'Photogrammetric Week '97' D. Fritsch & D. Hobbie, Eds., Wichmann Verlag, Heidelberg, 1997.

Automated interpretation of digital landscape models

KARL-HEINRICH ANDERS, Stuttgart

ABSTRACT

The purpose of this paper is to provide an overall view of the methods of *spatial data mining* and its applications to digital landscape models. Spatial data mining can be defined as the deduction of information that is not explicitly stored in a given spatial data model. The subject spatial data mining represents the integration of several fields, including machine learning, database systems, data visualization, statistics, information theory and computational geometry. The automation of spatial analysis functions has two main aspects. On the one hand it deals with the automation of spatial operators conventionally used in a GIS-program for complex analysis applications, e.g. site planning. Such an application usually involves a sequence of operations, e.g. classification, buffering, selection, etc. This process is controlled by the human operator according to a "model" he has in mind. Automation of such a process requires to make explicit this model and apply it to the data. In this way, e.g. a model for site planning can be created. The second aspect concerns data mining. This approach is used to find connections in the data which are not known in advance - therefore no model exists - which are however implicit in the data.

1. INTRODUCTION

Nowadays there exist 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. The availability of these quantities of digital spatial data naturally leads to the questions of how these amount of information can be handled in an effective way and how different data sets can be integrated and maintained in a unifying way. Usually different data bases are established for different purposes: as an example, national mapping agencies derive topographic information, urban planning institutions capture data for city planning, traffic management companies demand detailed road information for the autonomous navigation of cars - all of them are operating on the same physical reality, namely the earth's surface, and sometimes even on the same type of objects. The problem is that usually the underlying application drives both the data model and the level of detail. Thus although a huge amount of spatial data is already available - the question remains whether the data can also be used beyond their original purpose. This question is closely linked to the problem of derivation of information, which is only implicitly given in the data set.
Imagine a data set representing topographic objects as visualized in Figure 1. The data - taken from the German ALK[1] - describe legal aspects. Although in the data set primarily only lines and text symbols are stored, we as humans can easily read more information from it: we can identify individual parcels (e.g. Figure 2) and road crossings, distinguish city area from rural region, detect the neighbourhood between roads and built-up areas, ... Thus a human viewer can not only see the explicit information, but also the implicit one. The paper deals with the problem to automate this functionality of exploiting also implicit information in a spatial data base and making it explicit for different application areas.
The paper is organized as follows: First an overview of possible applications for the derivation of implicit information from spatial data sets is given. After that the technological background for this is described. Section 4 describes one application in detail. Finally a discussion and an overview of possible extensions of the approach is given.

[1] Automatische Liegenschaftskarte: Digital Cadastral Map

'Photogrammetric Week '97' D. Fritsch & D. Hobbie, Eds., Wichmann Verlag, Heidelberg, 1997.

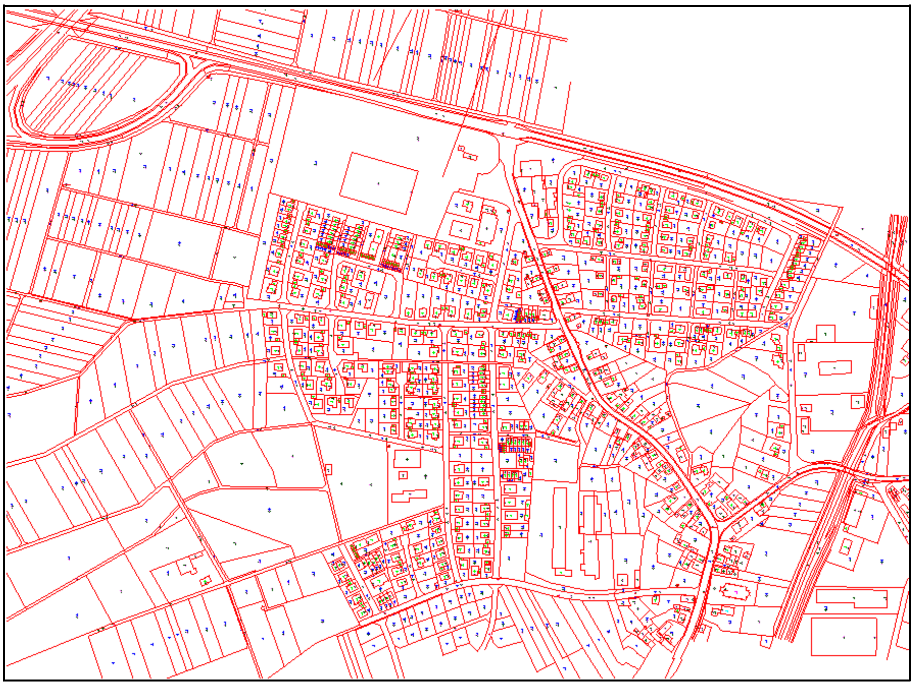

Figure 1: Digital cadastral map.

