

Automatische Interpretation Digitaler Landschaftsmodelle

Karl-Heinrich Anders

1 Einleitung

Die Notwendigkeit der automatisierten Fortführung von raumbezogenen Daten wird heutzutage immer deutlicher, da die große Menge an raumbezogenen Daten, die in digitaler Form vorliegen, stetig ansteigt. All diese raumbezogenen Daten basieren überwiegend auf verschiedenen Modellen der Landschaft. Unterschiede in diesen Modellen betreffen z. B. die Objektart, den Grad an Generalisierung oder die geometrische Genauigkeit der erfaßten Landschaftobjekte. Eine rein interaktive Bearbeitung und Analyse großer raumbezogener Datenbanken ist extrem zeit- und kostenintensiv. Besonders die operationelle Analyse raumbezogener Daten zu Fortführungszwecken wird in naher Zukunft an ihre Grenzen der Realisierbarkeit stoßen. Diese Problematik erfordert, Methoden des sogenannten *Spatial Data Mining* oder *Knowledge Discovery in Spatial Databases (KDSD)* in Geo-Informationssystemen zu integrieren. Unter dem Begriff Knowledge Discovery in Spatial Databases versteht man die Ableitung markanter, impliziter und vorher unbekannter Informationen aus großen räumlichen Datenbanken. Das Aufgabenfeld von KDSD integriert die Gebiete des Maschinellen Lernen, Datenbanksysteme, Datenvisualisierung, Statistik, Informationstheorie und algorithmischer Geometrie. Einen Einblick in das relativ junge Gebiet des Spatial Data Mining erhält man durch folgende Literatur : (Anand, Bell & Hughes 1993), (Bell, Anand & Shapcott 1994), (Ester, Kriegel & Xu 1995), (Ester, Kriegel, Sander & Xu 1996), (Fotheringham & Rogerson 1994), (Guan & Bell 1991), (Han, Nishio & Kawano 1994), (Koperski & Han 1995), (Koperski, Adhikary & Han 1996), (Lu, Han & Ooi 1993), (Mohan & Kashyap 1988), (Molenaar 1996), (Ng & Han 1994), (Regnauld 1996), (Shaw & Wheeler 1994).

Die Anwendung der automatischen Interpretation von digitalen Landschaftsmodellen zur Datenfortführung kann man in drei Gebiete einteilen :

- Ableitung von Vorinformationen für die bildbasierte Fortführung
- Fortführung basierend auf raumbezogenen Daten gleichen Maßstabs
- Fortführung basierend auf raumbezogenen Daten mit unterschiedlichem Maßstab.

Innerhalb unseres Projektes *Automatische Interpretation Digitaler Landschaftsmodelle* untersuchen wir drei Bereiche

- Objektorientierte Modellierung von ATKIS
- Ableitung von Vorinformationen aus digitalen Landschaftsmodellen
- Fortführung von digitalen Landschaftsmodellen.

Im ersten Kapitel gehen wir kurz auf die objektorientierte Modellierung von ATKIS ein. Die Ableitung von Vorinformationen aus digitalen Landschaftsmodellen untersuchten wir anhand der 3D-Gebäuderekonstruktion basierend auf Bild-, Laser- und 2D-Grundriß-Daten. Diese Untersuchungen werden näher im dritten Kapitel beschrieben. Die Fortführung, gestützt auf raumbezogene Daten mit unterschiedlichem Maßstab, wird im vierten Kapitel beschrieben.

2 Abbildung von ATKIS auf ein objektorientiertes Datenmodell

Das Amtliche Topographisch Kartographische Informationssystem (ATKIS) (ATKIS 1988) ist ein objektstrukturiertes Datenmodell für die Modellierung und Erfassung raumbezogener Daten. In ATKIS werden im wesentlichen die objektorientierten Konzepte *Klasse* und *Aggregation* verwendet. Ein ATKIS-Objekt besteht aus einem Objektteil, der die geometrischen Informationen beinhaltet. ATKIS-Objekte wiederum können zu sogenannten Komplexen-Objekten aggregiert werden. Im Gegensatz zur Aggregation wird das Konzept der Klasse in ATKIS nicht im Sinne des objektorientierten Paradigmas verwendet. ATKIS wird hierarchisch in *Objektbereich*, *Objektgruppe* und *Objektart* eingeteilt. Diese Hierarchie kann als eine Art Klassenhierarchie im objektorientierten Sinne gedeutet werden. Sie stellt jedoch nur eine rein begriffliche Gliederung der in ATKIS erfaßten räumlichen Objekte dar. Die Hierarchiestufen Objektbereich und Objektgruppe besitzen jedoch keine eigenen Attribute und haben keine weitere nähere Bedeutung. ATKIS stellt somit keine objektorientierte Klassenhierarchie zur Verfügung, die in mehreren Objektarten vorkommenden Attribute in Oberklassen zusammenfaßt.

Aufgrund des fehlenden Konzepts der *Vererbung* kann die Definition des ATKIS-Datenmodells nicht redundanzfrei erfolgen. In jeder Objektart müssen die Attribute direkt definiert werden. Dadurch wird die Wiederverwendung und Erweiterbarkeit des Datenmodells erschwert und ist nur mit einem höheren Aufwand möglich. Globale Änderungen können nicht in einer entsprechenden Oberklasse ausgeführt werden, sondern müssen explizit in allen Unterklassen (im Fall von ATKIS Objektarten) vorgenommen werden. Die Verwendung aller objektorientierten Konzepte der Datenmodellierung in ATKIS hätte natürlich auch den Vorteil, daß das externe Datenmodell ATKIS direkt ohne Konvertierung in ein objektorientiertes Datenmodell integriert werden könnte, was von Vorteil für Hersteller objektorientierter Geoinformationssysteme wäre.

Innerhalb unseres DFG-Projekts *Automatische Interpretation Digitaler Landschaftsmodelle* erzeugen und verwenden wir ATKIS-Daten. Deshalb müssen wir in der Lage sein, ATKIS-Daten in unserem objektorientierten Datenmodell darstellen zu können. Um nun die ATKIS-Daten innerhalb eines objektorientierten Systems nutzen zu können, muß das ATKIS-Datenmodell in geeigneter Form auf ein voll objektorientiertes Datenmodell abgebildet werden. Unser Datenmodell lehnt sich an das Standard Datenmodell *SAIF* (Spatial Archive Interchange Format) (*Spatial Archive and Interchange Format: Formal Definition Release 3.2* 1995) an. Die in den Abbildungen verwendete *OMT-Notation* (Object Modeling Technique) wird in (Rumbaugh, Blaha, Premerlani, Eddy & Lorensen 1991) beschrieben. Das Symbol der Raute steht bei der OMT-Notation für die Aggregation und das Dreieck für die Vererbung. In unserem Datenmodell werden grundlegende *Containerklassen* verwendet, wie sie jedes objektorientierte Datenbanksystem zur Verfügung stellt (Abb. 1). Ein geographisches Objekt wird durch die Klasse *GeoObjekt* beschrieben (Abb. 2). Jedes GeoObjekt besitzt ein Objekt *Thematik* (Abb. 3), das zur thematischen Modellierung dient und ist als assoziative Zugriffsstruktur implementiert. Die Klasse *Raumbezug* dient zur Modellierung der Objektgeometrie (Abb. 4 bis 7). Um verschiedene Relationen zwischen räumlichen Objekten modellieren zu können, wird die Klasse *Relation* (Abb. 8) eingeführt.

