Change Detection and Integration of Topographic Updates from ATKIS to geoscientific data sets

Sester, Monika; Gösseln, Guido von

Institut für Kartographie und Geoinformatik, Universität Hannover

1. Overview

Although different data sources are usually involved when geoscientific or environmental problems have to be solved, they can be a disadvantage at the same time. Despite the fact that all geoscientific data sets containing topographic information rely on the same source, the earth surface, they show significant differences due to different acquisition methods, formats and thematic focus, different sensors, level of generalisation, and even different interpretation of a human operator. Another problem which occurs while working with different data sets is the problem of temporal consistency: Even if the data sets are originally related to the same objects, different update cycles in the different thematic data sets lead to significant discrepancies. Observing this problem it is obvious that harmonisation, change detection and updating of different data sets is necessary to ensure consistency, but hardly practicable when performed manual.

This project deals with different aspects of data integration, namely integration of different vector data sets, integration of vector and raster data, as well as providing an underlying data structure in terms of a federated data base, allowing a separate, autonomous storage of the data, however linked and integrated by adapted reconciliation functions for analysis and queries on the different data sets (Sester et al., 2003).

In this paper, which gives an overview of the work done so far, there will be a concentration on the work of the Institute of Cartography and Geoinformatics (ikg), namely the semantic and geometric integration of vector data: Methods for the automatic integration, change detection and update between heterogeneous data sets.

2. Used data sets

As it has been specified in (Sester et al., 2003) three data sets are used in this project all at a scale of 1:25000:

- **ATKIS** - the topographic data set,
- **GK** - the geological map and
- **BK** - the soil-science map.

When going from analogue to digital maps, new possibilities for data handling and analysis appear: basically, the combination of different data sets in a geo-information system (GIS) is enabled.

Simple superimposition of different data sets already reveals visible differences (Fig. 1). These differences can be explained by comparing the creation of the geological, the soil-science map and ATKIS (Goesseln & Sester, 2003).

As for ATKIS the topography is the main thematic focus, for the geo-scientific maps it is either geology or soil science – however they are related to the underlying topography. The connection between the data sets has been achieved by copying the thematic information from topographic to the geo-scientific maps at that point of time the geological or soil-science information is collected.

While the geological content of these data sets will keep its actuality for decades, the topographic information in these maps do not: In general, topographic updates are not integrated unless new geological information has to be inserted in these data sets (Ad-Hoc, 1994 & 2002).

The update period of the feature classes in ATKIS varies from one year up to three months – in general, 10% of the objects have to be updated per year (LGN 2003).

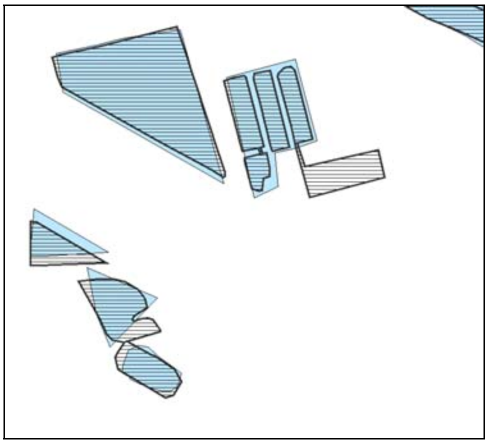

Fig. 1 : Simple superimposition of ATKIS (dark border, hatched) and geological map GK 25 (solid fill).

These differences in acquisition, creation and updating lead to discrepancies, making these data sets difficult to integrate. In order to identify changes in the data sets and update the changes, the following steps are needed: identification of corresponding objects in the different data sets, classification of possible changes, and finally update of the changes.

3. Data Integration

3.1 Semantic Integration

Firstly, semantic differences between these data sets must be described to avoid comparing “apples and oranges”.