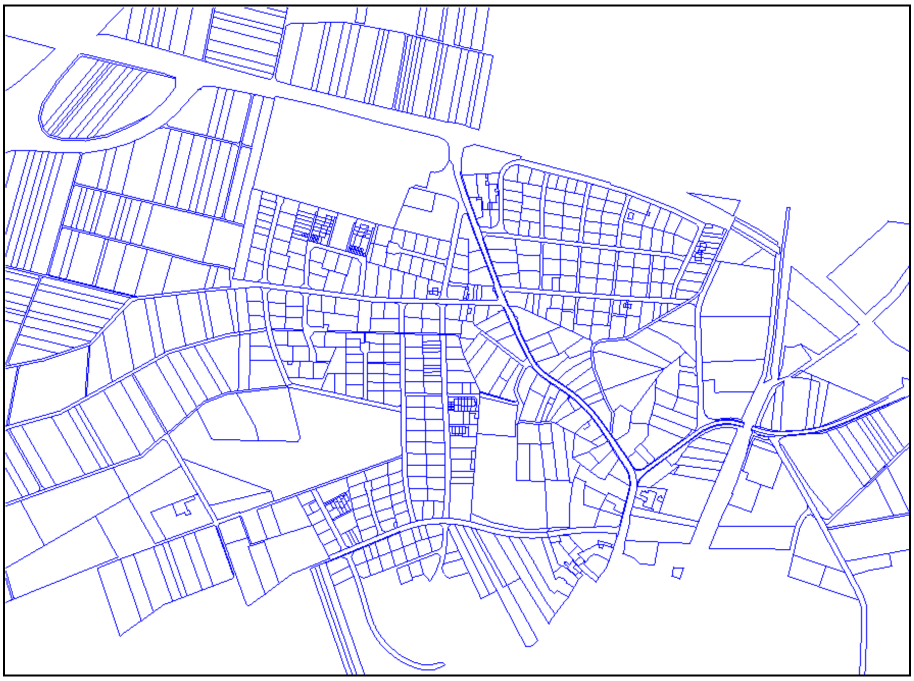

Figure 2: Extracted parcels.

2. MOTIVATION AND POSSIBLE APPLICATIONS

Looking for implicit information in data bases is a discipline called Data mining - up to now mainly applied to descriptive data sets. An example is the prediction of the audience of BBC-customers based on audience figures like type of program, time of showing, nature of competing programs (ISL, 1996). The application of this technique to spatial data bases has been addressed only recently. Here however there is a wide field of possible applications of this technique to spatial databases.

'Photogrammetric Week '97' D. Fritsch & D. Hobbie, Eds., Wichmann Verlag, Heidelberg, 1997.

Models for Image Interpretation:

Automatic image interpretation heavily relies on models to guide the interpretation process. These models are usually designed by hand and then applied to the image data. On the other hand huge collections of object models are already available in terms of objects stored in digital databases like CAD-systems or GIS. This information however cannot be used directly, because usually not all the necessary information for image interpretation is stored explicitly. For example (Plietker, 1994) uses ATKIS-data[2] to find roads in aerial orthophotos. In order to transfer the ATKIS-model to the image interpretation process the data has to be enriched by general knowledge about it, e.g. the fact that roads are captured with their middle axis and with a predefined width. (Haala & Anders, 1996) give another example of the use of individual object instances from a GIS for image interpretation: the extraction of buildings from stereo data using 2D-building groundplans. As only 2D information is available, hypotheses about the 3D-form and shape of the buildings have to be inferred from the data set (Figure 3). Data interpretation can also be used to derive an intensional, generic description from the extensional set of individual object instances in the data base. Thus a generic object model can be achieved. This general model comprises not only object attributes but also relations between objects. The latter being very important since they specify the context (Sester, 1995).

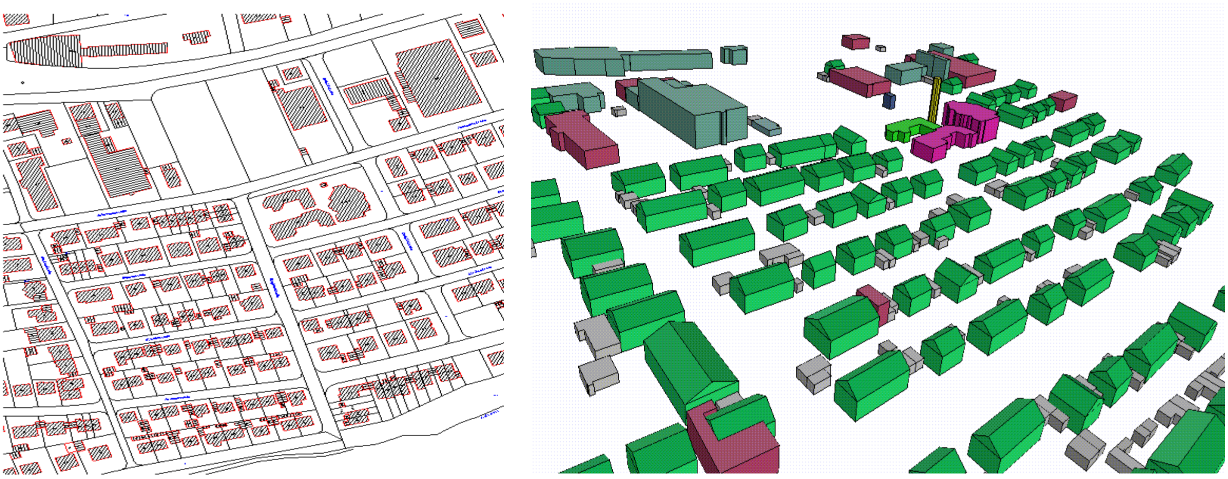

Figure 3: Generated 3D-building hypothese (right) based on the 2D digital cadastral map (left).

