OBJECT-BASED VERIFICATION AND UPDATE OF A LARGE-SCALE TOPOGRAPHIC DATABASE

S. Müller, C. Heipke

IPI – Institute of Photogrammetry and GeoInformation, Leibniz Universität Hannover,
Nienburger Str. 1, 30167 Hannover, Germany, {mueller, heipke}@ipi.uni-hannover.de

Commission IV, WG IV/2

KEY WORDS: Change detection, updating, databases, aerial, reference data

ABSTRACT:

It is well known that there is a growing need for consistent and up-to-date GIS-data at various scales. This paper describes a system for semiautomatic quality assessment and update of an existing large scale topographic database. The necessary reference information is derived from current digital aerial orthophotos with a ground sampling distance of 0.1m via automatic image analysis. The advantage of the system is that it reduces the manual efforts of a human operator to a minimum. To overcome the limitations of the automatic components the human interaction is focused on those objects, for which no reliable assessment result can be achieved within the automatic verification process. The efficiency of the whole system depends on the performance of the used image analysis component, because the number of correct objects for which not enough evidence can be found in the imagery is the limiting efficiency factor. The focus of the paper is on the assessment of the most important object classes of a large scale topographic database maintained at the Ministry of Municipal and Rural Affairs (MOMRA), Kingdom of Saudi Arabia. Two different applications are implemented and evaluated. The first is the verification of the existing GIS-data and the second is the update of the GIS-data. Both applications were implemented using the software GeoAIDA that is a so-called knowledge-based system which was developed at the Institut für Informationsverarbeitung (TNT), Leibniz Universität Hannover, Germany. The automatic image analysis is based on a supervised texture classification. Manual evaluation results of the verification and the update approach of the MOMRA GIS-data in different test regions are shown and the results are discussed.

1. INTRODUCTION

In this paper two applications are implemented that are described in the following.

The quality of GIS-data is essential for virtually all applications. To validate the consistency of GIS-data and reality normally manual inspection of the data is necessary. To reduce the manual effort during verification a system called WiPKA (knowledge-based photogrammetric-cartographic workstation) (Busch et al., 2004) (Busch et al., 2006) was developed at the Leibniz Universität Hannover. The system was originally developed for verification and update of GIS-data with a scale of 1:25,000. In this paper cadastral data with very high resolution are verified. The adaption of the system to high resolution GIS-data is described and represents the first impelmented application. The used vector data stem from the large scale database of the Ministry of Municipal and Rural Affairs (MOMRA), Kingdom of Saudi Arabia. The data include polygon, line and point objects. In the paper only polygonal objects are processed. The necessary reference information is taken from current orthophotos with a ground sampling distance of 0.1m.

Update of GIS-data is just as important as quality control. For the second application again the system WiPKA is employed and tested with the large scale MOMRA database. During the update task new objects still missing from the database are searched for. The procedure runs automatically using the same orthoimages as reference information.

The funcitionality of automatic verification and upate of existing GIS-data using image data as reference information is currently not available in commercial GIS systems. The paper shows that parts of manual human efforts can be replaced with the proposed approach.

2. DATA USED

2.1 Raster Data

The available raster data consists of aerial orthophotos with a ground sampling distance of 0.1m with the spectral channels red, green and blue. Figure 1 shows an example of the high resolution orthophotos. The images stem from the region Huraymila, Kingdom of Saudi Arabia.

Figure 1: Example for the RGB orthophotos.

2.1 Vector Data

The used MOMRA database covers an area of 182ha and includes polygon, line and point data from various object classes. In total the test site includes 4497 objects. In the following, an extract of the object class catalogue is given:

Polygonal objects:

- building under construction
- building-general-outline
- building-government
- building-ruin
- car park-paved outline
- car park-unpaved outline
- garden polygon
- mosque building
- misc-structure
- palm grove polygon
- swimming pool outline
- recreation area outline

Line objects:

- bridge
- concrete edge
- ditch
- road unpaved edge

Point objects:

- car park symbol
- chimney symbol
- contour index annotation
- mosque symbol

The vector data are available as shapefiles and in DGN format. In Figure 2 the original orthophoto is illustrated superimposed with an extract of all vector data object classes.

