Das Minimum-Maximum-Modell - Eine qualitative Beschreibung der geometrischen Genauigkeit und ihre Integration in GIS

Michael Glemser, Sandra Krauß, Christian Schneider
Institut für Photogrammetrie
Universität Stuttgart

1 Einführung

Geo-Informationssysteme (GIS) erfassen, verwalten, analysieren und präsentieren raumbezogene Daten. In den meisten Systemen erfolgt dabei eine Strukturierung der Daten in Form von räumlichen Objekten, deren Lage und Ausdehnung durch eine *geometrische Komponente* und deren Thematik durch eine *attributive Komponente* beschrieben werden. Die Geometrie eines Objektes folgt einer Randbeschreibung, durch die das Innere scharf vom Äußeren getrennt wird. Für flächenhafte Objekte z.B. ist der Rand durch eine Anzahl an Umringslinien definiert. Besonders für natürliche Objekte trifft aber die scharfe Abgrenzung im allgemeinen nicht zu. Z.B. ein Wald kann nicht exakt von den ihn umgebenden Wiesen abgegrenzt werden. Die scharfe Abgrenzung entspricht einem Abstraktionsprozeß, der die erzeugten GIS-Daten zu einem gewissen Grade unsicher macht. Man spricht hier von der *Abstraktionsunschärfe*, die in der Natur der Objekte begründet liegt. Weiterhin sind die zur Erfassung notwendigen Beobachtungsmethoden mit zufälligen, systematischen und groben Fehlern behaftet und verursachen so eine zusätzliche Unsicherheit in der Position der Ränder. Da solche Einflüsse auf alle Daten in gleichem Maße wirken, ist immer vom Bestehen einer gewissen Unsicherheit auszugehen. Die auf die Geometrie bezogene Unsicherheit wird als *geometrische Genauigkeit* bezeichnet.

Innerhalb eines GIS können sehr unterschiedlich genau Datenbestände vorkommen (z.B. hochgenaue Katasterdaten und ungenaue geologische Daten), die sich aber trotzdem ohne Probleme gemeinsam verarbeiten lassen. Um dennoch auch nach komplexen Verarbeitungsschritten Aussagen zur Genauigkeit des Ergebnisses treffen zu können, muß eine Integration der geometrischen Genauigkeit in die Systemumgebung erfolgen. Sie dient dazu dem Anwender den Umgang mit ungenauen Daten zu erleichtern. Die Integration erfordert gezielte Erweiterungen bei der Verwaltung und Analyse. Im speziellen sind Datenhaltungskonzepte zu entwickeln und alle Analysefunktionen zu ergänzen, um eine Berücksichtigung und Fortpflanzung der geometrischen Genauigkeit auf das Ergebnis zu ermöglichen. Zusätzlich muß sich um eine ansprechende Visualisierung der Unsicherheit gekümmert werden. Diesen Aufgaben widmet sich der vorliegende Beitrag.

Zunächst wird näher auf die Beschreibung der geometrischen Genauigkeit in GIS eingegangen. Dazu erfolgt die Formulierung eines Genauigkeitsmodells - *das Minimum-Maximum-Modell*. Seine Integration in die Umgebung eines bestehenden GIS ist Thema eines weiteren Abschnitts. Welche Auswirkungen die geometrische Genauigkeit auf die Analysefunktionen besitzen, wird an drei Beispielen verdeutlicht: am *Punkt-im-Polygon-Test*, an der *geometrischen Verschneidung* und an der *Festlegung topologischer Relationen*. Zum Schluß werden die gefundenen Erkenntnisse zusammengefaßt.

2 Minimum-Maximum-Modell

Die Integration der geometrischen Genauigkeit in GIS erfordert die Definition eines Genauigkeitsmodells, das die notwendigen Beschreibungsmaße bereitstellt. Dieser Arbeit wird das *Minimum-Maximum-Modell* zugrundegelegt. Doch vor einer genauen Definition des Modells, soll der Begriff der geometrischen Genauigkeit näher erläutert werden.