Multiscale Representation / Model Generalization:

Each data acquisition process starts with modeling the data to be captured. The model is determined by the application, and thus describes that part of the reality, which is "seen" from the system. In order to make use of data also for other applications than they were originally acquired for, a transformation of the data has to be performed. This presumes to have rules for the transition between different data models. The more different the models are, the more complex are the relations between the data sets and thus is the derivation of the rules or transformations. This is even more true for the case when the data stems from different resolutions of detail (in terms of geometry, semantics and time). E.g. ATKIS specifies built-up areas - in order to derive them from ALK (which describes legal boundaries graphically - including buildings) a relation of scales 1:1 500 with 1:25 000 has to be found. This can be achieved using grouping operations, which base on spatial neighbourhood. See section 4 for an implementation of such an aggregation method.

(van Oosterom, 1995) describes a so-called reactive data structure - a data structure representing levels of detail. His GAP (generalized area partitioning)-tree is successfully applied to perform a

[2] Authoritative Topographic Cartographic Information System

'Photogrammetric Week '97' D. Fritsch & D. Hobbie, Eds., Wichmann Verlag, Heidelberg, 1997.

generalization on-the-fly - primarily for visualization purposes. His aggregation rule is defined a priori and bases exclusively on object adjacency.
The approach given by (Molenaar, 1996) specifies a semantic model of the aggregation hierarchies in advance. The aggregation itself then is based on spatial adjaceny.

Use of data sets for different applications / map matching:
When the scales of the different applications are similar, map-matching techniques can be applied. This process allows for a transformation of one data set into the other. Usually matching is based on geometric similarity (e.g. position, form, ...); but also semantic similarity is of great use. An example for such a matching is the link between GDF[3] road data and ATKIS roads. Although the data sets reflect the same physical reality (namely roads) and are based on the same resolution (1:25 000), due to slightly different underlying data models different object instances are acquired. The approach by (Walter & Fritsch, 1995) uses relational matching to find the correspondences between the data sets. Another application is the integration of application dependent information into existing data sets, e.g. the integration of environmental information into an official topographic map.

Derivation of Meta-Data:
In order to allow for a more flexible, open and independent use of digital data, so-called meta-information is of great importance. Meta-data specifies general knowledge about the data (e.g. object types, extremal values, range of values, ...). The problem with meta-data is that it is usually not stored explicitly with the data, but mostly is only known by its intelligent user. This information can be derived from actual object instances of the data set, by spatial interpretation techniques.

Extraction of High Level Structure:
City planners need information about typical structures in a city. In order to derive this knowledge, an analysis and statistical evaluation of the facts and the relations between individual objects has to be performed. This process can be automated based on a given dataset by spatial data interpretation. This presumes to have a model of what is to be found. An example is a data set consisting of roads; if road crossings have to be found, a model for a crossing has to be available. This interpretation task can directly be compared to image interpretation, which also needs models in order to find objects in an array of gray values.
Potentially also relations are interesting, which are not known or clear in advance, and kind-of hidden in the data. Therefore techniques are necessary to automatically derive dependencies, relations and statistical information from the data. For example a spatial and statistical analysis of data in a city can reveal dependencies of certain buiding types and special regions, etc. Another example are the requirements for regional planning purposes, where information about the population density is very important. This should be derived relating to special regions, along roads, ... (Koperski & Han, 1995) give an example of spatial data mining where they analyse the adjacency of canadian cities and other topographic objects. The result is that huge cities tend to be close to the sea.

3. SPATIAL DATA MINING

Data mining or *knowledge discovery in databases* (Piatetsky-Shapiro & Frawly, 1991), (Holsheimer & Siebes, 1994), (Fayyad et al., 1996) can be defined as the discovery of interesting, implicit, and

[3] Geographic Data File: standard for road data in Europe

'Photogrammetric Week '97' D. Fritsch & D. Hobbie, Eds., Wichmann Verlag, Heidelberg, 1997.

previously unknown knowledge from large databases (Frawley et al., 1991). The subject spatial data mining is the extension of data mining from relational and transactional databases to spatial databases. Nowadays huge amounts of spatial data have been captured in various applications, ranging from remote sensing, to geographic information systems, environmental and planning. The human ability to analyse this large spatial databases manually is far exceeded. That makes it necessary to automate the information (knowledge) discovery to support a human operator.
The subject spatial data mining represents the integration of several fields, including machine learning (Michalski et al., 1984), database systems, data visualization, statistics, information theory and computational geometry. Spatial data mining techniques have wide applications in geographic information systems and remote sensing. These methods can be used for understanding spatial data, discovering relationships between spatial and nonspatial data, construction of spatial knowledge-bases, query optimization, characterization of spatial data, etc. There are different forms of knowledge discovery in spatial databases, like:

