AUTOMATIC INTERPRETATION OF DIGITAL MAPS FOR DATA REVISION

Karl-Heinrich Anders and **Dieter Fritsch**
Institute of Photogrammetry
Stuttgart University
P.O.B. 106037
70049 Stuttgart / Germany
kh.anders@ifp.uni-stuttgart.de

Commission IV, Working Group 3

KEY WORDS: Revision, Updating, GIS, Spatial Data Interpretation, Semantic Modelling, Object Oriented

ABSTRACT:

The amount of spatial data in digital form increases continuously. Governments and companies capture digital spatial data directly or convert existing analogous map data. The resulting large spatial databases have to be updated continuously; they often include a variety of implicit information which are in most cases not used. Because of the huge efforts in time and costs the automation of data revision is a topic of growing interest. The automatic interpretation of digital maps or digital landscape models (DLM) makes it possible to deduce information that is not explicitly stored in the data model. With this implicit information it is possible to support tasks like computer based revision of digital landscape models, automatic production of thematic maps and automatic data generalization. This article will concentrate on data revision based on aerial images, and the revision based on different digital maps. The used object-oriented data model will be described in detail. Very often a DLM is represented by several data models which might be differ slightly from each other. For this reason, the corresponding data models have to be converted into a uniform model. Furthermore, a set of operators for the object recognition in digital maps will be described and how they can be used for data revision.

1 INTRODUCTION

Nowadays there exists a huge amount of spatial data in digital form. All over the world governments and companies capture spatial data in digital form or convert existing analogous map data. In Germany, for example, the large scale database ALK (official digital cadastral map of Germany) and the medium scale database ATKIS (Authorative Topographic Cartographic Information System) have been initiated and realized. Besides these official basis information systems the car industry and related vendors of information technology offer digital map data for traffic management systems. All these information systems are based on different models of the landscape, i.e. they include only special parts of the landscape, have differences in the generalization and accuracy of the captured landscape phenomena. Because of the huge efforts in time and costs the automation of data revision is a topic of growing interest. It is important to note, that in the following we will refer to large digital databases and not deal with scanned raster data. Approaches on the interpretation of digital raster maps are given by [Illert 1990], [Illert, A. 1991], [Meng 1993] and [Carosio 1995].

The automatic interpretation of digital maps or digital landscape models allows to deduce information that is not explicitly stored in the data model. For example a human operator who looks at a part of the digital cadastral map (ALK) shown in figure 1 has no problems to recognize roads, crossroads or areas of residential buildings or industrial areas, although all these kinds of spatial objects are not explicitly stored in the ALK. On the contrary a computer system has access only to the information about the geometry of agriculture areas, ground plans of buildings and text symbols. Therefore, methods have to be found being able to deduce further information about spatial relationships of stored objects. On the other hand, these methods should also be able to create new objects from the existing ones.

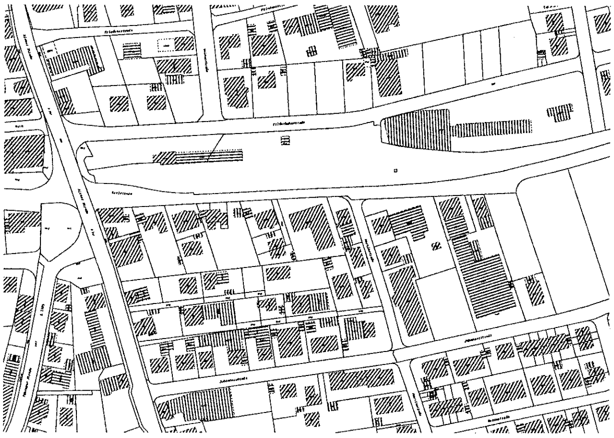


Figure 1: Example for the German ALK

With this implicit information it should be possible to support tasks like computer based revision of digital landscape models, automatic production of thematic maps and automatic generalization. This article describes, how implicit information can be used to support the data revision process based on digital aerial photos, and a revision based on different digital maps.