2.1 Geometrische Genauigkeit

Werden zwei Datensätze miteinander fusioniert, die identische Inhalte besitzen, so tritt häufig das Problem auf, daß eigentlich identische Geometrien (z.B. Straßenverläufe) lagemäßig gegeneinander verschoben sind. Zur Herstellung der Identität sind aufwendige Verfahren zur *Homogenisierung* notwendig. Die zu beobachtende Verschiebung ist aber nichts anderes als ein Effekt der in den Daten enthaltenen geometrischen Genauigkeit. Jeder Datensatz repräsentiert demnach genau eine Realisierung von theoretisch unendlich vielen möglichen Darstellungen. Erfassungseinflüsse, aber auch die Natur der Objekte bewirken, daß keine Realisierung der anderen gleicht (Caspary, 1992; Glemser, 1994). Die geometrische Genauigkeit äußert sich demnach als eine Art Variation in der Geometrie. Auch nach einer erfolgten Homogenisierung, sind die *homogenen Objekte* nicht übergeordnet genau. Sie repräsentieren nur eine bestimmte Realisierung (z.B. die mittlere Geometrie). Ein großes Problem ist dadurch gegeben, daß man einer Realisierung allein nicht ansehen kann, welchen Genauigkeitswert sie besitzt. Nur über den Weg von Mehrfacherfassungen oder aus bereits vorhandenem Wissen kann die geometrische Genauigkeit ermittelt werden. Die bekannten Werte sind dann zum Zwecke der weiteren Nutzung explizit und individuell für jedes Objekt zu modellieren.

Die geometrische Genauigkeit bildet die Ursache für die Entstehung von unsicheren Relationen zwischen den Objekten. Unter Objektrelationen werden Nachbarschaftsbeziehungen verstanden, die häufig in GIS genutzt werden, um selektive Anfragen an den Datenbestand auszuführen. Z.B. könnten alle Objekte gesucht sein, die ein ausgewähltes Objekt *berühren*. Diese Art von Anfragen sind als *topologische Operationen* bekannt. Da aufgrund der geometrische Genauigkeit der Rand eines Objektes als variabel aufzufassen ist, fallen die Beziehungen nicht mehr eindeutig aus. Sie sind ebenfalls unsicher. Man kann diese Art der Unsicherheit als *topologische Genauigkeit* der Relationen bezeichnen. Ihr Entstehen und Ausmaß ist alleinig von der geometrischen Genauigkeit abhängig. Um eine Bewertung der unsicheren Relationen durchzuführen, ist daher auf die Auswirkungen der geometrischen Genauigkeit auf die Beziehungen der Objekte einzugehen.

2.2 Modelldefinition

Wie in der vorangegangenen Diskussion erarbeitet, sind die geometrischen Primitive (Punkte, Linien, Flächen) nur zu einem bestimmten Grade genau. Die Genauigkeit äußert sich durch eine Variation in der Lage und Ausdehnung, die für die einzelnen Objekte möglich sind. Diese räumliche Variation gilt es zu beschreiben. Als erstes soll eine Eingrenzung erfolgen. Die Eingrenzung erfüllt die Aufgabe, den Bereich des Vorkommens möglicher Realisierungen der Geometrie räumlich einzuschränken. Sie äußert sich im allgemeinen in Form eines Bandes um die geometrischen Primitive und entspricht damit dem bekannten *Toleranzband* (Perkal, 1956; Chrisman, 1982). Das Band beinhaltet alle möglichen Realisierungen und somit auch die *wahre* Position der Primitive (sofern eine solche überhaupt existiert). Daraus kann einerseits eine maximal mögliche Ausdehnung generiert werden, die als *Maximum-Geometrie* (kurz: *Maximum*) zu bezeichnen ist. Außerhalb des Maximums existieren keine weiteren Realisierungen der Objektgeometrie. Andererseits existiert für flächenhafte Objekte zusätzlich eine minimale Ausdehnung, die *Minimum-Geometrie* (kurz: *Minimum*). Für sie gilt, daß die umschlossene Ausdehnung als sicher anzunehmen ist. Des weiteren kann über eine Schätzung einer mittleren Ausdehnung des Objektes eine *mittlere Geometrie* (kurz: *Mittel*) formuliert werden. Bestehende Datensätze enthalten jeweils nur eine geometrische Beschreibung pro Objekt, die aber in einer guten Näherung dem Mittel gleicht. Aus diesem Grund ist das Mittel als *traditionelle Geometrie* (kurz: *Normalgeometrie*) aufzufassen. Zur Vollständigkeit wird festgelegt, daß im Falle von punkt- und linienförmigen Objekten Minimum und Mittel identisch sind. Formal beschreibt das Modell die Geometrie des Objektes durch das 3-Tupel:

$$Geometrie = \{Maximum, Mittel, Minimum\}$$

Abb. 1 veranschaulicht die getroffenen Festlegungen an einem flächenhaften Beispielobjekt.