- **Spatial object characteristics**
  Description of spatial objects at a high level of abstraction. To find such types of description *concept hierarchies* and *generalization* operators are needed.
- **Spatial object clusters**
  Classification of spatial objects based on their *spatial neigbourhood* which has to be defined in a suitable way. A possible information in a building database can be clusters of buildings including mainly industrial buildings.
- **Spatial association rules**
  Knowledge about the correlation between spatial objects based on topologic (*adjacency, intersection, ...*) and geometric (*distance between objects, size of objects, ...*) relations.
- **Spatial patterns**
  A spatial pattern describes the characteristic structure of the distribution of the spatial objects. For example a group of buildings can have a *linear* structure or the structure of a *cluster*. Spatial patterns can be used for a spatial trend analysis of spatial objects.

which all needs their own spatial data mining method.

3.1 Spatial data mining methods

Spatial data mining methods can be divided into five classes:

Statistical methods:
Statistical analysis of spatial data is the most common approach for analyzing spatial data ((Fotheringham & Rogerson, 1994), (Shaw, 1994)). It handles very well numerical data and usually comes up with realistic models of spatial phenomena. The major disadvantage of the statistical approach is the assumption of statistical independence among the spatially distributed data. This causes problems as many spatial data are in fact interrelated, i.e., spatial objects are influenced by their neighbour objects (Koperski et al., 1996).

Knowledge generalization methods:
Generalization-based methods are used to find characteristic and discriminant rules in a database. The method of generalization-based knowledge discovery requires the existence of background knowledge which is represented in the form of concept hierarchies. A widely used technique for discovering generalized knowledge is the machine learning technique *learning from examples*

combined with methods for *knowledge generalization*(e.g., decision-tree induction). An extension and modification of these techniques to spatial databases is described in (Lu et al., 1993).

Cluster analysis:

Cluster analysis is a branch of statistics that has been studied for many years. The goal of this method is to cluster objects into classes, based on their features, which maximize the similarity of class objects and minimize the similarity of objects from different classes. There are *probability-based* and *distance-based* clustering methods. Cluster analysis represents a type of *unsupervised learning*. The advantage of this method is that interesting structures or clusters can be found directly from the data without any background knowledge. Modified clustering techniques are described by (Ng & Han, 1994), (Ester et al., 1995) and (Ester et al., 1996).

Search for spatial associations rules:

These methods look for characteristic rules that link one or more spatial objects to other spatial objects. Spatial association rules are of the form $X \rightarrow Y\ (s\%, c\%)$, where X and Y are spatial or nonspatial predicates, $s\%$ the support and $c\%$ the confidence of the association rule ((Koperski & Han, 1995), (Bollinger, 1996)). There are various kinds of spatial predicates that could constitute a spatial association rule, e.g. neighbourhood (distance) information, topologic relations (*intersection*, *overlap*, etc.) or spatial orientation (*right_of*, *north_of*, etc.).

Aggregation-, approximation-methods :

These methods use spatial relations (topologic or geometric) to create new complex objects (Knorr & Ng, 1995) or to find patterns, structures (Regnauld, 1996) in the spatial database. The geometry of complex objects is described by an appropriate approximation of the aggregated geometries (e.g. convex hull). This type of spatial data mining methods often use concepts of computational geometry (Preparata & Shamos, 1985). An example for that kind of information is a created rule as the following: *80% of all houses in that cluster have a saddle roof.*

3.2 Spatial neighbourhood

A basic relation in spatial data mining is the *spatial neighbourhood relation* between spatial objects. This relation is important for spatial aggregation and spatial approximation operators. But there is no common definition for the subject spatial neighbourhood because the definition depends on the application. Generally one can define three types of spatial neighbourhood:

topologic neighbourhood

The neighbourhood is defined through the topology between the geometry of spatial objects. In the most cases the *adjacency* relation is used for that, but also relations *intersection* or *is_inside* are possible.

metric neighbourhood

The neighbourhood is defined through the distance between spatial objects. In this case there are three possibilities:

- **Distance buffer**
  All objects with a distance to a given object smaller than a given value are neighbours.
- **Nearest neighbour**
  Only the object with the smallest distance to a given object is called a neighbour. In this case no value for the distance is necessary and only one neighbour object is possible.

- **Delaunay triangulation**
  A given value for the distance is also not necessary if the neighbourhood is defined through the *Delaunay Triangulation* (Preparata & Shamos, 1985) of all points which describes the geometry of the spatial objects. In this case more than one neighbour object is possible.

semantic neighbourhood

The neighbourhood of spatial objects is defined through their semantic and not through their spatial attributes. Two objects are neighbours when their semantic is similar. For example a tree and a hedge can be neighbours because both belongs to the semantic vegetation. But the used definition depends on the application like the definition of spatial neighbourhood in general.

