2(3), pp. 217–28.

Schardt, M., Kenneweg, H., Faber, L. and Sagischewski, H., 1998. Fusion of different data-level in geographic information system. *In: D. Fritsch, M. Englich and M. Sester (eds), ISPRS Commission IV symposium on GIS - Between visions and applications.*

Silverman, B., 1986. *Density estimation for statistics and data analysis.* Chapman & Hall.

Walter, V., 1998. Automatic revision of remote sensing data for gis database revision. *In: D. Fritsch, M. Englich and M. Sester (eds), ISPRS Commission IV symposium on GIS - Between visions and applications.*

Walter, V., 2004. Object-based classification of remote sensing data for change detection. *ISPRS Journal of photogrammetry and remote sensing* 58, pp. 225–238.