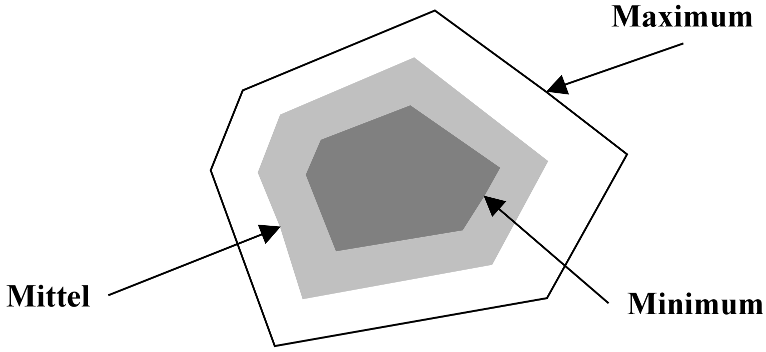
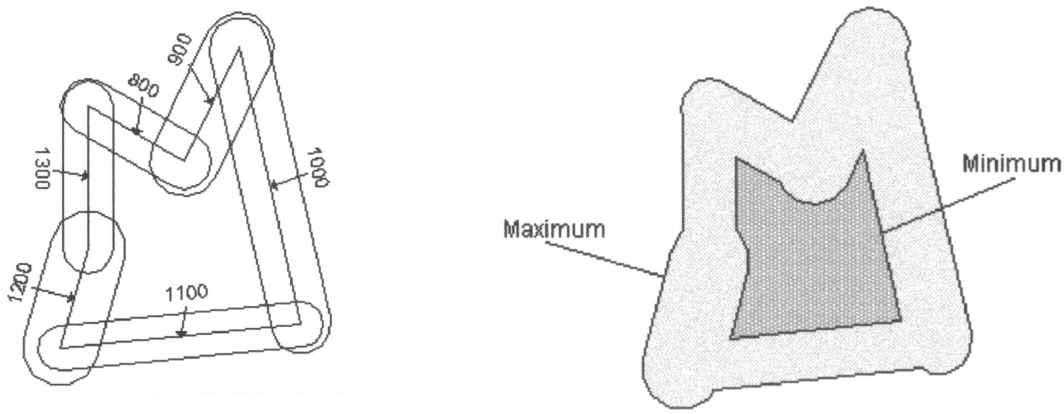


Abb. 1: Darstellung des Minimums, des Mittels und des Maximums eines Flächenobjektes

Die Anwendung des Modells setzt die individuelle Erfassung des Minimums, des Mittels und des Maximums für jedes Objekt voraus. Eine direkte Erfassung dieser Geometrien ist zwar theoretisch denkbar, doch praktisch kaum zu realisieren. Ein besserer Weg ergibt sich durch Ergänzung der seither üblichen Vorgehensweise mit Erfassung der Normalgeometrie um die Angabe eines Epsilon-Wertes (ε) zu jedem Primitiv. Dieser Wert veranschaulicht die Breite des Toleranzbandes und kann zur Bildung von Minimum und Maximum eingesetzt werden. Dazu werden zunächst Puffer mit Breite ε um alle Primitive gebildet (Abb. 2(a)). Die Vereinigung der Puffer läßt das Toleranzband entstehen. Im Falle von punkt- oder linienförmigen Objekten ist damit das Maximum erzeugt. Bei flächenhaften Objekten entsteht das Maximum durch eine zusätzliche Vereinigung des Bandes mit der Normalgeometrie, während das Minimum durch Differenzbildung von Normalgeometrie und Band zu erzeugen ist (Abb. 2(b)).



(a) (b)

Abb. 2: Pufferbildung mit ε um die Randlinien eines flächenhaften Objektes (a) zur Erzeugung von Minimum und Maximum (b)