4. EXAMPLE: ALK/ATKIS - FROM PARCELS TO SETTLEMENT AREAS

4.1 Basic assumptions and available data

In Germany the ATKIS data base is established in three levels of detail, namely 1:25 000, 1:200 000, 1:1000 000. This data set comprises topographic objects of seven different classes (e.g. settlement, traffic, vegetation, ...). On the other hand, cadastral information in scales 1:1500 ...1:2000 is acquired for the ALK. In ALK legal boundaries, as well as useage or type of objects are specified. In order to facilitate and automate the time consuming update process, the idea arises to derive ATKIS data from ALK. In this way only the more detailed data set (ALK) would have to be updated manually - whilst the other can be deduced from it. Currently the update is done separately for each data set; ATKIS is acquired (in Baden-Württemberg) based on a manual interpretation of orthophotos, while ALK stems from direct geodetic measurement. Using ALK for ATKIS update involves generalization or aggregation operations, respectively. First of all the links between the data sets have to be established. In the following, an approach will be demonstrated showing the derivation of built-up areas in ATKIS from ALK data. A so-called object-catalogue of ATKIS specifies, how the objects are to be created and captured. The following "rule" describes how built-up areas (object id 2111) have to be acquired (*ATKIS-Objektartenkatalog*, n.d.).

> “Built-up areas are areas that are exclusively or mainly used for living. Besides buildings, there are also objects serving for the provision of needs of the whole area like small workshops, organizations of clerical, cultural, social, or sanitary purpose.
> The boundaries between a built-up area and neighbouring areas is given by the boundaries of the parcels of land the buildings are standing on.
> Only areas greater than 10 hectars have to be captured; if the area is smaller, it has to be acquired only in special cases.
> The geometric object type is a point (for areas less than 10 hectars) or an area for those greater than 10 hectars."

The first step is to implement such a rule. The specification above reveals a link between the data sets, namely the fact that built-up areas consist of adjoining residential areas. Residential areas are parcels of land comprising building objects like houses, garages, ... This however is information which is stored (explicitly and implicitly) in ALK. As the rule implies to aggregate neighbouring building-objects to a settlement area, the question is how to define neighbourhood in this application. First of all, in order to come from ALK to ATKIS, the underlying data models have to be examined in detail in order to make them explicit and thus usable. This will be revealed in the following semantic modeling process.

4.2 Semantic modeling, data structures, and implementation

The digital cadastral map (ALK) is provided by the landsurvey administration. Possible formats the user can get the data are SICAD/open data exchange format (SQD) or DXF. In these formats these data sets contain only geometric features and text elements. These elements are points, lines, areas and text features. SICAD/open additionally provides the feature parcel number. Thus the data is a mere graphical representation (cf. Figure 1). But in high level functions in GIS we need more complex objects with a specific meaning. For an human operator it is no problem to extract spatial objects from this data set when it is displayed on the computer screen.
In order to describe how spatial objects are represented in a digital cadastral map, we use *semantic modeling* as a conceptual method to analyse how a human operator detects such kinds of objects. Figure 4 shows the semantic model which we use to describe parcels of land and residential areas. To represent the spatial objects in the computer we use an object oriented data model presented in (Anders & Fritsch, 1996), (Fritsch & Anders, 1996). *Spatial objects* are aggregates of *thematic objects* which describe the semantics of an object, and *geometric objects* which represent the spatial position and form of an object. This structure directly allows to compute the topological relation *adjacent_to* among the features, because the incidence relations for the topological objects point, line and area are stored explicitly in our data model. As every line object ``knows" to which area it belongs, the adjacency between areas can be computed. The whole implementation bases on an object oriented database in conjunction with an object oriented programming language. The advantage is that we are able to represent semantic relations like *is_a* and *part_of* directly in our data model.

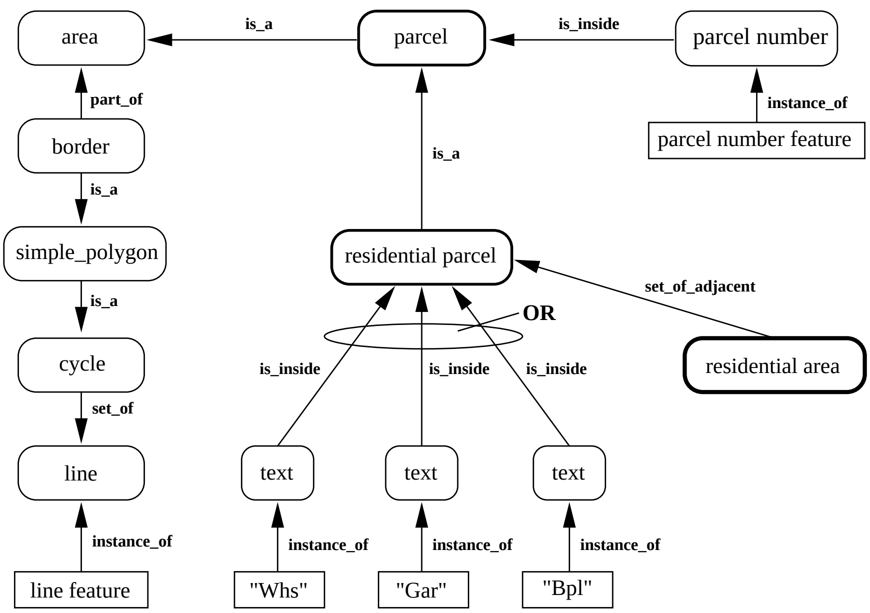

Figure 4: Semantic model of parcels and residential areas.