Wenn man nun die Definition von ATKIS-Objekten redundanzfrei auf ein objektorientiertes Datenmodell abbilden will, stößt man auf das Problem, daß in ATKIS die Attribute einer Objektart

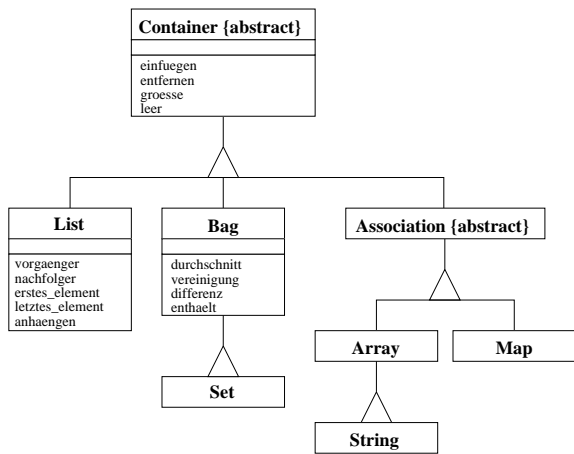


Abbildung 1: Container-Klassen.

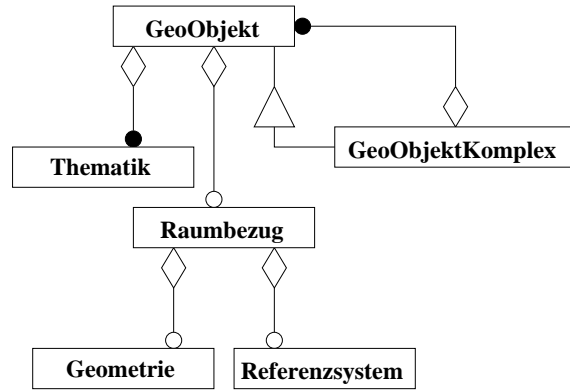


Abbildung 2: Klasse GeoObjekt.

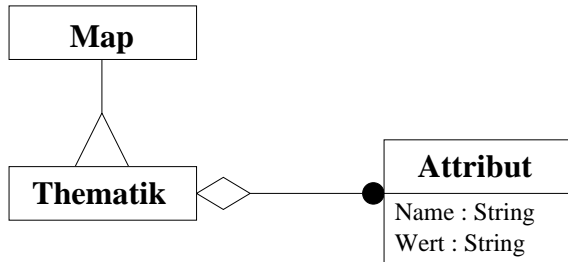


Abbildung 3: Klasse Thematik.

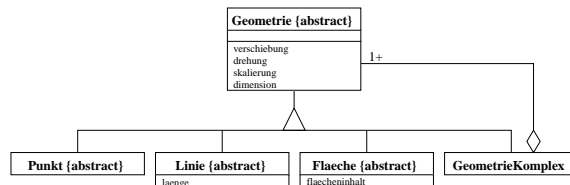


Abbildung 4: Klasse Geometrie.

nicht immer eindeutig im Zusammenhang mit der Klassenhierarchie stehen. Das soll anhand der ATKIS-Objektart Straße verdeutlicht werden (Abb. 9). Um eine redundanzfreie Klassenhierarchie aufbauen zu können ist es möglich, neue Zwischenklassen einzuführen (Abb. 10). Diese Hilfsklassen besitzen jedoch nur wenig semantische Informationen. Eine andere Möglichkeit ist, jedem Objekt einer bestimmten Objektgruppe alle Attribute, die in dieser Gruppe vorkommen, zu vererben und für nicht erfaßte Attribute einen Standardwert, wie z. B. *Nicht Erfaßt*, einzuführen. Weitere Untersuchungen haben gezeigt, daß diese zweite Variante die geeignetere ist. Abbildung 11 zeigt am Beispiel Straße, wie ein ATKIS-Objekt in unserem Datenmodell dargestellt werden kann. Diese Arbeiten wurden in (Fritsch & Anders 1996) veröffentlicht.

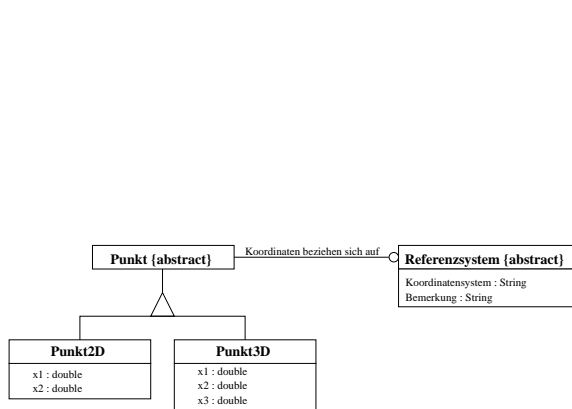


Abbildung 5: Klasse Punkt.

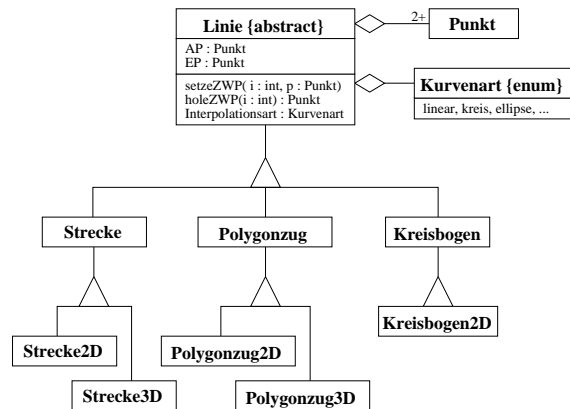


Abbildung 6: Klasse Linie.

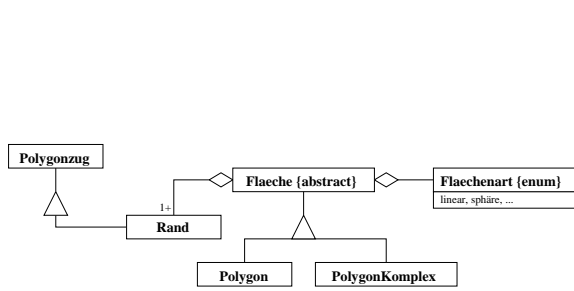


Abbildung 7: Klasse Fläche.