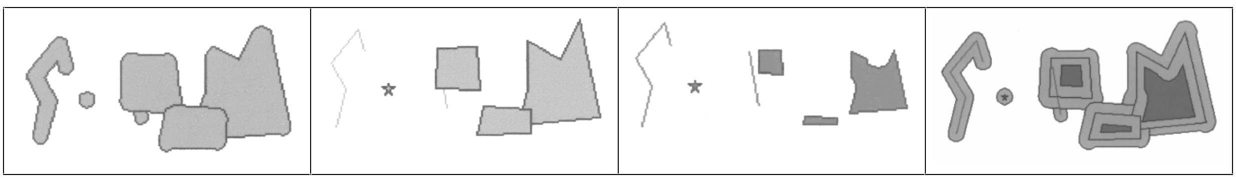
Der Vorteil dieses Modells liegt in seiner einfachen Anwendbarkeit. Die bisherige Arbeitsweise mit Erfassung der Normalgeometrie kann weiterhin beibehalten werden. Lediglich eine Ergänzung um die Erzeugung von Minimum und Maximum hat zu erfolgen. Wenn diese durch die Angabe eines Epsilon-Wertes erfolgt, dann kann deren Berechnung weitgehend automatisiert durchgeführt werden (Schneider, 1998). Ein Nachteil stellt das Fehlen einer Angabe zur Verteilung der Realisierungen innerhalb des Toleranzbandes dar. Alle möglichen Ausdehnungen eines Objektes sind gleich wahrscheinlich. Aus diesem Grund lassen sich nur *qualitative* Bewertungen zur Genauigkeit einer bestimmten Aussage (z.B. Bewertung des bekannten *Punkt-im-Polygon-Problems* (Abschnitt 4.1)) ableiten. Hierin ist der entscheidende Unterschied zu einem *stochastischen Ansatz* zu sehen, der explizit die Verteilung der Realisierungen miteinbezieht (Glemser, 1996).

3 Integrationsaspekte

Die Integration des Genauigkeitsmodells erfordert Erweiterungen in allen Komponenten (Erfassung, Verwaltung, Analyse und Präsentation) eines GIS (Glemser, 1994). Zunächst ist die Einbindung in das Datenmodell zu klären. Generell bieten sich zwei unterschiedliche Möglichkeiten. Wird Minimum und Maximum über den Epsilon-Wert gebildet, so bietet sich an, den Wert direkt als zusätzliches Attribut getrennt für jedes Primitiv zu verwalten (*Attributansatz*). Ein graphischer Weg eröffnet sich, indem man das 3-Tupel für jedes Objekt herstellt und dann Minimum, Mittel und Maximum ebenenweise als geometrische Objektbestandteile verwaltet (*graphischer Ansatz*). In der Regel dürften beide Möglichkeiten in allen gängigen GIS-Systemen unter Verwendung zusätzlicher Verweistabellen zu realisieren sein. Für die hier erfolgten Untersuchungen wurde das Modell im GIS-Paket MapInfo umgesetzt (Schneider, 1998). Der Attributansatz besitzt den Nachteil, daß Maximum und Minimum erst aufwendig (i.d.R. mehrfach) berechnet werden müssen. Dagegen verhält sich der Attributansatz im Vergleich zum graphischen Ansatz äußerst speicherplatzsparend.

Wie bereits im vorigen Abschnitt angedeutet ist eine praktikable Erfassungsweise durch Angabe eines Epsilon-Wertes möglich. Um einen solchen Wert anzugeben, bedarf es der Durchführung von Tests, innerhalb denen durch Auswertung von Mehrfacherfassungen, Wissen über die Variation der Realisierung erarbeitet wird. Die Tests sollten beispielhafte Objekte aller relevanter Objektarten umfassen.

Bei der Visualisierung der Daten im System sollte sich zunächst für den Nutzer keine Änderung ergeben. Der Nutzer arbeitet hauptsächlich mit der Darstellung der Normalgeometrie. Das Genauigkeitsmodell bleibt im Hintergrund. Das System reagiert dagegen auf Anfrage (*Wie genau sind die dargestellten Objekte?*) und zeigt dann zusätzlich Minimum und Maximum der Objekte (einzeln oder gemeinsam) an. Die Darstellung von verschiedenen Objekttypen (Punkte, Linien, Flächen) ist in Abb. 3 zu sehen.



(a) (b) (c) (d)

Abb. 3: Beispieldatensatz mit Darstellung des Maximum (a), des Mittels (b), des Minimums (c) und einer Gesamtansicht (d)

Auf die Auswirkungen der Integration der geometrischen Genauigkeit wird vertiefend im anschließenden Abschnitt eingegangen.

4 Anwendung

Für Analysen bedeutet die Integration des Genauigkeitsmodells, daß am Ende das Ergebnis mit einer Bewertung der Qualität versehen werden kann. Dazu ist es notwendig die angewendeten Analysefunktionen so zu erweitern, daß sie mit den Daten auch die Genauigkeit verarbeiten können. Die Verarbeitung der Genauigkeit entspricht dabei einer Übertragung oder *Fortpflanzung* der Genauigkeit auf das Ergebnis. Der Nutzer erhält somit nicht nur das Ergebnis, sondern stets auch ein Bewertung der Güte. Die Beurteilung bleibt aber weiterhin dem Nutzer überlassen. Jedoch werden nun Vergleiche zwischen unabhängig voneinander erzielten Ergebnisse ermöglicht. Im folgenden werden die Auswirkungen der Integration an den in GIS häufig genutzten Analysen des Punkt-im-Polygon-Test, der geometrischen Verschneidung und der Bestimmung der topologischen Relation erläutert.