The first problem is that in the given representation of the digital cadastral map parcels are not stored as area features. Only the borderlines of the parcels are stored as line features. The buildings are directly stored as area objects - but this is only true for most of the cases. Therefore the first step in our approach is to compute all possible areas based on the given list of line features. An area is represented through a *simple polygon* which is a cycle of line features in the net of lines.
In the next step the relation between geometric objects parcels and the corresponding parcel number features have to be established. This can be achieved by geometric calculations of a point-in-polygon test between all possible combinations of parcels and numbers. This results in the following figure 2.

'Photogrammetric Week '97' D. Fritsch & D. Hobbie, Eds., Wichmann Verlag, Heidelberg, 1997.

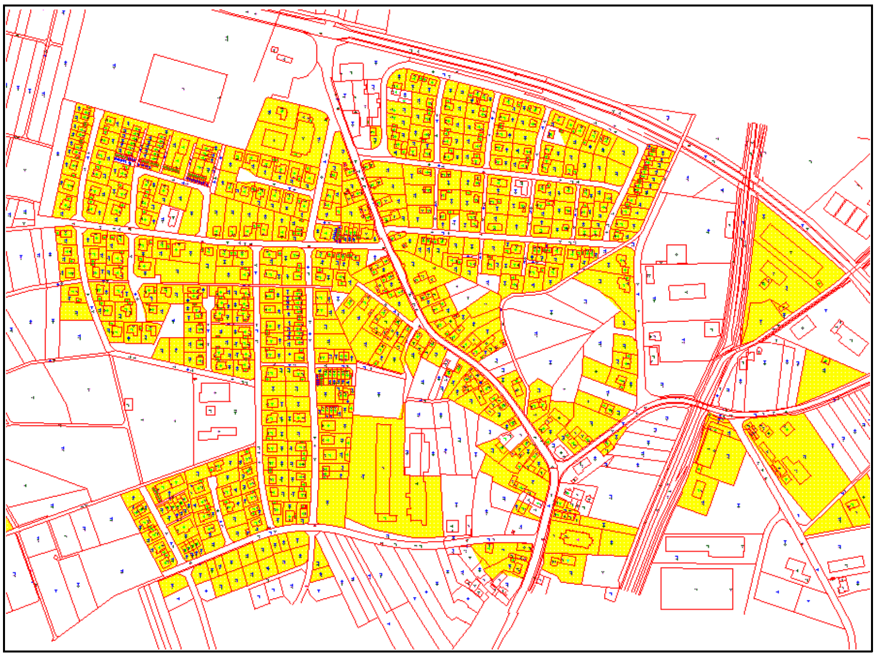

Figure 5: Built-up areas generated from ALK.

Then we have to choose all that parcels which probably have a residential use. Therefore we take all text features which are instances of *'Whs'*, *'Gar'*, *'Bpl'* (building, garage, building lot) and again make a point-in-polygon test for all possible combinations to find the relations to the corresponding areas. In the last step we have to group all computed parcels with residential use to create *residential area* objects. The semantic rule for grouping of these areas is the adjacency principle (see right hand side of Figure 4). I.e. all residential areas which are connected via a common border are merged. This can easily be performed using the topological relations between the areas which is directly extractable from our data model.

Figure 5 gives the result of the grouping and shows the built-up areas in that region. This result was achieved only based on the assumption of a spatial aggregation of adjoining residential objects. Comparing these ATKIS-objects with original ATKIS-data reveals the following differences (see Figure 6):

- The main source of the difference are: possibly different acquisition dates, different acquisition rules, ...
- The boundaries of the areas do not exactly correspond: this is due to the fact that in ATKIS the roads are stored as line objects only by their middle axis - the adjacent objects then meet at that axis; in our case as we take the road model from ALK (which is area based), the built-up areas consequently border the ALK roads. An extension would be to also generalize the roads (find middle axis from road areas) and then generalize both data types in common.
- In the current approach only a subset of all possible building objects have been modeled. Thus industrial areas or areas with mixed usage (like industrial buildings and residential buildings in the same parcel) are included in our result - in ATKIS however they should belong to a new class. Therefore the model has been modified reflecting these facts. Figure 7 shows the improved result.