Figure 2: Orthophoto superimposed with vector data.

3. APPROACH

3.1 Verification

The workflow of the investigated application is analogous to that of the system WiPKA and is illustrated in Figure 3. It works as follows: Current image data is used as reference information and together with GIS-data serves as input for an automatic image analysis component. In this component image processing operators are applied to the image data. The results of these operators are subsequently used to evaluate single GIS-data objects. The evaluation of a GIS-data object is done by help of an evaluation catalogue. The result of the automatic image analysis component is an evaluation ("accept" or "reject") of each investigated GIS-data object, given as a so-called traffic light diagnostics: accepted objects receive a green light, rejected objects a red one. Rejected objects can then be visualized using a standard GIS e.g. ArcGIS, and a human operator decides manually, if they are correctly rejected.

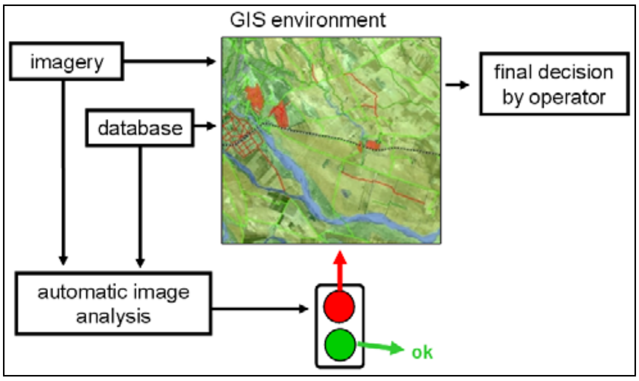

Figure 3: GIS-data verification workflow.

With the current version of WiPKA the following GIS-data object classes of the delivered vector data could be processed:

- building-general-outline
- garden-polygon
- misc-structure
- mosque building
- recreation area outline
- palm grove polygon

The system WIPKA is based on the image anlysis system GeoAIDA (Bückner et al., 2002). GeoAIDA is a so-called knowledge-based system using semantic networks for knowledge respresentation. The image anlysis as well as the assessment of the vector data is performed by GeoAIDA. The whole procedure is described in the following.

First the GIS-data objects of interest are selected from the vector data. In Figure 4 the original orthophoto serving as input is illustrated. In Figure 5 it is shown superimposed with the GIS-data to be verified. For the verification application a supervised texture classification algorithm (Gimel'farb, 1997) is applied to the orthophoto. In this approach, Markov random fields in combination with Gibbs-potentials are used for modelling individual pixels and relations between pixel pairs in a local neighbourhood. This approach has shown to be an adequate means to describe texture properties of high resolution orthophotos (Müller, 2007). The optimal Gibbs-potentials are learnt from given samples applying a maximum likelihood estimation. A segmentation and labelling of a given image consists in finding piecewise homogenous regions using a maximum a posteriori (MAP) estimation which involves simulated annealing.

In Figure 6 the per-pixel result of the texture classification operator is illustrated. The following four classes were manually trained before applying the texture operator:

- building
- vegetation
- street surface
- soil surface

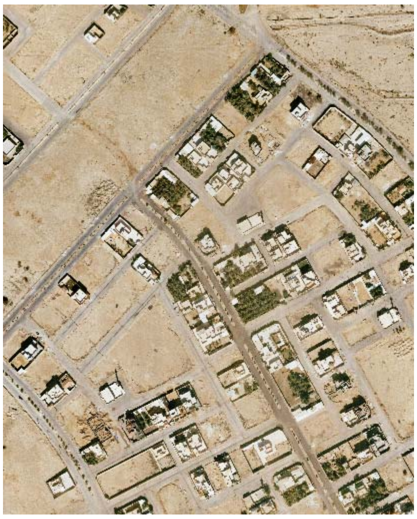

Figure 4: Orthophoto used as input: Original RGB.