4.1 Punkt-im-Polyon-Test

Der Punkt-im-Polygon-Test stellt fest, ob sich ein bestimmter Punkt innerhalb eines flächenhaften Objekts befindet oder nicht. Es existieren eine Reihe von Algorithmen, die dieses Problem lösen (Bill, 1996). Sie liefern jeweils eine eindeutige Antwort: *innerhalb* oder *außerhalb*. Der Rand wird dabei meist zum Inneren gezählt. Durch die Beachtung der geometrischen Genauigkeit geht die Eindeutigkeit verloren. Dies trifft insbesondere dann zu, wenn sich der Punkt innerhalb des Toleranzbandes befindet. Das Ergebnis des Tests wird unsicher.

Zur Bewertung der Unsicherheit wird die übliche Bestimmung dahingehend erweitert, daß der Test für alle Geometrien (Minimum, Mittel, Maximum) ausgeführt wird. Das Ergebnis für das Mittel entspricht der traditionellen Lösung. Man erhält dafür als Aussage *Punkt innerhalb* oder *Punkt außerhalb* des Objektes. Durch die weiteren Geometrien erfolgt nun zusätzlich eine Bewertung der Unsicherheit der traditionellen Aussage. Dazu wird diese Aussage um die Begriffe *sicher* und *möglicherweise* ergänzt (Blakemore, 1984). Liegt z.B. der Punkt außerhalb der Normalgeometrie, aber noch innerhalb des Maximums, dann gilt für das Testergebnis: *Punkt liegt möglicherweise außerhalb*. Im mathematischen Sinne erweitert sich der Ergebnisraum E auf vier mögliche Fälle:

$$E = \{sicher\ außerhalb, möglicherweise\ außerhalb, möglicherweise\ innerhalb, sicher\ innerhalb\}$$

Tab. 1 beinhaltet ein Bewertungsschema, das alle möglichen Bewertungsfälle abdeckt. Abb. 4 verdeutlicht schematisch die verschiedenen Aussagemöglichkeiten.

Punktlage	*Bewertung*
$Punkt \in Minimum$	*Sicher* innerhalb
$Punkt \in Mittel \quad \cap \quad Punkt \notin Minimum$	*Möglicherweise* innerhalb
$Punkt \in Maximum \quad \cap \quad Punkt \notin Mittel$	*Möglicherweise* außerhalb
$Punkt \notin Maximum$	*Sicher* außerhalb