Enabling the adaptation of updates from one data set to another leads to the problem of integration of heterogeneous data sets. Regarding to (Walter & Fritsch, 1999) the integration task which must be applied to these data-sets is classified as the most elaborate, handling data sets which are stored in heterogeneous sources and differ in data modelling, scale, thematic content, acquisition method, accuracy and temporal update.

In the first phase of this project, the topographic feature class “water areas” has been chosen as a candidate for developing and testing (other will follow), because of the existence of this topographic element in all data sets. To ensure a correct and fully automatic process, the detection of changes and the correct linking between semantic partners is a must.

3.2 Geometric Integration

Following the semantic integration, differences in geometric representation have to be identified and removed. Geological and soil-science maps are single-layered data sets which consist only of polygons with attribute tables, while ATKIS is a multi-layered data-structure with points, lines and polygons, together with attribute tables. The different data models used in ATKIS and the geoscientific data sets are resulting in more discrepancies in the geometric representation requiring a harmonisation procedure before the establishing of links between corresponding objects could be done.

3.3 Harmonisation

Water objects in ATKIS are represented in two different ways: Water areas and rivers exceeding a certain width are represented as polygons. Thinner rivers are digitised as lines and are assigned additional attributes, referring to some classified ranges of widths. The representation of water objects in the geo-scientific maps is always a polygon. For the first implementation a simple buffer algorithm has been chosen, using the line representation from ATKIS as centre line and the width attribute. Another problem is the representation of grouped objects in different maps resulting in different relation cardinalities that have to be integrated: 1:1, 1:0, 1:n, and n:m. In a first step different criteria like area, position and shape are used to identify relations between the data-sets, which enables the detection of 1:1 and 1:0 relations. The usage of different algorithms to reveal these relations are the main-topic at this point of time in the project. Another problem which occurs comparing the data-sets is the partial representation of objects. For this case different methods of similarity calculations are tested to ensure a proper detection of “partial-digitised” objects. These methods calculate the similarity between the geometry of objects in two data-sets using points located on the shape of each objects. For every point of an object in one data-set a nearest-neighbour is searched in the amount of points placed on the shape of the other object. Evaluation of similarity is done by comparing the sum of points placed on both objects with the amount points for which a nearest neighbour has been found using the criteria of proximity. Analysing 1:n and n:m relations is done by grouping single objects in one data-set and compare these “test-groups” from one data-set with a single-object (1:n) or another “test-group” (n:m) from the other data-set. Comparation of two groups or a group and a single object is done by calculating a convex hull for the “test-groups”.

4. CHANGE DETECTION

Objects which have been selected through semantic and geometric integration and have been considered as a matching pair will be investigated for change detection. A simple intersection of corresponding objects is used for the change detection. Yet, the mentioned differences may cause even more problems which are visible as discrepancies in position, scale and shape. Therefore firstly, a local transformation will be applied, leading to a better geometric correspondence of the objects. To this end, the iterative closest point (ICP) algorithm from (Besl & McKay, 1992) has been implemented to achieve the best fitting between the objects from ATKIS and the geo-scientific elements using a rigid transformation. The result of this iterative procedure, is the best fit between the objects, and a link between corresponding objects in the different data set is established .

Evaluating the transformation parameters allows to classify and characterize the quality of the matching: in the best case, the scale parameter should be close to 1; also rotation and translation should not be too large – assuming, that the registration of the data sets is good. At this state of the implementation ATKIS is, because of the highest accuracy and the more frequent updates, taken as the reference, therefore the geometrie of the geoscientific data sets is matched to the reference geometrie without using any constraints and no intermediate result will be calculated.

4.1 Intersection and evaluation

Following these steps, intersection of objects for a proper change detection leads into a more promising result as simple intersection. The analysis and the classification of these segmented intersection result into different change situations, is a semantic problem and will be conducted in close collaboration with experts from geology and soil science, who are also partners in the project.