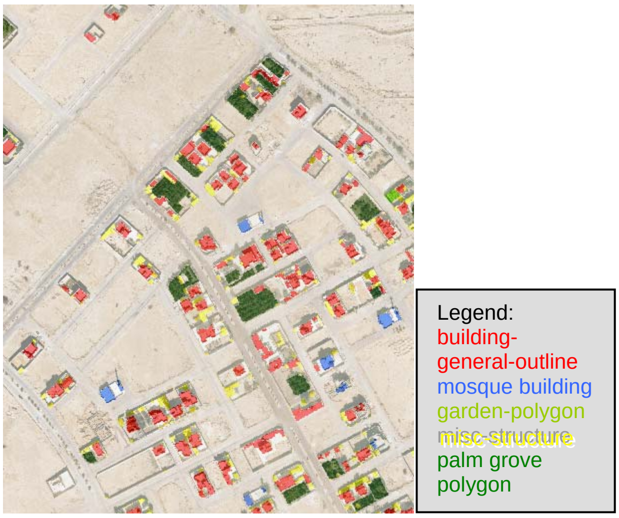

Figure 5: Orthophoto superimposed with GIS-data.

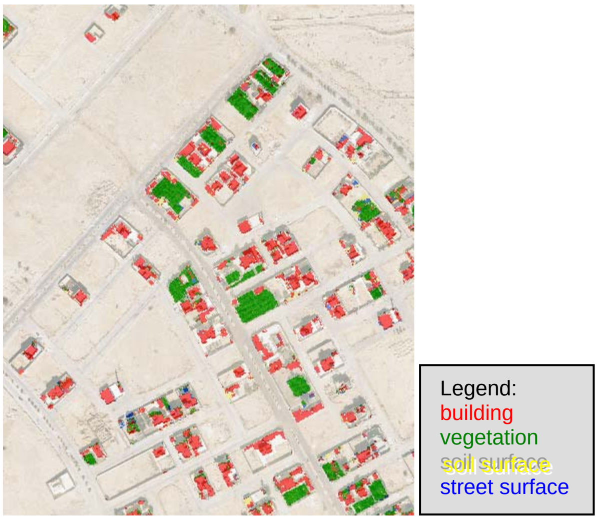

Figure 6: Orthophoto superimposed with texture classification (results shown only inside GIS data depicted in Figure 5).

The subsequent evaluation of single GIS-data objects is based on an evaluation catalogue shown in Table 1. For example an object of the GIS-data object class “building-general-outline” is accepted if at least 50% of all pixels belong to the texture class building. A palm grove polygon, as another example, is rejected if more than 65% inside the object do not belong to the texture class vegetation. The explicit distinction between all GIS-data object classes is not possible. For example a misc-structure object labeled as mosque building leads to no rejection but will be accepted.

GIS-data object class	**correct texture classes**	**max. error rate [% pixel]**
building-general-outline	building	50
garden-polygon	vegetation, soil-surface	50
misc-structure	building	50
mosque building	building	50
recreation area outline	vegetation, soil-surface	50
palm grove polygon	vegetation	65

Table 1: Evaluation catalogue.

3.2 Update

While during verification the quality control of existing GIS-data is investigated, during update new objects still missing from the database are searched for. Therefore, the second implemented application, called “GIS-update”, works without any GIS-data as input. A comparison between the GIS-data automatically extracted from the image and the available GIS-data then provides a measure of quality of the automatic update process.

First, the same test site as in the verification application is investigated. Again the knowledge-based system GeoAIDA and the orthophoto are used. Since the second application works without GIS-data, the semantic network had to be adapted. In comparison to the verification application the semantic network used here does not contain nodes that refer to existing GIS-data objects. **Fehler! Verweisquelle konnte nicht gefunden werden.**In the update application, the nodes building_update, street_update and vegetation-update represent new GIS-data objects.