Tab. 1: Bewertungsschema beim Punkt-im-Polygon-Test

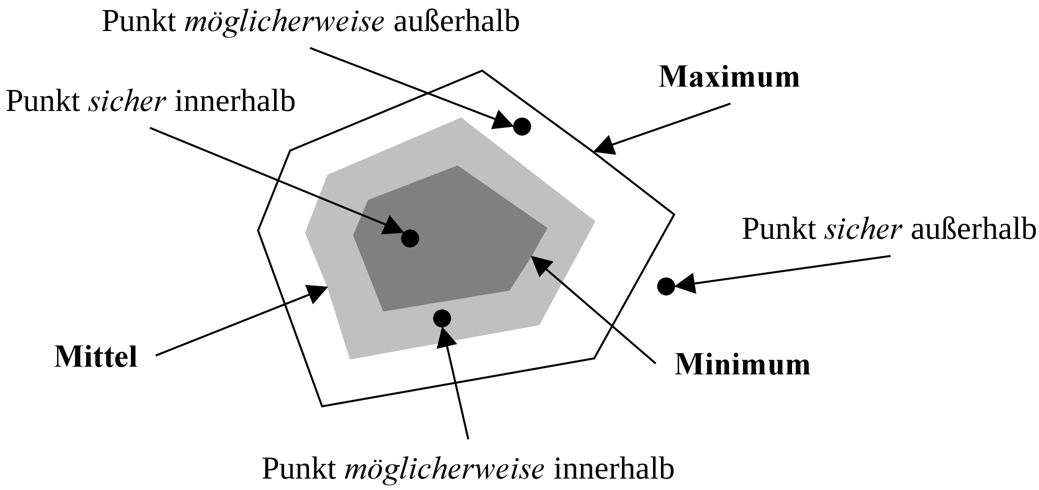


Abb. 4: Verschiedene Bewertungen der Punkt-im-Polygon-Aussage

4.2 Geometrische Verschneidung

Die geometrische Verschneidung hat die Aufgabe die geometrischen Formen der Objekte (Punkte, Linien, Flächen) miteinander zu verschneiden, um aus den sich überlappenden Teilen neue Objekte (Schnittobjekte) zu bilden (Bill, 1996). Sie entspricht damit im mengentheoretischen Sinn einer *logischen UND-Operation*. Diese Funktionalität gehört ebenfalls zu den elementaren Basisfunktionen und ist nahezu in allen GIS-Systemen vorhanden. Da als Ergebnisse wiederum Objekte entstehen, besteht die Fortpflanzung der geometrischen Genauigkeit in einer Berechnung von Minimum und Maximum des Schnittobjekts.

Die Erweiterung äußert sich zunächst in einer Untersuchung mit welcher Sicherheit ein Schnitt vorliegen kann (Schnittprüfung). In einem zweiten Schritt werden dann die neuen Objekte gebildet. Die Untersuchung startet mit der ebenenweisen Verschneidung der drei Geometrien der beteiligten Objekte A und B: Minimum A mit Minimum B, Mittel A mit Mittel B, Maximum A mit Maximum B. Nicht alle Schnittbildungen müssen eine Ergebnismenge erzielen. Je nach nichtleerer Ergebnismenge kann festgestellt werden, wie sicher ein Schnitt vorliegt. Wiederum erfolgt die Bewertung durch das Hinzufügen der Begriffe *sicher* und *möglicherweise*. Tab. 2 stellt die verschiedenen Bewertungsmöglichkeiten in Form eines Schemas zusammen.

Schnitt von			
Minimum	Mittel	Maximum	Bewertung
$E = O$	$E = O$	$E = O$	Objekte schneiden sich *sicher nicht*
$E = \emptyset$	$E = O$	$E = O$	Objekte schneiden sich *möglicherweise nicht*
$E = \emptyset$	$E = \emptyset$	$E = O$	Objekte schneiden sich *möglicherweise*
$E = \emptyset$	$E = \emptyset$	$E = \emptyset$	Objekte schneiden sich *sicher*

Tab. 2: Bewertung des Schnitts ($E = O$- leere Ergebnismenge, $E = \emptyset$- nichtleere Ergebnismenge)

An dieser Stelle bietet sich dem Nutzer nochmals die Gelegenheit die Erzeugung von Schnittobjekten zu steuern, indem er angibt, welche Objekte er konkret erhalten möchte. Dazu hat er sich zu entscheiden, wie sicher seine Schnittobjekte sein sollen. Beispielsweise kann er wählen, daß nur sichere Objekte, d.h. die durch einen sicheren Schnitt der Eingangsobjekte entstehen, erzeugt werden sollen. Dieses Kriterium ist dann für alle Objektschnitte zu prüfen.

Nach erfolgreicher Prüfung ist das vollständige Geometrietupel des Schnittobjekts herzustellen. Es setzt sich aus den Schnitten der jeweiligen Ebenen, die zur Schnittprüfung bereits erzeugt wurden, zusammen. Liegt innerhalb einer Ebene (z.B. für das Minimum) eine leere Ergebnismenge vor, so ist diese durch einen punktförmigen Repräsentanten (z.B. den Schwerpunkt einer anderen Geometrie) zu füllen. Mehrere Verschneidungsbeispiele sind in Abb. 5 abgebildet.

Abb. 5: Ergebnis der geometrischen Verschneidung des Beispieldatensatzes

Die Berücksichtigung der geometrischen Genauigkeit bedeutet eine Verbesserung des Verschneidungsergebnisses gegenüber der traditionellen Lösung. Dies wird daran deutlich, daß auch solche Schnittobjekte gebildet werden können, die sich nur in den Maximum-Geometrien der Eingangsobjekte überlappen. Ein solches Ergebnis kann im traditionellen Ansatz nicht erzielt werden. Es ist aber besonders für kritische Analysen, bei denen z.B. mit Sicherheit das gemeinsame Auftreten zweier Eigenschaften ausgeschlossen werden soll, von essentieller Bedeutung.