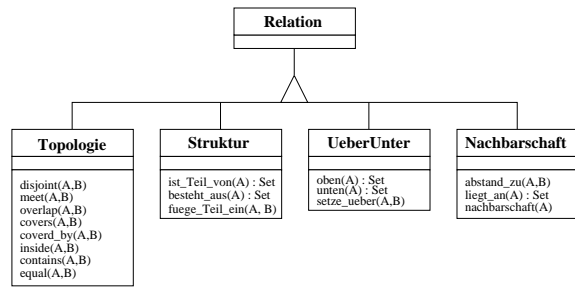


Abbildung 8: Klasse Relation.

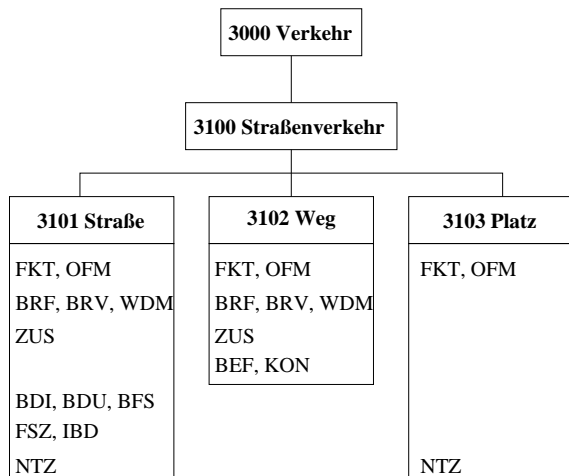


Abbildung 9: ATKIS Objektbereich 3000 Verkehr.

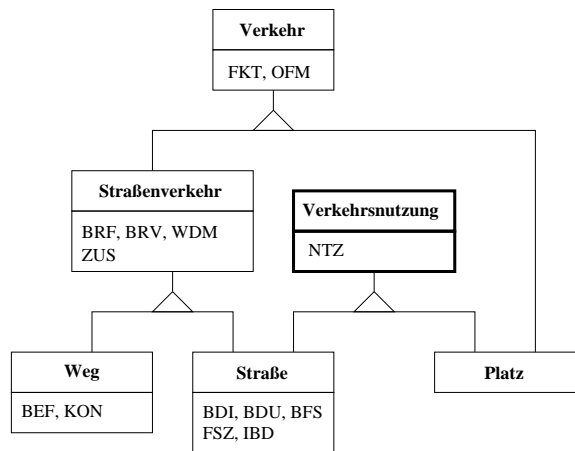


Abbildung 10: Redundanzfreie Klassenhierarchie für Verkehr durch Einführung von Zwischenklassen.

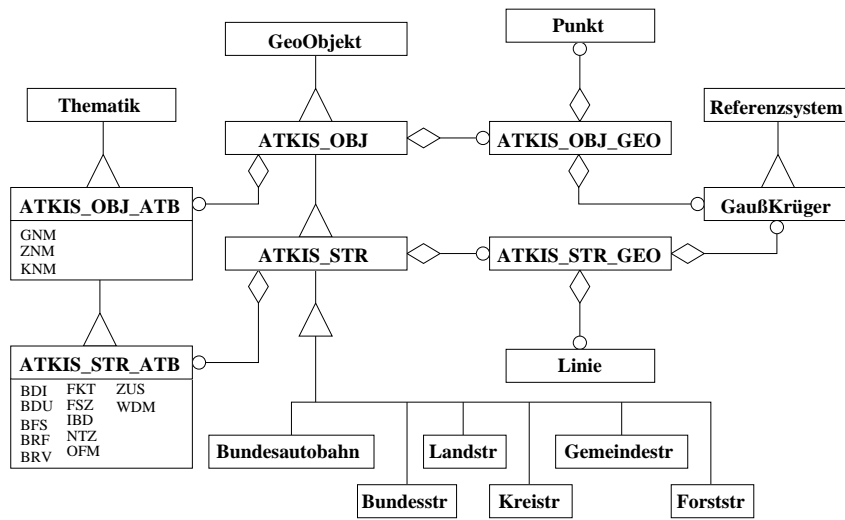


Abbildung 11: Beispiel für ATKIS-Objektart Straße.

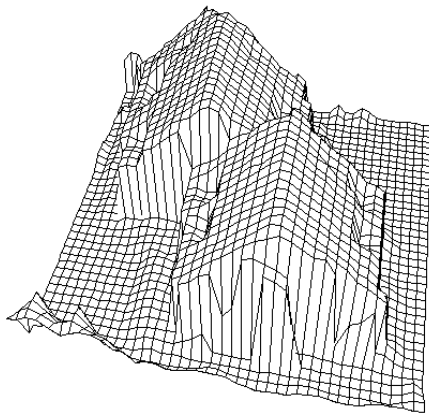


Abbildung 12: Ausschnitt aus einem DHM, erfaßt mit einem Laserscanner.

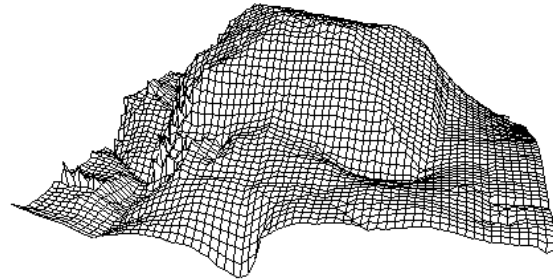


Abbildung 13: Ausschnitt aus einem DHM, erfaßt durch Stereobildzuordnung.

3 Ableitung von 3D-Gebäudehypothesen für die 3D-Gebäude-rekonstruktion, basierend auf digitalen Oberflächenmodellen

Die Nachfrage nach digitalen 3D-Stadtmodellen von Seiten der Stadtplaner nimmt kontinuierlich zu. Daraus ergab sich die Frage, ob sich vorhandene 2D-Gebäudedaten nutzen lassen, um den Prozeß der 3D-Gebäuderekonstruktion aus Luftbildern oder Laserdistanzdaten zu unterstützen (Haala & Anders 1996).

Bereits vorhandene Gebäudegrundrisse in einem zweidimensionalen Geo-Informationssystem sind eine zuverlässige Zusatzinformationsquelle für die Erfassung qualifizierter Gebäudehöhenmodelle aus direkt mit Laserscanner (Abb. 12) bzw. indirekt durch Stereobildzuordnung (Abb. 13) erfaßten Digitalen Höhenmodellen (DHM). Ein Beispiel für solche Daten ist das amtliche Liegenschaftskataster (ALK), in der sowohl Grundrißinformationen als auch die jeweilige Nutzungsart der Gebäude enthalten ist. Neben der Aktualisierung dieses Bestandes ist die Erweiterung der vorhandenen zweidimensionalen Daten durch Informationen über die dritte Dimension ein wichtiger Anwendungsbereich. Daten über die dritte Dimension werden besonders auf dem Gebiet genauer und dichter Höhenmodelle in Stadtgebieten benötigt. Durch die Analyse bzw. Interpretation der vorhandenen Grundrisse können beispielsweise Kontrollparameter



Abbildung 14: Bildausschnitt vom Testgebiet.