The supervised texture classification algorithm (Gimel’farb, 1997) is used in the same way as in the verification application. Thus, the result of the texture classification is the same as before and is shown in Figure 6. Based on this per-pixel result new GIS-data objects are searched. For the update application the new object classes correspond to the classes of the texture operator. Thus, the possible new object classes are:

- new building object
- new vegetation object
- new street object

The texture class soil surface is not taken as an object class but as simple background. For a new GIS-data object to be established a number of conditions have to be fulfilled. These conditions are collected in Table 2. For example, a new building object has to belong to the texture class building and has to have a minimum size of 60m² and a minimum width of 6m. A new street object has to be classified as texture class street surface with a minimum size of 1000m², a minimum width of 3m and a maximum width of 40m. The used thresholds have been derived from studying the MOMRA feature catalogue.

GIS-data object class	correct texture class	min. area [m²]	min. width [m]	max. width [m]
new building object	building	60	6	---
new vegetation object	vegetation	200	14	---
new street object	street surface	1000	3	40

Table 2: Conditions for new GIS-data objects.

4. RESULTS

4.1 Verification Results

The evaluation result for the input orthophoto (cp. Figure 4) is shown in Figure 7 in terms of the traffic light diagnostics. Only very few objects are rejected and thus marked in red.

Figure 7: Orthophoto superimposed with traffic light diagnostics (cp. Figure 3).

For an independent evaluation of the results of the automatic verification approach Table 3 should be consulted. Here, the manual decision of a human operator is compared to the automatic evaluation of the system WiPKA for each single GIS-data object. If an object is accepted by both, the human and the machine, the decision is called “true positive”. Automatically accepted objects that are manually rejected are “false positive” decisions and constitute undetected errors. All objects that are rejected automatically need manual inspection. The true positive decisions represent the efficiency of the system.

GIS-data object	automatic verification result: acceptance	automatic verification result: rejection
decision by human operator: correct	true positive (efficiency)	false negative (interactive final check)
decision by human operator: incorrect	false positive (undetected errors)	true negative (interactive final check)

Table 3: Evaluation of traffic light diagnostics.

For the Huraymila test site no reference data about the GIS-data objects was available. A manual inspection of the GIS-data led to the result, that the data is very precise and contains nearly no errors. This leads to a simpler evaluation matrix shown in Table 4. 65.5% of all GIS-data objects are automatically accepted, accordingly 34.5% are rejected and false negative decisions. Most of the rejected objects belong to the class misc_structure (86.7%). For the 182ha the computing time on an Intel Core2Quad 2.4GHz CPU system was 96min.

The existence of the relatively large number of rejected objects belonging to the object class misc_structure has several reasons:

- sometimes misc_structure objects are covered by vegetation and therefore not visible in the orthophoto
- often misc-structure objects are of quite small size (e.g. 5m²) and have no structure/texture and therefore cannot be properly recognised using the employed texture classification approach

<table>
<tr><th>GIS-data object</th><th>automatic verification result: acceptance</th><th colspan="2">automatic verification result: rejection</th></tr>
<tr><td rowspan="3">decision by human operator: correct</td><td rowspan="3">2946 objects 65.5%</td><td colspan="2">1551 objects 34.5%</td></tr>
<tr><td>1345 objects 86.7%</td><td>misc_structure</td></tr>
<tr><td>161 objects 10.4%</td><td>building-general-outline</td></tr>
</table>

Table 4: Evaluation result for complete Huraymila test site, 4497 objects.

In Table 5 the same evaluation is made without the object class misc-structure. The number of false negative decisions is reduced to 17.3%. Some examples for still appearing false negative decisions are:

- palm_grove_polygon objects that were recently planted and/or contain only few vegetation pixels
- building-general-outline objects with rare roof materials
- recreation_area_outline polygons having no contrast to the surrounding terrain and hence not being visible in the orthophoto

GIS-data object	verification result: acceptance	verification result: rejection
decision by human operator: correct	983 objects 82.7%	206 objects 17.3%

Table 5: Evaluation result for complete Huraymila test site without object class misc-structure, 1189 objects.