At this time of the project three different classes have been identified: the intersection segments can be classified according to the respective classifications of the area the segments are intersected from in the original data sets:

- **Type I** : Area is defined as water area in both maps, no adaptation required,
- **Type II** : Area in the geoscientific data set has been any type of soil, but is defined as water-area in the reference data set ATKIS; therefore the attribute of classification will be changed in the geoscientific map,
- **Type III** : Area in geoscientific data set has been water-area, but is now updated as soil-type. Therefore a new soil-classification is required.

While Type I and II require only geometric corrections and can be handled automatically, Type III needs more of the operators attention.

The resulting segments from the intersection process will be filtered using a predefined area-threshold. Segments undergoing these threshold – which has to be defined in close collaboration with the geoscientific partners – are taken as geometric discrepancy and will be closed using the attributes of the surrounding neighbours. Areas or segments which are larger than the given threshold are reported to the operator. There are different ways a water area can disappear, either through different natural (e.g. erosion) or man-made (e.g. refill) processes which have influence to the new soil type. This new soil type could not be derived automatically, but there are different proposals which could be offered to the user by the software. As a result a visualisation will be produced showing all the areas where an automatically evaluation of the soil situation could not be derived or only a proposal could be delivered and manual "field work" must be performed (Fig. 2).

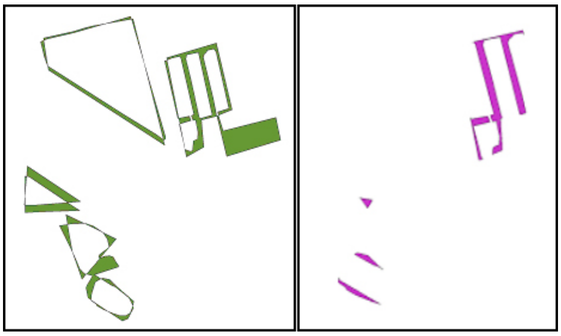

Fig. 2 : left: Results of simple intersection; right: Results of intersection with ICP matching and area-threshold filtering

The visualisation of Type III segments will reduce the amount of human resources needed to detect the topographic changes between the geoscientific data sets and ATKIS.

5. CONCLUSION

In this paper the ongoing research on semantic and geometric integration has been presented. The selection of the topographic element water, the automatic merging of the segmented objects and the use of the ICP-algorithm showed very good results.

In the near future the semantic catalogue will be expanded to cover all topographic elements which are represented in each of the three data sets, german digital topographic map (ATKIS) and the geoscientific maps from geology and soil-science. The introduction of punctual and linear elements will enhance the process of geometric integration.

Additional discussions with the external geo-scientific partners will ensure the creation of a fully-functional and automatic process.

5.1 Literature

Ad-hoc AG Boden, 1994. Bodenkundliche Kartieranleitung. Hannover, Germany, pp. 27-44.

Ad-hoc AG Geologie, 2002. Geologische Kartieranleitung – Allgemeine Grundlagen. Geologisches Jahrbuch G 9. Hannover, pp. 19 ff.

Besl, P. & McKay, N.,1992. A Method for Registration of 3-D Shapes, Trans. PAMI, Vol. 14(2), pp. 239-256.

Goesseln, G. v. & Sester, M., 2003. Semantic and geometric integration of geoscientific data sets with atkis – applied to geo-objects from geology and soil science. In: Proceedings of ISPRS Commission IV Joint Workshop, Stuttgart, Germany, September 8-9, 2003.

LGN, 2003. ATKIS in Niedersachsen und in Deutschland. In: Materialien zur Fortbildungsveranstaltung Nr. 1/2003, Hannover.

Sester, M., Butenuth, M., Goesseln, G. v., Heipke, C., Klopp, S., Lipeck, U., Mantel, D., 2003. New methods for semantic and geometric integration of geoscientific data sets with ATKIS – applied to geo-objects from geology and soil science. In: Geotechnologien Science Report, Part 2, Koordinierungsbüro Geotechnologien, Potsdam.

Walter, V. & Fritsch, D,. 1999. Matching Spatial Data sets: a Statistical Approach, International Journal of Geographical Information Science 13(5), 445–473.