Contents

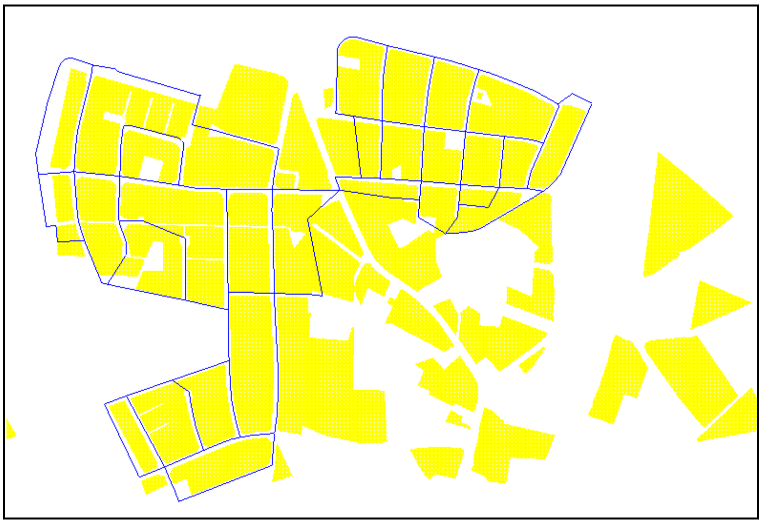

Figure 6: Overlay of the original ATKIS built-up areas (black lines) and automatically derived areas (shaded areas).

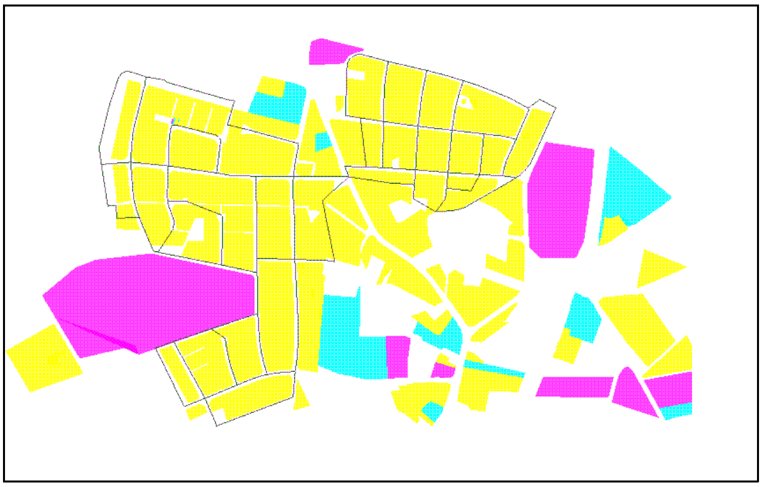

Figure 7: Built-up areas generated from ALK categorized in: residential area (light grey), area of mixed use (grey) and industrial area (dark grey).

4.3 Extensions

The result of the previous steps shows, which aggregation level is achievable based on an initial semantic modeling of the knowledge implicitely given in the data and explicitely given in terms of an - rather abstract - object description catalogue. In order to compensate for possible exceptions to this, different strategies have to be applied depending on the type of relation to be modeled.
If the relations are mainly geometric, a triangulation structure can be used to determine the neighbourhood. This process can be followed by a spatial clustering operation, e.g. minimal spanning trees (Regnauld, 1996). In the same way e.g. higher level structures of ATKIS can be derived by application of further aggregation procedures. E.g. the boundary of a city (ATKIS object id 2101) is described in the object catalogue as comprising all the built-up and industrial areas, and also including the traffic and vegetation areas. In this case the aggregation cannot base on the adjacency principle

alone, but has also to consider possible gaps in between. This can be solved with a triangulation of the settlement objects, followed by a spatial clustering. The hull around these clustered objects then represents the boundary of a city.
If however the relations between the objects are not known explicitely, machine learning techniques (Michalski et al., 1984) can be applied. This approach should be used, when examples for such relations in the data set exist - when however their reasons or explanations are not known. Then the input to an ML-program is a set of examples specifying the data sharing the relations, together with a corresponding classification (the type or name of a relation). The ML-system is then able to reveal the underlying links and describe them in terms of semantic concepts or rules.

5. CONCLUSION

The scope of the paper was to motivate the usefulness and the need for techniques of spatial data base interpretation. 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 important to automize the interpretation and revision of digital landscape models. Spatial data mining techniques are one tool to make the data re-use possible and also allow for a utilization of the data beyond their original purpose. A list of possible applications and methods for spatial data mining have been presented.
The example showed the usefulness of such techniques in the domain of data abstraction (model generalization) for the update of digital databases like ATKIS with the help of existing ones of a higher level of detail, namely ALK. The result shows that the fusion of two databases can very efficiently be implemented based on a semantic data model and on spatial proximity operators. In the example the transfer of the semantic rules for object generation into an aggregation scheme is straightforward. Already the implementation of these first steps reveal the great potential of such techniques for data re-use. In other cases however, these links might not be that direct. Then the above described spatial data mining techniques related to machine learning techniques have to be applied.
The automated interpretation of digital landscape models is far from being solved, but in times when more and more spatial data bases are build up indepentently to (enormously) costs, there is a strong need for having tools and methods to use them in a unifying way. It is obvious that no general purpose data base can ever be generated. Also, it is getting more and more clear that each geographic process has its own level of detail where it is best understood - and at this level it should be monitored and managed. Thus there is a basic need for techniques to analyse spatial data in an automatic way.

6. REFERENCES

Anders, K.-H., D. Fritsch (1996): Automatic interpretation of digital maps for data revision, *in* 'International Archives of Photogrammetry and Remote Sensing', Vol. 31/4, ISPRS, Vienna, Austria, pp. 90-94.