Abbildung 15: Ausschnitt aus der verwendeten ALK.

für die Segmentierung des DHM bzw, für die anschließende dreidimensionale Rekonstruktion bestimmt werden. Abbildung 14 zeigt eine Luftaufnahme des Testgebiets und Abbildung 15 den dazu gehörenden Ausschnitt aus der ALK.

3.1 Analyse der Gebäudeinformationen

Bei geringen Genauigkeitsanforderungen an die zu erfassende dreidimensionale Form eines Gebäudes kann eine Beschreibung durch einfache Blöcke erfolgen. In diesem Fall kann der Gebäudegrundriß aus dem existierenden GIS ohne weitere Analyse verwendet werden; die benötigte Gebäudehöhe kann dann als die maximale Höhe, innerhalb des Grundrisses, aus dem DHM bestimmt werden. Für Anwendungen wie Visualisierungen ist häufig ein mittlerer Detaillierungsgrad ausreichend, für den neben dem Grundriß noch der Dachtyp (z. B. Flach- oder Walmdach) sowie einige wenige Höhenwerte (z. B. First-, Trauf- und Fußbodenhöhe) zu bestimmen sind. Durch die Analyse der ALK-Daten können aus der Form und Lage der Gebäudegrundrisse und den ebenfalls vorhandenen Nutzungsarten Hypothesen (Abb. 17, Abb. 18) über die dreidimensionale Form gebildet werden. Die unterschiedlichen Hypothesen können anschließend mit Hilfe der DHM- und Bilddaten verworfen bzw. verifiziert werden.

Die Erzeugung von 3D-Gebäudemodellen mit Hilfe der Grundrißinformationen kann natürlich nicht eindeutig gelöst werden. Zusätzlich zu den unbekanntem Gebäudehöhen sind die unterschiedlichsten Dachformen möglich, wie z. B. Flach-, Pult-, Sattel- oder Walmdach. Da die Gebäudenutzung einen möglichen Hinweis auf die Dachform und Dachhöhen bietet, nutzen wir die in der ALK enthaltenen Informationen über die Gebäudenutzung, um die Gebäude in die sechs Klassen *Wohn-, Industrie-, Bürogebäude, Kirche, Turm* und *Garage* einzuteilen. Jede dieser Klassen ordnen wir einen Dachtyp und Dachhöhen wie folgt zu :

- Garage : Flachdach, Dachhöhe 3m
- Wohn-, Bürogebäude : Satteldach, Traufhöhe 6m, Firsthöhe 9m (das entspricht 2 Geschossen mit 3m Höhe und ein Dachgeschoß von 3m Höhe)

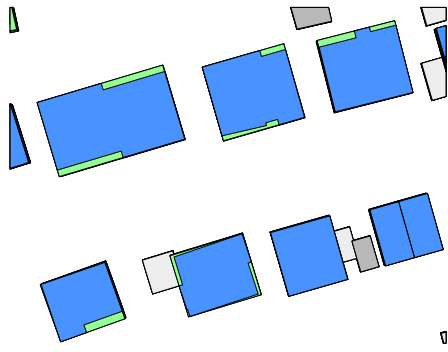


Abbildung 16: Generalisierung des Grundrisses durch minimal umschließendes Rechteck.

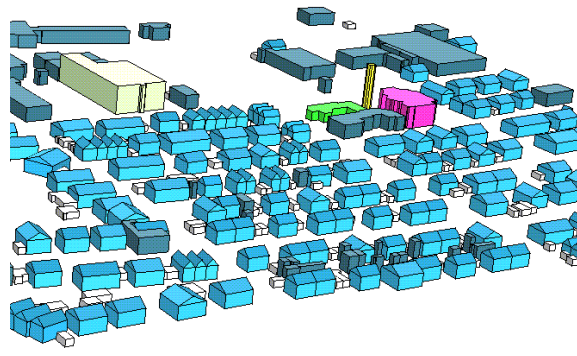


Abbildung 17: 3D-Gebäudehypothesen basierend auf den ALK-Gebäudegrundrissen.

- Industriegebäude : Flachdach, Dachhöhe 15m (3 Geschosse mit 5m Höhe)
- Kirche : Flachdach, Dachhöhe 12m
- Turm : Flachdach, Dachhöhe 25m
- Kindergarten : Flachdach, Dachhöhe 5m
- andere : Flachdach, Dachhöhe 7,5m

Diese Werte sind natürlich rein heuristisch, aber sie sollen auch nur eine erste Nahrung für das Gebäude liefern. Durch den anschließenden Verifikationsprozeß wird die Hypothese dann bestätigt und verbessert, oder sie wird verworfen und eine neue Hypothese gebildet (z. B. andere Firstrichtung oder anderer Dachtyp).

Bevor wir jedoch das 3D-Modell erzeugen führen wir noch eine Generalisierung des Gebäudegrundrisses durch. Flach- und Pultdächer folgen in den meisten Fällen der Form des Grundrisses. Bei Walm- oder Satteldächern ist das jedoch nicht der Fall. Kleine Erker oder Unregelmäßigkeiten im Grundriß werden vom Dach verdeckt. Man kann in den meisten Fällen davon ausgehen, daß die Grundfläche eines solchen Daches rechteckförmig ist. Als Generalisierung des Grundrisses berechnen wir deshalb ein minimal umschließendes Rechteck wie es in Abbildung 16 dargestellt ist. Die Firstrichtung wird dann durch die längste Seite dieses Rechtecks festgelegt. Basierend auf all den aufgezählten Annahmen erzeugen wir dann ein hypothetisches 3D-Stadtmodell. In Abbildung 17 ist das Ergebnis dargestellt. Ein Vergleich dieses Ergebnisses mit Abbildung 14 zeigt jedoch noch erhebliche Unterschiede auf. Dies geht daraufhin zurück, daß wir bisher nur die Grundrisse allein betrachtet haben. Ein Grundriß in der ALK repräsentiert ein bestimmtes Besitzverhältnis, und angrenzende Grundrisse gehören im allgemeinen zu einem Gebäude und besitzen somit ein gemeinsames Dach. Deshalb bestimmen wir topologisch benachbarte Grundrisse (Grundrisse, die sich ein Geometrieelement Linie teilen) und gruppieren diese Grundrisse, falls sie die gleiche Nutzungsart aufweisen, zu einem neuen Grundriss zusammen (Umhüllendes Polygon). Dieses neue Polygon wird dann gegebenenfalls generalisiert und dann ein 3D-Modell erzeugt. Das Ergebnis dieser verbesserten Hypothesenbildung zeigt Abbildung 18. Ein nochmaliger Vergleich mit Abbildung 14 zeigt, daß sich allein mit diesen einfachen Annahmen ein relativ gutes Stadtmodell erzeugen läßt.