4.3 Bewertung topologischer Relationen

Bei der Herstellung von topologischen Relationen zwischen Objekten ist die gegenseitige Lage zueinander zu betrachten. Dabei sind die exakten metrischen Verhältnisse unwichtig; die relative Lage ist entscheidend. Folgende mögliche Relationen können unterschieden werden (Abb. 6):

- Getrennt (Disjoint)
- Berührend (Meet)
- Überlappend (Intersect)
- Enthalten (Contains)
- Enthält (Inside)
- Berührend und enthalten (Covers)
- Berührend und enthält (Covered By)
- Identisch (Equal)

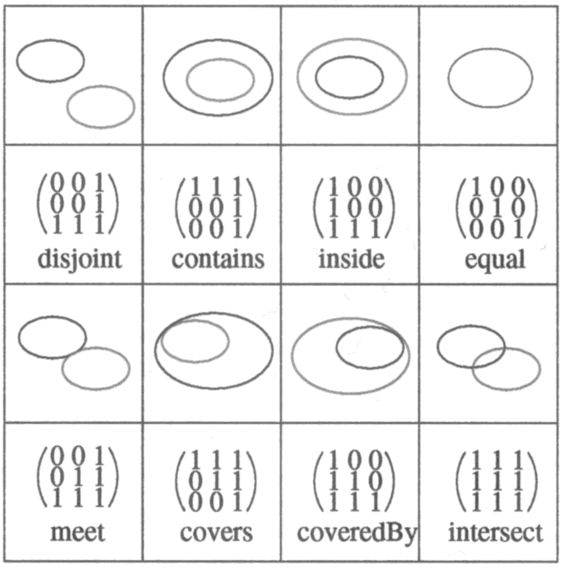


Abb. 6: Topologische Relationen und deren mathematische Repräsentation im Relationenmodell mit Hilfe von Matrizen

Es existiert ein mathematisch begründetes Relationenmodell, das sich zur Beschreibung der Relationen eignet (Egenhofer, Herring, 1990). Dafür wird zwischen den Objektbestandteilen *Inneres*, *Rand* und *Äußeres* unterschieden. Bildet man alle möglichen Schnitte der Bestandteile zweier Objekte erhält man eine eindeutige Aussage, welche Relation vorliegt. Bei der Schnittbildung ist die konkrete geometrische Form des jeweiligen Schnitts unbedeutend. Lediglich die Existenz ist von Interesse. Die Schnittergebnisse lassen in übersichtlicher Form innerhalb einer Matrix (*Intersectionmatrix*) anordnen. Jede Relation verfügt über ein ganz bestimmte und daher eindeutige Matrixzusammensetzung, die somit zur Zuweisung der vorliegenden Relation genutzt werden kann (Abb. 6). Beziehungen zwischen den Relationen lassen sich durch ein Relationen-Graphen (Conceptual-Neighborhood-Graph) ausdrücken (Abb. 7) (Egenhofer, Al-Taha, 1992; Clementini, Di Felice, 1996;Winter, 1996). Er beschreibt den Übergang von Relationen, wenn Änderungen in der Geometrie angebracht werden und veranschaulicht so die Nachbarschaften zwischen den Relationen.

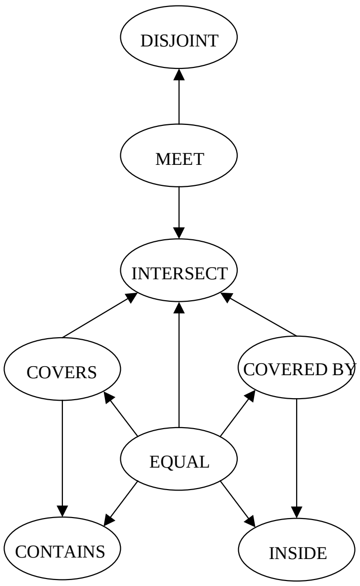


Abb. 7: Relationengraph (Conceptual-Neighborhood-Graph) nach Winter (1996)

Aus dem Graph kann abgeleitet werden, daß einige Relationen *stabil*, d.h. während Änderung über längere Zeit bestand haben und andere dagegen *labil* sind, d.h. sofort bei Änderung in andere übergehen. Bei den labilen Relationen handelt es sich um diejenigen, die aufgrund der speziellen Betrachtung des Randes vorliegen. Dazu zählen: *Meet, Covers, Covered By,* und *Equal*. Da bei Berücksichtigung der geometrischen Genauigkeit der Rand stets unsicher ist, werden diese Relationen ebenfalls entsprechend unsicher vorliegen. Aus diesem Grund konzentrieren sich die weiteren Betrachtungen zunächst auf die stabilen Relationen: *Disjoint*, *Intersect*, *Contains* und *Inside*. Um eine einfachere Bewertung der vorliegenden Relationen zu ermöglichen, wird der Relationengraph in zwei Teile aufgetrennt (Winter, 1996). Die Teilung erfolgt zentral durch die Relation *Intersect*, die die Überlappung der Objekte kennzeichnet. Aufgrund des Grades der Überlappung wird diese Relation aufgeteilt in eine *schwache* Überlappung (*Weak Intersect*) und eine *starke* Überlappung (*Strong Intersect*). Es entstehen die Relationengruppen (*Relationencluster*) C_1 und C_2 mit:

$$C_1 = \{Weak\ Intersect, Disjoint\}$$
$$C_2 = \{Strong\ Intersect, Contains, Inside\}$$