The following chapters will outline possible fields of interest, which can be supported by an automatic interpretation of DLM. Our approach designing a knowledge based shell to be linked with an object oriented database depends on two key issues

- which data model is used within the DLM ?
- which collection of simple objects with corresponding methods seems to be adequate ?

2 INTERPRETATION OF DIGITAL LANDSCAPE MODELS

The automatic interpretation of digital landscape models allows us to deduce information that is not explicitly stored in the data model itself. For the automatic production of thematic maps based on a given DLM (e.g. a map of industrial areas) a spatial and semantic analysis of the DLM objects is necessary. In the area of data revision the interpretation can be used to support :

- **a revision based on automatic image processing**
- **a revision based on different databases of similar scale**
- **a revision based on databases of different scale**

This topics will be described in the next sections.

2.1 Image Processing

Existing digital maps, which have to be updated, offer also prior information contents for data revision based on digital aerial photographs. The geometry and thematics of a digital map can contribute much to solve the image interpretation problem, for instance to extract object attributes like shape, texture, size etc. from the imagery or to predict regions of interest for which alterations are to be expected, e.g. new housing estate. The link of the existing map and the aerial imagery is bidirectional which means, the symbolic scene description of a map is imported using an E/R model in the image space. Then questions of image interpretation can be answered, for instance, what type of features in the image is to be expected (shape, size, texture etc.). On the contrary extracted image features may represent update map information, e.g. the boundary line of a street, the ring polygon of a house etc. This knowledge can be used to control further steps of processing.

This kind of using spatial data as prior information requires matching techniques between image objects and digital map objects. In particular the mapping of image objects, which do not exist in the digital map, allows to deduce implicit information like geometric or neighbourhood relations. With this implicit information we are able to set up hypotheses in the sense to interpret the image features in terms of real world objects. Therefore, we need operations which match image features (lines, pixel regions, etc.) with existing vectorial map data.

In [Haala & Anders 1996] the large scale database ALK is used to predict 3D building models for a 3D building reconstruction in digital aerial photos. Figure 2 shows constructed building hypotheses based on the shape and spatial relationships of the ground plans included in the ALK. Another approach for updating the ATKIS DLM 200 database using satellite images is described in [Klaus-Jürgen Schilling, Thomas Vögtle, Peter Müßig 1994].

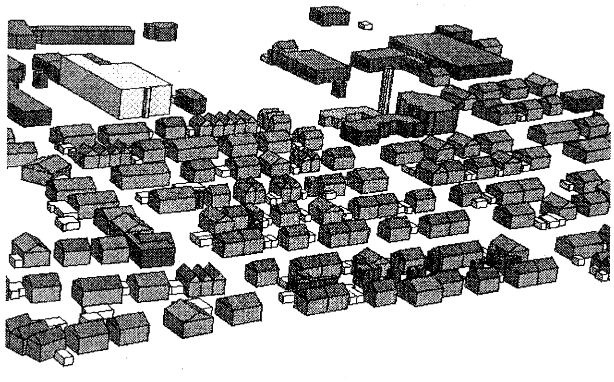


Figure 2: 3D building hypotheses based on the ALK

2.2 Databases of Similar Scale

If one class of spatial objects in a given DLM should be revised using another DLM, which contains the same kind of spatial objects in a more or less equal scale, one has to match both digital maps. The matching of two given vector databases is also called *Conflation*, which comes from the Latin *con flare* meaning "blow together" [Maureen Lynch and Alan Saalfeld 1985]. After the matching process is carried out, it should be possible to detect differences between the two databases. Two basic problems occur within the matching of two digital maps

- **differences in accuracy** of data capture (figure 3) and
- **differences in data modelling** of a spatial object (figure 4)

which can be overcome the better the more implicit information can be used. In [Walter & Fritsch 1995] an approach based on relational matching for the matching of ATKIS road-objects with GDF road-objects is described.