Die Erzeugung von 3D-Gebäudehypothesen benötigt jedoch noch zusätzliche Metainformationen, die nicht in der ALK enthalten sind. Die Gebäudeparameter (Dachform, Dachhöhen, Firstrichtung) hängen auch von den Regeln der Stadtplaner ab. Solche zusätzlichen Bedingungen sind :

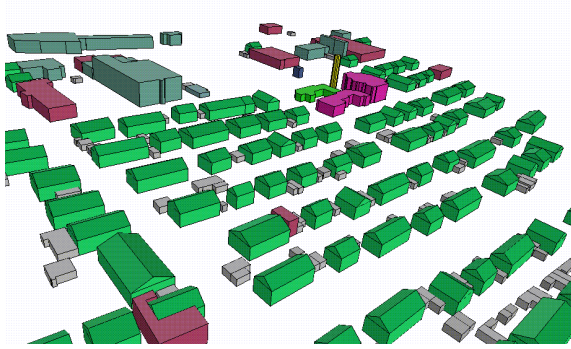


Abbildung 18: 3D-Gebäudehypothesen unter Berücksichtigung benachbarter Gebäude.



Abbildung 19: 3D-Gebäudehypothesen für einen Teilbereich von Stuttgart Vaihingen.

- Die Lage des Gebäudes zu einer Straße, da heutzutage die Firstrichtung meistens parallel zur Straßenrichtung verläuft.
- Die Nachbarschaft zu anderen, nicht angrenzenden, Gebäuden, da in den meisten Fällen benachbarte Gebäude (Häuserreihen) den gleichen Dachtyp, ähnliche Dachhöhe und gleiche Firstrichtung besitzen.
- Die Stadtlage des Gebäudes (Zentrum, Stadtrand, Industriegebiet) hat auch einen Einfluß auf die Dachform und Höhe eines Gebäudes.

Es fehlen jedoch zur Zeit noch Statistiken über die Häufigkeit von bestimmten Dachformen und ihre Abhängigkeit von Stadtlage und Gebäudenutzungsart. Ebenso fehlen noch Statistiken über die Abhängigkeit der Gebäudehöhe von der Stadtlage und Größe einer Stadt. Mit Hilfe solcher statistischen Maße könnte man ein stochastisches Stadtmodell aufstellen und die Hypothesenbildung erheblich verbessern. Anhand kleiner Gebiete sind wir dabei, solche Statistiken aufzustellen, aber in Anbetracht des Gesetzes von den großen Zahlen, ist es fraglich wie evident diese Werte sein werden. Die Einbeziehung räumlich benachbarter Straßen und Gebäude (Grundrisse, die nicht aneinander grenzen) für die Generierung der 3D-Gebäudehypothesen ist Ziel unserer nächsten Arbeiten.

3.2 Verifikation der Gebäudehypothesen

Zur Verifikation unserer 3D-Gebäudehypothesen führen wir ein Matching der Modellkanten und Modellflächen mit 3D-Kanten und 3D-Flächen aus dem DHM durch. Das DHM kann direkt durch einen Laserscanner oder indirekt durch eine Stereobildzuordnung erfaßt worden sein. Ein Beispiel für ein Laser-DHM ist in Abbildung 12 dargestellt. Abbildung 13 zeigt ein Beispiel für eine Stereozuordnung. Für die Extrahierung von Raumkanten und Ebenen aus Distanz- bzw. Höhendaten benutzen wir die Klassifikation der DHM-Punkte durch die Gauß'sche und mittlere Krümmung der Flächen. Für Anwendungen im Nahbereich werden diese Verfahren bereits seit einiger Zeit eingesetzt (Besl 1988). In diesem Bereich stehen schon länger Sensoren zur Verfügung, die eine genaue und dichte Erfassung der Oberflächengeometrie ermöglichen. Da DHM-Daten durch die mittlerweile verfügbaren flugzeuggetragenen Laserscanner mit ähnlicher Qualität bestimmt werden können, bietet sich eine Adaption der Verfahren für die Bearbeitung solcher Flächen an (Haala & Anders 1997).

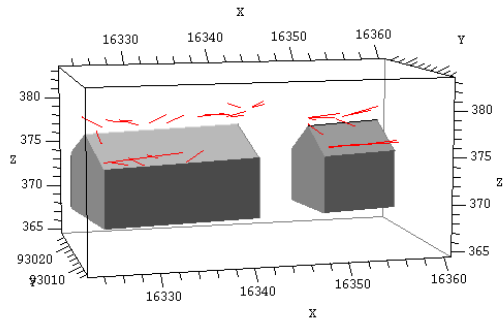


Abbildung 20: 3D-Ansicht der Gebäudehypothesen und zugeordneten DHM-Kanten.

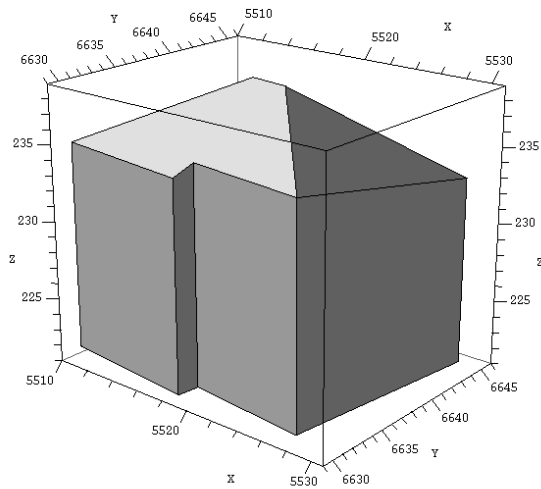


Abbildung 21: Zugeordnetes und verbessertes Gebäudemodell.

Durch die im DHM beobachteten Raumkanten und Ebenen führen wir nun eine Verifikation durch, d.h. wenn zu wenige Beobachtungen der 3D-Hypothese zugeordnet werden können verwerfen wir die Hypothese. Ansonsten verbessern wir die Hypothese durch Ausgleichung der Kanten und Ebenen. In Abbildung 21 ist ein so erzeugtes Gebäudemodell dargestellt. Die Projektion dieser Modellkanten in das Orthophoto, wie in Abbildung 22 dargestellt, zeigt die gute Übereinstimmung mit dem realen Gebäude auf.