4.2 Update Results

Figure 8 shows all automatically detected new GIS-data objects in the region shown of Figure 4. The comparison of the existing GIS-data (Figure 5) to the automatically detected new objects shows a good correspondence of the data for buildings and vegetation objects. Nearly all existing structures in the scene could be automatically detected and given the right object class. However, street objects could not be detected successfully due to the lack of contrast with respect to bare soil. Some additional difference between the two datasets can be observed:

- complex building structures are detected as one contiguous object - in the existing GIS-data these objects consist of several single objects
- again misc-structure objects are sometimes too small to be automatically detected

Table 6 shows the update results of a manual evaluation for a sample size of about 44ha. For the reason stated obove, the object class new street object was not considered in the evaluation. The detection rate for automatically detected objects in comparison to the existing GIS-data is 92.8%. The rate of false alarms, that are regions wrongly classified as objects, is 7.2% in this case.

Figure 8: Automatically detected new objects.

new building/ vegetation object	automatically detected	not automatically detected
existing in reference GIS-data	64 objects 92.8%	5 objects 7.2%
not existing in reference GIS-data	5 objects 7.2%	---

Table 6: Evaluation of GIS-data update results for building and vegetation objects, sample size 44ha.

5. CONCLUSIONS AND OUTLOOK

The automatic verification results show that about 65% of all objects could be verified automatically. This leads to a productivity gain of about 200%. In order not to commit any undetectable errors the automatic process decided to call in a human operator for a final decision for about 35% of all GIS-data objects. Most of these objects belong to the GIS-data object class misc-structure that are often covered by vegetation and therefore are not visible in the image data. Moreover misc_structure objects are mostly of quite small size (e.g. 5m²) and have no significant structure or texture to detect them automatically. The number of objects needing manual intervention drops to only about 17% without considering the class misc_structure. This result is comparable with the experience made in the WiPKA project during the last years using data from around the world.

The update results show, that for two investigated object classes new building object and new vegetation object the detection rate is about 93% at a rate of false alarms of about 7%. However, not surprisingly the results were found not to be less favourable when compared to the verification application. In a number of cases the approach has difficulties to detect new street objects automatically due to the small contrast between streets and the surroundings. Also, the number of object classes that can be detected automatically is reduced in comparison to the MOMRA object class catalogue. In addition, complex building structures were sometimes detected as one contiguous object, whereas they may be composed of different single buildings in reality.

As a next step it should be ascertained that the results achieved in this project are truly representative for the whole Kingdom of Saudi Arabia, or at least for its fast developing suburban areas. In this regard more empirical tests are needed.

This work has been carried out using very high aerial orthophotos with a ground sampling distance of 0.1m. This is not to say, however, that space images could not be used at all for verifying or updating the MOMRA database. High resolution space images, e.g. from the IKONOS, Quickbird or Wordview satellites, should be investigated in future. Additionally, more object classes will also be investigated whether additional object classes can be automatically verified and/or updated, and with which quality.

6. LITERATURE

Busch, A.; Gerke, M.; Grünreich, D.; Heipke, C.; Liedtke, C. E.; Müller, S., 2004: Automated Verification of a Topographic Reference Dataset: System Design and Practical Results. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XXXV, Part B2, Istanbul, pp. 735-740.

Busch, A.; Gerke, M.; Grünreich, D.; Heipke, C.; Helmholz, P.; Liedtke, C.-E.; Müller, S., 2006: Automated Verification of a topographic dataset using IKONOS imagery. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XXXVI/4, Goa, pp. 134-139.

Bückner, J.; Pahl, M.; Stahlhut, O.; Liedtke, C.-E., 2002: A knowledge-based system for context dependent evaluation of remote sensing data. In: L. J. v. Gool (ed.), DAGM-Symposium, Lecture Notes in Computer Science, Vol. 2449, Springer, Zurich, Switzerland, pp. 58–65.

Gimel'farb, G. L., 1997: Gibbs fields with multiple pairwise pixel interactions for texture simulation and segmentation. In: Rapport de recherche RR-3202, INRIA, Sophia Antipolis France.

Müller, S., 2007: Extraktion baulich geprägter Flächen aus Fernerkundungsdaten zur Qualitätssicherung flächenhafter Geobasisdaten, Dissertation, ibidem-Verlag, Stuttgart.