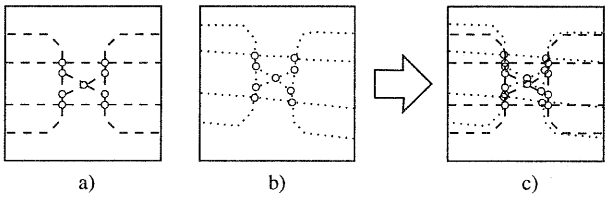
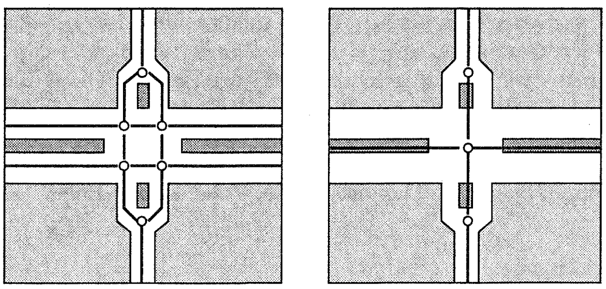


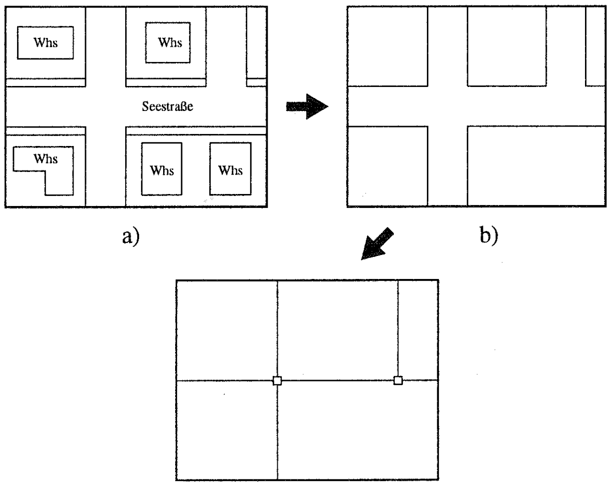
Figure 3: a) dataset A, b) dataset B of the same area, c) differences in the accuracy of data (after [Walter & Fritsch 1995])



a) b)

Figure 4: Example for the different modelling of the same crossroads (after [Walter & Fritsch 1995])

2.3 Databases of Different Scale



a) b) c)

Figure 5: a) ALK data, b) extracted roadsides, b) generalized ATKIS road-objects

Besides the data revision based on digital aerial photographs it is also possible to use an existing digital map. This dataset can include more information (larger scale) or information which is captured with a higher accuracy than the digital map which has to be revised. The official digital cadastral map (ALK) of Germany contains spatial information of large scale (1:250 - 1:2500) in several layers, but is unfortunately not based on objects. Therefore the usage of modern object oriented techniques requires data conversion to built up object classes and object hierarchies for further processing. On the contrary the medium scale geographic information system ATKIS is based on classes of several landscape objects and includes very similar information contents as the ALK. The large scale dataset may serve as data source to deduce medium scale information, e.g. roads and further linear elements.

One aim of our work is to deduce ATKIS road objects from the ALK because in both DLM's exists the spatial object road, although in the ALK roads are captured with a higher accuracy. The effort in time and costs of data capture for the information system ATKIS could be reduced. A two step approach is applied which consists of object recognition (figure 5b) and object generalization (figure 5c). The object recognition is carried out using well-known methods of image processing to extract the geometry of the roads. Among those especially operators for the *similarity, continuation, unity, symmetry, closeness and parallelism* are used. The object generalization is necessary because the ALK describes the geometry of roads by areas and ATKIS represents the geometry by the central axes of the lanes. Therefore we will use a vector geometry based algorithm described in [Olson 1995]. Additionally it is necessary to build correct ATKIS road objects using the central axes.