4 Fortführung von ATKIS-Daten basierend auf ALK-Daten

In Deutschland wird ATKIS in drei unterschiedlichen Genauigkeitsgraden erfasst (1:25000, 1:200000, 1:1000000). Diese Daten umfassen topographische Objekte aus sieben Klassen (Siedlung, Verkehr, Vegetation, ...). Ebenso wird das amtliche Liegenschaftskataster (ALK) in den Maßstäben 1:1500 bis 1:2000 erfasst. Die ALK beinhaltet Flurstücksgrenzen, Gebäudegrundrisse, Nutzungsarten und Straßennamen. Da der Prozeß der Fortführung zeit- und kostenintensiv ist, entstand die Idee, ATKIS-Daten aus ALK-Daten abzuleiten. Somit müßte nur noch die detailliertere ALK manuell fortgeführt werden. Zur Zeit werden beide Datenbestände getrennt voneinander fortgeführt. Da es grundsätzlich möglich ist, ATKIS-Daten aus ALK-Daten abzuleiten, konnten wir anhand des ATKIS-Objektbereichs *Siedlung* aufzeigen. Durch die Informationen über Grundstücksgrenzen, Nutzungsarten und den implizit gegebenen geometrischen und topologischen Informationen enthält die ALK alle notwendigen Informationen, die man für die Ableitung von Siedlungsobjekten benötigt. In Abbildung 23 ist der von uns verwendete ALK-Datensatz abgebildet. Uns lag dieser Datensatz im Austauschformat *SQD-Format* des Geo-Informationssystems *SICAD/open* vor. Die Ableitung von ATKIS-Objekten des Typs *Wohnsiedlung* (ATKIS-Objektart 2111) aus ALK-Daten beinhaltet Generalisierungs- und Aggregationsoperationen, die im folgendem näher erläutert werden. Zur Beschreibung, wie räumliche Objekte in der ALK repräsentiert werden, benutzen wir die semantische Modellierung als eine konzeptionelle Methode um zu analysieren, wie ein Operateur diese Objekte erkennt. Abbildung 24 zeigt das von uns verwendete Semantische Netz für die Beschreibung von Flurstücken und Wohnsiedlungsbereichen.

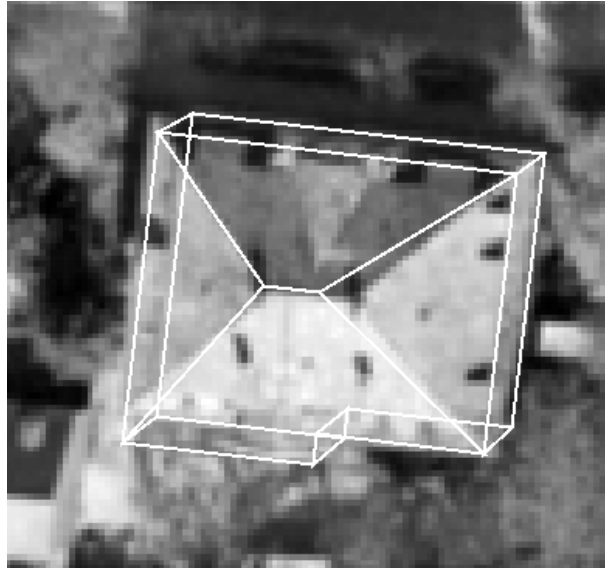


Abbildung 22: Projektion des erzeugten 3D-Modells in das zugehörige Orthophoto.

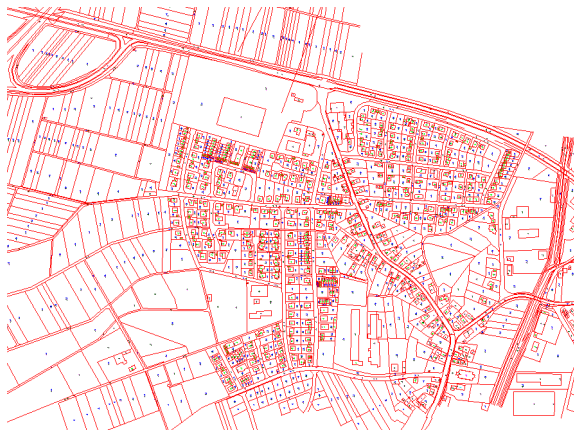


Abbildung 23: Ausschnitt aus einem ALK-Datensatz.

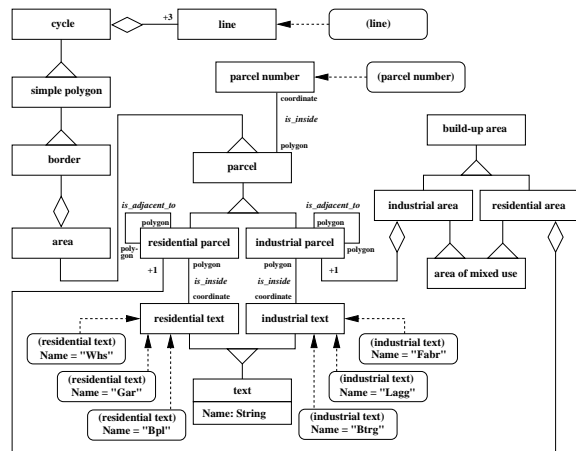


Abbildung 24: Semantisches Modell für Flurstücke und Wohnsiedlungen.

4.1 Aggregations- und Generalisierungsoperatoren

Zur Ableitung neuer räumlicher Objekte benötigt man im allgemeinen Operatoren zur Aggregation und Generalisierung von räumlichen Objekten (Regnaud 1996), (Molenaar 1996). Für die Aggregation räumlicher Objekte benötigt man eine geeignete Definition der *Objektnachbarschaft* und ein Aggregationskriterium, wie z. B. ein gleiches Objektattribut. Wir haben dazu in unserem objektorientierten Datenmodell Operatoren für die Aggregation bezüglich

- topologischer und
- räumlicher Nachbarschaft

implementiert. Bei der Berechnung der topologischen Nachbarschaft werden die Adjazenz-Relationen im Datenmodell genutzt. Die *räumliche Nachbarschaft* berechnen wir durch eine *Delaunay-Triangulation* (Preparata & Shamos 1985) aller Objektpunkte. Die Delaunay-Triangulation ermöglicht uns die Definition der räumlichen Nachbarschaft ohne Angabe eines Abstandsmaß, da diese Art der Triangulation Punkte nur mit ihren nächsten Nachbarn verbindet. Für die Generalisierung der geometrischen Form gegebenen oder aggregierten räumlichen Objekten haben wir Operatoren zur Erzeugung der *konvexen Hülle* (Preparata & Shamos 1985) und minimal umschließenden Rechteck implementiert. Im folgenden Abschnitt beschreiben wir unseren Ansatz zur Ableitung von Wohnsiedlungsobjekten aus ALK-Daten. Bei dieser Vorgehensweise wurde ausschließlich die topologische Nachbarschaft zur Aggregation räumlicher Objekte verwendet.

4.2 Ableitung der ATKIS-Objektart *Wohnsiedlung*

Das erste Problem war, daß in dem uns zur Verfügung gestellten ALK-Datensatz, die Flurstücke nicht als flächenhafte Objekte gespeichert waren. Nur die Flurstücksgrenzen waren als linienhafte Objekte gespeichert. Die Gebäudegrundrisse waren wiederum als Flächenobjekt gespeichert worden. Deshalb ist der erste Schritt in unserem Ansatz, anhand der Linienobjekte und deren topologischen Beziehungen untereinander, alle möglichen geschlossenen Flächen zu berechnen. Eine Fläche wird in unserem Fall durch ein einfaches Polygon repräsentiert. Das einfache Polygon wiederum stellt sich als geschlossene Masche in dem durch die Linienobjekte aufgespannten Netz dar.