ATKIS-Objektartenkatalog (n.d.): Technical report, Landesvermessungsamt Baden-Württemberg, Stuttgart.

Bollinger, T. (1996): 'Assoziationsregeln - Analyse eines data mining Verfahrens', Informatik Spektrum **19**(5), 257.

Ester, M., H.-P. Kriegel, J. Sander, X. Xu (1996): A density-based algorithm for discovering clusters in large spatial databases with noise, in 'Proceedings of 2nd. International Conference on Knowledge Discovery and Data Mining (KDD-96)'.

Ester, M., H.-P. Kriegel, X. Xu (1995): Knowledge discovery in large spatial databases: Focusing techniques for efficient class identification, in 'Advances in Spatial Databases (Proc. 4th Symp. SSD-95)', Portland, ME, pp. 67-82.

Fayyad, U., G. Piatetsky-Shapiro, P. Smyth, R. Uthurusamy (1996): Advances in Knowledge Discovery and data Mining, AAAI/MIT, Menlo Park, CA.
Fotheringham, S., P. Rogerson (1994): Spatial Analysis and GIS, Taylor and Francis.
Frawley, W., G. Piatetsky-Shapiro, C. Matheus (1991): Knowledge discovery in databases: An overview, in G. Piatetsky-Shapiro & W. Frawley, eds, 'Knowledge Discovery in Databases', AAAI/MIT Press, Menlo Park, CA, pp. 1-27.
Fritsch, D., K.-H. Anders (1996): 'Objektorientierte Konzepte in Geo-Informationssystemen', GIS - Geo-Informations-Systeme **9**(2), pp. 2-14.
Haala, N., K.-H. Anders (1996): Fusion of 2d-gis and image data for 3d building reconstruction, in 'International Archives of Photogrammetry and Remote Sensing', Vol. 31/3, ISPRS, Vienna, Austria, pp. 285-290.
Holsheimer, M., A. Siebes (1994): Data mining: The search for knowledge in databases, Technical Report CS-R9406, CWI, Amsterdam, The Netherlands.
ISL (1996): Product description demo of data mining tool "clementine".
Knorr, E., R. Ng (1995): Applying computational geometry concepts to discovering spatial aggregate proximity relationships, Technical report, University of British Columbia.
Koperski, K., J. Adhikary, J. Han (1996): Knowledge discovery in spatial databases progress and challenges, in 'Proceedings of Workshop on Research Issues on Data Mining and Knowledge Discovery', Montreal, QB.
Koperski, K., J. Han (1995): Discovery of spatial association rules in geographic information databases, in 'Advances in Spatial Databases (Proc. 4th Symp. SSD'95)', Portland, ME, pp. 47-66.
Lu, W., J. Han, B. Ooi (1993): Discovery of general knowledge in large spatial databases, in ' Proc. of 1993 Far East Workshop on Geographic Information Systems (FEGIS'93)', Singapore, pp. 275-289.
Michalski, R., J. Carbonell, T. Mitchell (1984): Machine Learning - An Artificial Intelligence Approach, Springer-Verlag, Berlin.
Molenaar, M. (1996): The role of topologic and hierarchical spatial object models in database generalization, in M. Molenaar, ed., 'Methods for the Generalization of Geo-Databases', number 43, Netherlands Geodetic Commission, Delft, The Netherlands, pp. 13-35.
Ng, R., J. Han (1994): Efficient and effective clustering method for spatial data mining, in 'Proc. of 1994 Int. Conf. on Very Large Data Bases (VLDB'94)', Santiago, Chile, pp. 144-155.
van Oosterom, P. (1995): The GAP-tree, an approach to 'on-the-fly' map generalization of an area partitioning, Taylor & Francis, pp. 120-132.
Piatetsky-Shapiro, G., W. Frawley (1991): Knowledge Discovery in Databases, AAAI/MIT Press, Menlo Park, CA.
Plietker, B. (1994): Semiautomatic revision of street objects in atkis database dlm 25/1, in 'International Archives of Photogrammetry and Remote Sensing', ISPRS Commission IV, Atlanta, USA, pp. 311-317.
Preparata, F., M. Shamos (1985): Computational Geometry: An Introduction, Springer-Verlag, New York.
Regnauld, N. (1996): Recognition of building clusters for generalization, in M. Kraak M. Molenaar, 'Advances in GIS Research, Proc. of 7th Int. Symposium on Spatial Data Handling (SDH)', Vol. 2, Faculty of Geod. Engineering, Delft, The Netherlands, p. Session 4B.
Sester, M. (1995): Lernen struktureller Modelle für die Bildanalyse, Vol. C441, Deutsche Geodätische Kommission, München.
Shaw, G., D. Wheeler (1994): Statistical Techniques in Geographical Analysis, David Fulton, London.
Walter, V., D. Fritsch (1995): Matching techniques for road network data in different data models, in J. Soliman D. Roller, '28th International Symposium on Automotive Technology and Automation', Automotive Automation Limited, Croydon, England, pp. 633-640.