3 SYSTEM DESIGN

The aim of our work is to develop a shell onto a given object oriented spatial database to deduce implicit information we are looking for. Our system setup (figure 6) is based on the object oriented database Objectivity/DB from Objectivity Inc., because an object oriented data model has some advantages described in the next chapter. A general interpretation process needs a type of knowledge representation to describe the semantic models which will be used by the interpretation process to control the usage of object methods (described in chapter 5) and the symbolic data processing. In principal the knowledge representation can be done by *rule base systems, blackboards, semantic networks or frame based systems*, a detailed description of the different types of knowledge representation can be found in [Reimer 1992]. We will use a frame based system because of the direct link between frames and object oriented data models. Frames are an object oriented type of knowledge representation and well suited for the object oriented data model.

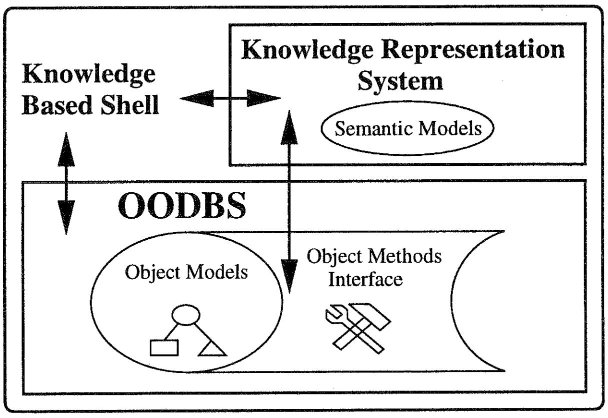


Figure 6: Scheme of the system design

4 DATA MODEL

Most of the existing spatial information systems are based on different models of the landscape because they include only special parts of the landscape (e.g. traffic information systems) and have differences in the generalization or accuracy of the captured landscape objects (differences in the scale of the DLM's). Also, there exists a difference between the data representation of the different information systems. The data can be made available in a graphic format (e.g. DXF), in a layer format (e.g. ALK) or in an object oriented format (e.g. SAIF). Therefore, we have to convert the corresponding data models into a uniform model. We have chosen an object oriented data model to represent spatial objects because of the natural representation of real world objects in contrast to their representation in relational databases [Riekert 1993, L.Raynal, B.David & G.Schorter 1995, Fritsch & Anders 1996].

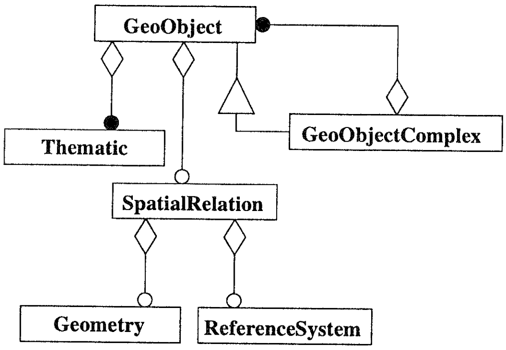


Figure 7: Description of the spatial object class

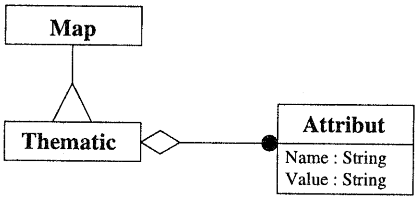


Figure 8: Thematic class

Our object oriented data model used is derived from the data model of the Canadian standard *SAIF* (Spatial Archive and Interchange Format) described in [*Spatial Archive and Interchange Format: Formal Definition Release 3.2* 1995]. In the following figures we use the *OMT-Notation* (Object Modeling Technique) described in detail in [Rumbaugh, Blaha, Premerlani, Eddy & Lorensen 1991]. In the OMT-Notation the rhomb symbol represents the object oriented feature *aggregation* (has_a relation) and the triangle symbol represents the feature *inheritance* (is_a relation). Our developed data model uses *container classes* provided by the used object oriented database system. Geographic objects are modelled by the class *GeoObject* (figure 7). Parts of the class *GeoObject* are the classes *Thematic* and *SpatialRelation*. The class *Thematic* (figure 8) is implemented as a associative map and represents the non-spatial attributes of a geographic object. The class *SpatialRelation* is used to represent the spatial attributes and describes the kind of geometry and reference system which belongs to the geographic object. The class *Geometry* shown in figure 9 provides the basic geometric objects described by the classes *Point, Line, Area* and *GeometryComplex*. The class *GeometryComplex* is used to build geometries, which are composed of the classes *Point, Line* and *Area*. Every object from the class *Line* is defined by two objects of the class *Point* (start point and end point of the line) and any number of additional points depending on the type of line interpolation (e.g. polygon, spline). An object of the class *Area* is described by at least one border line of the area (areas with holes includes two or more border lines), which is modelled by the class *Line*.