Im nächsten Arbeitsschritt bestimmen wir, in welchen von diesen Flächen eine Flurstücknummer liegt. Dies wird durch einen Punkt-in-Polygon-Test für alle möglichen Kombinationen von Flächen und Flurstücksnummern erreicht. Durch eine raumbezogene Zugriffsstruktur kann der zeitliche Aufwand für solche geometrischen Operationen noch erheblich verringert werden. Das Ergebnis nach diesem Schritt ist in Abbildung 25 dargestellt. Nachdem alle Flurstücke ermittelt wurden, sind nun all diejenigen Flurstücke auszuwählen, die einem Wohnsiedlungsbereich zugeordnet werden können. Dazu wählen wir aus allen vorhandenen Textsymbolen im ALK-Datensatz die Symbole *Whs*, *Gar*, *Bpl* (Wohnhaus, Garage, Bauplatz) aus, und führen wieder einen Punkt-in-Polygon-Test für alle möglichen Kombinationen dieser Symbole und den Flurstücksflächen aus.

Im letzten Arbeitsschritt wird eine Gruppierung aller extrahierter Wohnsiedlungsflächen vorgenommen, um ATKIS-Objekte des Typs Wohnsiedlung zu erstellen. Die semantische Regel für die Gruppierung solcher Flächen ist das Prinzip der Adjazents (rechter Teil in Abb. 24). Mit anderen Worten, alle diese extrahierten Flächen, die eine gemeinsame Flurstücksgrenze besitzen, werden zusammengefaßt. Diese Berechnung kann einfach ausgeführt werden, indem die

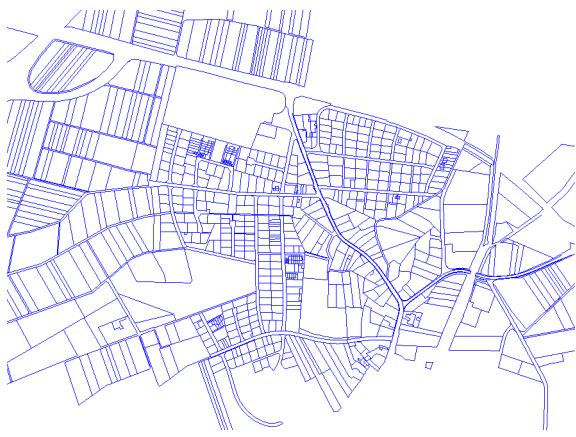


Abbildung 25: Ermittelte Flurstücke.

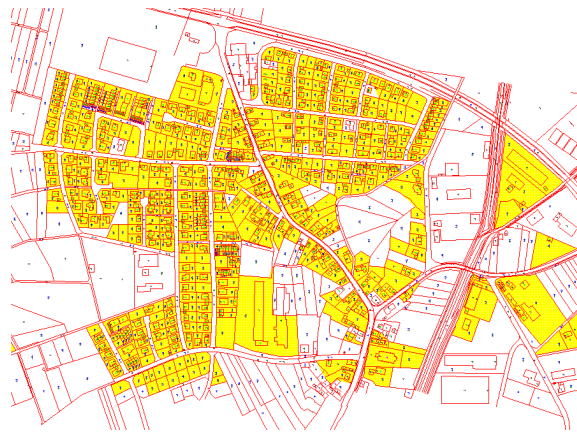


Abbildung 26: ALK-Daten mit erzeugten Wohnsiedlungen.

topologischen Relationen benutzt werden, die aus unserem objektorientierten Datenmodell direkt abgeleitet werden können. Abbildung 26 zeigt das Ergebnis dieser Gruppierung, überlagert mit den ALK-Daten. Dieses Ergebnis wurde allein durch Aggregation angrenzender Wohnsiedlungsflächen ermittelt. Der Vergleich dieser automatisch erzeugten ATKIS-Objekte mit original ATKIS-Objekten zeigt einige Unterschiede auf (Abb. 27). Ursachen für diese Unterschiede sind :

- Verschiedene Erfassungszeiträume der ALK- und ATKIS-Daten.
- Unterschiedliche Erfassungsregeln.
- Subjektive Interpretation des Operateurs bei der Erfassung der ATKIS-Daten aus Orthophotos.
- Die Grenzen der Wohnsiedlungsbereiche stimmen deshalb nicht exakt überein, da in dem ATKIS-Datensatz die Straßen nicht als Fläche erfaßt wurden, sondern durch ihre Mittelachsen repräsentiert werden. Da in ATKIS topologisch angrenzende Objekte durch die gleichen geometrischen Elemente dargestellt werden, treffen sich die Siedlungsgrenzen bei den Mittelachsen der Straßen. Unsere Grenzen entsprechen jedoch den genauen Grenzen zu den ALK-Straßenflächen. Man kann dieses Problem umgehen, indem man entweder gleichzeitig ATKIS-Straßenobjekte als Flächenobjekte aus den ALK-Daten ableitet oder einen Generalisierungsschritt einführt, der die Mittelachsen bestimmt und dann die Grenzen der Siedlungsflächen modifiziert. Man kann jedoch sagen, daß eine getrennte Ableitung von Siedlungsobjekten und Straßenobjekten nicht sinnvoll erscheint, sondern beide gemeinsam ermittelt werden müssen.
- Ein weiterer Unterschied entsteht dadurch, daß in diesem ersten Ansatz Flächen mit gemischter Nutzung (Industrie- und Wohngebäude in einem Flurstück) nicht modelliert wurden und somit ebenfalls in unserem Ergebnis enthalten sind. Wir erweiterten deshalb unser Modell um die ATKIS-Klassen *Wohnsiedlung*, *Industriegebiet* und *Fläche mit gemischter Nutzung* Abbildung 28 zeigt das verbesserte Ergebnis.

Zusammenfassend kann man sagen, daß unsere Ergebnisse, die Verwendung großmaßstäbiger Daten (z. B. ALK) zur automatischen Fortführung von Daten mittlerem Maßstabs (z. B. ATKIS), bestätigen. Unser pragmatischer Ansatz zeigt, daß die Verknüpfung unterschiedlicher räumlicher Datenbestände effizient mit Hilfe eines semantischen Modells und räumlicher Aggregierungs-

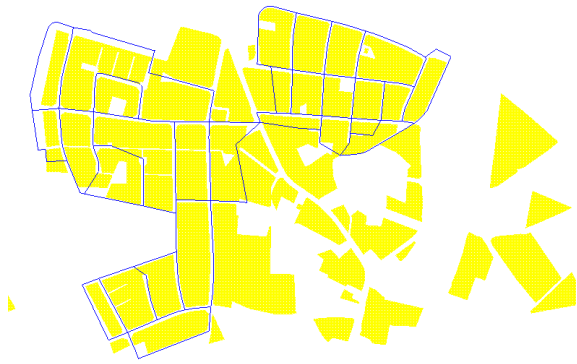


Abbildung 27: Überlagerung von original ATKIS Wohnsiedlungsbereichen (schwarz) und automatisch abgeleiteten Flächen (grau).