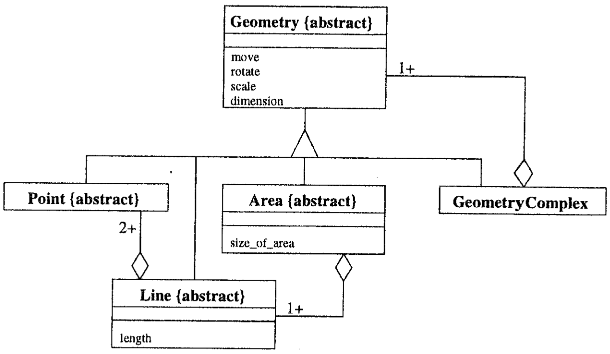


Figure 9: Geometry class (point, line, area)

5 METHODS

We can distinguish the following methods which every spatial object in the database has to provide :

- **relational operators** (figure 10) for
  - ▷ **semantic relations** (e.g. *type_of, is_a*)
    The interpretation process has to be able determining the kind of object it is looking at. For example, with this method we are able to determine the semantics of a spatial object (e.g. road or building).
  - ▷ **structural relations** (e.g. *part_of, has_a*)
    In order to get information about the composition of an object it provides methods to get all objects, which are a part of this spatial object, or all objects to which the spatial object belongs.
  - ▷ **topological relations** (e.g. *inside, disjoint, ...*)
    Basically we need information about the topological relations of a spatial object because we have implemented methods to test all topological relations as described in [Egenhofer & Herring 1990].
  - ▷ **neighbourhood relations** (e.g. *lies_beside, near_to*)
    The methods for the neighbourhood relations are not implemented at the moment. They will be used to deduce relations which cannot be determined only by topological relations. For example, the database query *Retrieve all buildings which lie beside road A*. This can only be solved using geometrical relations. Another problem is, that there exists no mathematical definition of this kind of relation.

  - ▷ **geometrical relations** (e.g. *is_parallel_to*)
    In our data model it is possible to determine whether two spatial objects are parallel (vertical) or not. Those objects with an area as geometry the diameter [Preparata & Shamos 1988] of the area is determined and used for the calculation. Also we can calculate the distance between objects.

- **operators for aggregation of spatial objects** (e.g. *convex hull, closest object pair, ..*)
  To create new generalized or grouped objects we have implemented a convex hull algorithm [Preparata & Shamos 1988] for different kinds of geometry classes (Point, Line, Area). Also we have implemented methods to determine the nearest objects to a given object. This problem can be solved by the *Voronoi Diagram* [Preparata & Shamos 1988] in the case of points. For lines and areas we use a centroid point.

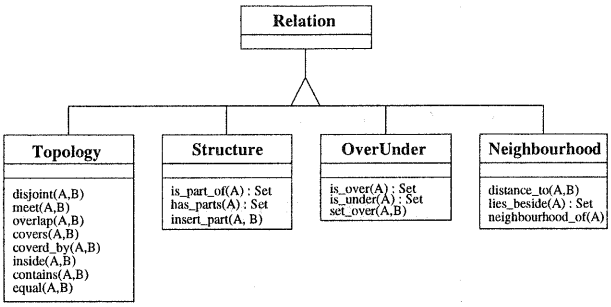


Figure 10: Class of object relations

6 CONCLUSION

Because of the increasing amount of spatial information in digital form and the importance of data actuality and data quality for economy, industry and commerce it will become more and more important to automize the revision of digital landscape models. Within this article we described an approach for data mining in spatial databases to deduce implicit given information. Also a knowledge based shell for an object oriented spatial database was designed. We described some examples of possible applications in the area of automatic data revision. The automatic interpretation of digital landscape models is a new field of interest for GIS research and far from being solved. For the time being we are only able to detect low level object information like geometrical or topological relations (based on the given object geometry) and we have some aggregation operators to create new objects developed. In the future work the frame based knowledge representation has to be developed to control the interpretation process and to deduce high level object information.

References

Carosio, A. [1995], Three-dimensional synthetic landscapes: Data acquisition, modelling and visualization, *in* D. Fritsch & D. Hobbie, eds, 'Photogrammetric Week '95', Herbert Wichmann Verlag, pp. 293–302.

Egenhofer, M. J. & Herring, J. R. [1990], A Mathimatical Framework for the Definition of Topological Relationships, *in* '4th Int. Symp. on Spatial Data Handling', Vol. 2, pp. 803–813.

Fritsch, D. & Anders, K.-H. [1996], 'Object Oriented Concepts in GIS', *GIS: GEO-INFORMATION-SYTEMS* **9**(2), 2–14.

Haala, N. & Anders, K.-H. [1996], 'Fusion of 2D-GIS and Image Data for 3D Building Reconstruction', Will be published in proc. of 18. ISPRS conference.

Illert, A. [1990], Automatische Erfassung von Kartenschrift, Symbolen und Grundrißobjekten aus der Deutschen Grundkarte 1:5000, Dissertation, Universität Hannover.

Illert, A. [1991], Automatic Digitization of Large Scale Maps, *in* 'Technical Papers ACSM-ASPRS Annual Convention', Vol. 6, Baltimore USA.

Klaus-Jürgen Schilling, Thomas Vögtle, Peter Müßig [1994], Knowledge Based Analysis of Satellite Images, *in* 'Symposium on Spatial Information from Digital Photogrammetry and Computer Vision', ISPRS Commission III, Munich, Germany, pp. 732–736.

L.Raynal, B.David & G.Schorter [1995], Building an OOGIS Prototype : Experiments with GEO$_2$, *in* 'ACSM/ASPRS Ann.Con., Charlotte, North Carolina, Tech.Papers, Auto-Carto 12', Vol. 4, pp. 137–146.

Maureen Lynch and Alan Saalfeld [1985], Conflation: Automated Map Compilation – A Video Game Approach, *in* 'American Society of Photogrammetry, editor, *Auto-Carto 7*.

Meng, L. [1993], Application of neural network in cartographic pattern recognition, *in* 'Proc. of 16. Int. Cartographic Conference', Vol. 1, Köln, pp. 192–202.

Olson, N. E. [1995], An Algorithm for Generating Road Center-Lines from Road Rights-of-Way, *in* 'ACSM/ASPRS Ann.Con., Charlotte, North Carolina, Tech.Papers, Auto-Carto 12', Vol. 4, pp. 11–21.

Preparata, F. P. & Shamos, M. I. [1988], *Computational Geometry*, Springer-Verlag, New York.

Reimer, U. [1992], *Einführung in die Wissensrepräsentation*, B.G. Teubner, Stuttgart.

Riekert, W.-F. [1993], Das Forschungsprojekt GODOT: Geodatenhaltung mit objektorientierten Techniken, *in* '3. Internationales Anwenderforum Geoinformationssysteme', Duisburg.

Rumbaugh, J., Blaha, M., Premerlani, W., Eddy, F. & Lorensen, W. [1991], *Object-Oriented Modeling and Design*, Prentice Hall, Inc.

Spatial Archive and Interchange Format: Formal Definition Release 3.2 [1995], Surveys and Resource Mapping Branch Ministry of Environment, Lands and Parks, Province of British Columbia, Canada.

Walter, V. & Fritsch, D. [1995], Matching Techniques for Road Network Data in different Data Models, *in* J. Soliman & D. Roller, eds, '28th International Symposium on Automotive Technology and Automation', Automotive Automation Limited, Croydon, England, pp. 633–640.