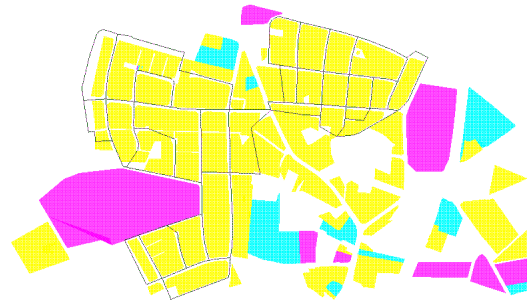


Abbildung 28: Automatisch abgeleitete Wohnsiedlungsbereiche (hellgrau), Industriegebiete (dunkelgrau) und gemischter Nutzung (grau).

und Approximationsoperatoren durchführbar ist. Ein geeignetes semantisches Modell ist jedoch nicht immer vorhanden (keine Erfassungsregeln vorhanden) oder kann nur schwer erstellt werden (Verknüpfung der Erfassungsregeln und dem verwendeten Datenbestand ist nicht direkt ersichtlich). Für solche Fälle ist es zwingend erforderlich, unser Vorgehen durch Methoden des *Maschinellen Lernens* (Michalski, Carbonell & Mitchell 1984) zu erweitern. Durch die Methode *Lernen durch Beispiele* könnten dann automatisch semantische Konzepte oder Regeln abgeleitet werden. Eine andere Möglichkeit besteht darin, von einer sehr allgemeinen Regel auszugehen, und beim Auftreten von Ausnahmen von dieser Regel, durch die Methode des *Inkrementelles Lernen*, das Verhalten des Systems zu verbessern. Zusätzlich ist es noch erforderlich, eine automatische Beurteilung der Ergebnisse zu ermöglichen; insbesondere bei der Verwendung maschineller Lernverfahren, da hier nur von *unscharfem Wissen* ausgegangen werden muß. Die Integration von maschinellen Lernverfahren und einer geeigneten Beurteilungsmethode sind die Ziele unseres Folgeantrags. Veröffentlichungen zu diesem Thema findet man in (Anders & Fritsch 1996) und (Anders & Sester 1997).

Literatur

- Anand, S., Bell, D. & Hughes, J. (1993), 'A general framework for database mining based on evidential theory', Internal Report, Dept. of Inf.Sys., University of Ulster at Jordanstown.
- Anders, K.-H. & Fritsch, D. (1996), Automatic Interpretation of Digital Maps for Data Revision, *in*: 'International Archives of Photogrammetry and Remote Sensing', Vol. 31(4), ISPRS, Vienna, Austria, Seiten 90–94.
- Anders, K.-H. & Sester, M. (1997), Methods of Data Base Interpretation - Applied to Model Generalization from Large to Medium Scale, *in*: W. Förstner & L. Plümer, Hrsg., 'Semantic Modeling for the Acquisition of Topographic Information from Images and Maps / SMATI 97', Birkhäuser, Bonn, Germany, Seiten 89–103.
- ATKIS (1988), *ATKIS-Gesamtdokumentation: Amtliches Topographisch-Kartographisches Informationssystem – Das Vorhaben der Landesvermessungsverwaltungen zum Aufbau Digitaler Landschaftsmodelle und Digitaler Kartographischer Modelle*, Landesvermessungsamt Nordrhein-Westfalen.

- Bell, D., Anand, S. & Shapcott, C. (1994), 'Database Mining in Spatial Databases', International Workshop on Spatio-Temporal Databases.
- Besl, J. (1988), *Surfaces in Range Image Understanding*, Springer series in perception engineering, Springer-Verlag, New York Berlin Heidelberg.
- Ester, M., Kriegel, H.-P. & Xu, X. (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, Seiten 67–82.
- Ester, M., Kriegel, H.-P., Sander, J. & Xu, X. (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)'.
- Fotheringham, S. & Rogerson, P. (1994), *Spatial Analysis and GIS*, Taylor and Francis.
- Fritsch, D. & Anders, K.-H. (1996), 'Objektorientierte Konzepte in Geo-Informationssystemen', *GIS GEO-Informationssysteme* 9(2), 2–14.
- Guan, J. & Bell, D. (1991), *Evidence Theory and its Applications*, Vol. 1, North-Holland.
- Haala, N. & Anders, K.-H. (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, Seiten 285–290.
- Haala, N. & Anders, K.-H. (1997), Acquisition of 3D urban models by analysis of aerial images, digital surface models and existing 2D building information, *in*: 'SPIE Conference on Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision III', SPIE, Orlando, Florida, Seiten 212–222.
- Han, J., Nishio, S. & Kawano, H. (1994), Knowledge Discovery in Object-Oriented and Active Databases, *in*: F. Fuchi & T. Yokoi, Hrsg., 'Knowledge Building and Knowledge Sharing', Ohmsha/IOS Press, Seiten 221–230.
- Koperski, K., Adhikary, J. & Han, J. (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. & Han, J. (1995), Discovery of Spatial Association Rules in Geographic Information Databases, *in*: 'Advances in Spatial Databases (Proc. 4th Symp. SSD'95)', Portland, ME, Seiten 47–66.
- Lu, W., Han, J. & Ooi, B. (1993), Discovery of General Knowledge in Large Spatial Databases, *in*: 'Proc. of 1993 Far East Workshop on Geographic Information Systems (FEGIS'93)', Singapore, Seiten 275–289.
- Michalski, R., Carbonell, J. & Mitchell, T. (1984), *Machine Learning - An Artificial Intelligence Approach*, Springer-Verlag, Berlin.
- Mohan, L. & Kashyap, R. (1988), An Object-Oriented Knowledge Representation for Spatial Information, *in*: 'IEEE Transaction on Software Engineering', number 5 *in*: 'IEEE Transaction on Software Engineering', Seiten 675–681.
- Molenaar, M. (1996), The role of topologic and hierarchical spatial object models in database generalization, *in*: M. Molenaar, Hrsg., 'Methods for the Generalization of Geo-Databases', number 43, Netherlands Geodetic Commission, Delft, The Netherlands, Seiten 13–35.

- Ng, R. & Han, J. (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, Seiten 144–155.
- Preparata, F. & Shamos, M. (1985), *Computational Geometry: An Introduction*, Springer-Verlag, New York.
- Regnauld, N. (1996), Recognition of building clusters for generalization, *in*: M. M. Kraak, Hrsg., 'Proceedings of 7th Int. Symp. on Spatial Data Handling (SDH)', Vol. 2, Faculty of Geod. Engineering, Delft, The Netherlands.
- Rumbaugh, J., Blaha, M., Premerlani, W., Eddy, F. & Lorensen, W. (1991), *Object-Oriented Modeling and Design*, Prentice Hall, Inc.
- Shaw, G. & Wheeler, D. (1994), *Statistical Techniques in Geographical Analysis*, David Fulton, London.
- 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.