Zur Entscheidung kann der Überlappungsfaktor *OF* genutzt werden:

$$OF(A,B) = \frac{Fläche(Norm_A \cap Norm_B)}{\min(Fläche(Norm_A), Fläche(Norm_B))}$$

Norm bezeichnet die Normalgeometrie der Objekte. Welche Gruppe gültig ist, definiert sich durch:

$$\begin{aligned} OF \le 0.5: \quad & C_1 \\ OF > 0.5: \quad & C_2 \end{aligned}$$

Die Festlegung einer Gruppe hat den Sinn, daß nur die Relationen dieser Gruppe zu bewerten sind. Für alle anderen Relationen kann pauschal angenommen werden, daß diese *sicher nicht* zutreffen werden. Damit ist stets nur die geringste Anzahl an Relationen wirklich zu bewerten.

Die Vorgehensweise bei der Bewertung der Unsicherheit von Relationen unterteilt sich in zwei Hauptschritte (Krauß, 1998). Im ersten Schritt werden die exakten Relationen für alle möglichen Kombinationen (insgesamt 9) der beiden Objektgeometrien ermittelt und die Ergebnisse in einer *Relationenmatrix* zusammengefaßt. Dazu gilt eine Kodierung der ermittelten Relation in folgender Form:

$$r_{rel}(A,B) \quad = \quad \left\{ \begin{bmatrix} 0 = DISJOINT(A,B) \\ 1 = INTERSECT(A,B) \\ 2 = CONTAINS(A,B) \\ 3 = INSIDE(A,B) \end{bmatrix} \right.$$

Die Relationenmatrix besitzt folgendes Aussehen:

$$R^9 = \begin{pmatrix} r_{rel}(Max_A, Max_B) & r_{rel}(Norm_A, Max_B) & r_{rel}(Min_A, Max_B) \\ r_{rel}(Max_A, Norm_B) & r_{rel}(Norm_A, Norm_B) & r_{rel}(Min_A, Norm_B) \\ r_{rel}(Max_A, Min_B) & r_{rel}(Norm_A, Min_B) & r_{rel}(Min_A, Min_B) \end{pmatrix}$$

mit *Max* als Maximum, *Norm* als Mittel und *Min* als Minimum. Theoretisch können insgesamt $n = 4^9$ verschiedene Matrizen entstehen, von denen aber nur $\overline{n} = 400$ in der Realität vorkommen. Ein Beispiel für das Aussehen der Relationenmatrix ist der Abb. 8 zu entnehmen.

Abb. 8: Beispiel für die Bestimmung der Relationenmatrix

Der zweite Schritt nutzt die Relationenmatrix zur Bewertung der möglichen Relationen. Vor der eigentlichen Bewertung wird zunächst mit Hilfe des Überlappungsfaktors *OF* eine Entscheidung über das momentan gültige Relationencluster herbeigeführt. Damit sind die zu bewertenden Relationen festgelegt. Die Beschreibung der Unsicherheit erfolgt wiederum in qualitativer Weise durch Hinzufügen eines Wertungsprädikats. Hierzu eignen sich ebenfalls die bereits bekannten Begriffe: *sicher*, *möglicherweise*, *möglicherweise nicht* und *sicher nicht*. Allen Relationen ist nun eine Wertung zuzuweisen. Als Entscheidungskriterium ist die Häufigkeit n_{rel} des Vorkommens der jeweiligen Relation in der Relationenmatrix zu bestimmen. Die abschließende Bewertung folgt dem folgenden Schema:

$$\begin{aligned} n_{rel} &= 9: && \textit{sicher} \\ 6 \le n_{rel} &< 9: && \textit{möglicherweise} \\ 0 < n_{rel} &< 6: && \textit{möglicherweise nicht} \\ n_{rel} &= 0: && \textit{sicher nicht} \end{aligned}$$

Je ein Beispiel für die Bewertung innerhalb der beiden Relationengruppen sind in Abb. 9 bzw. in Abb. 10 zu sehen.

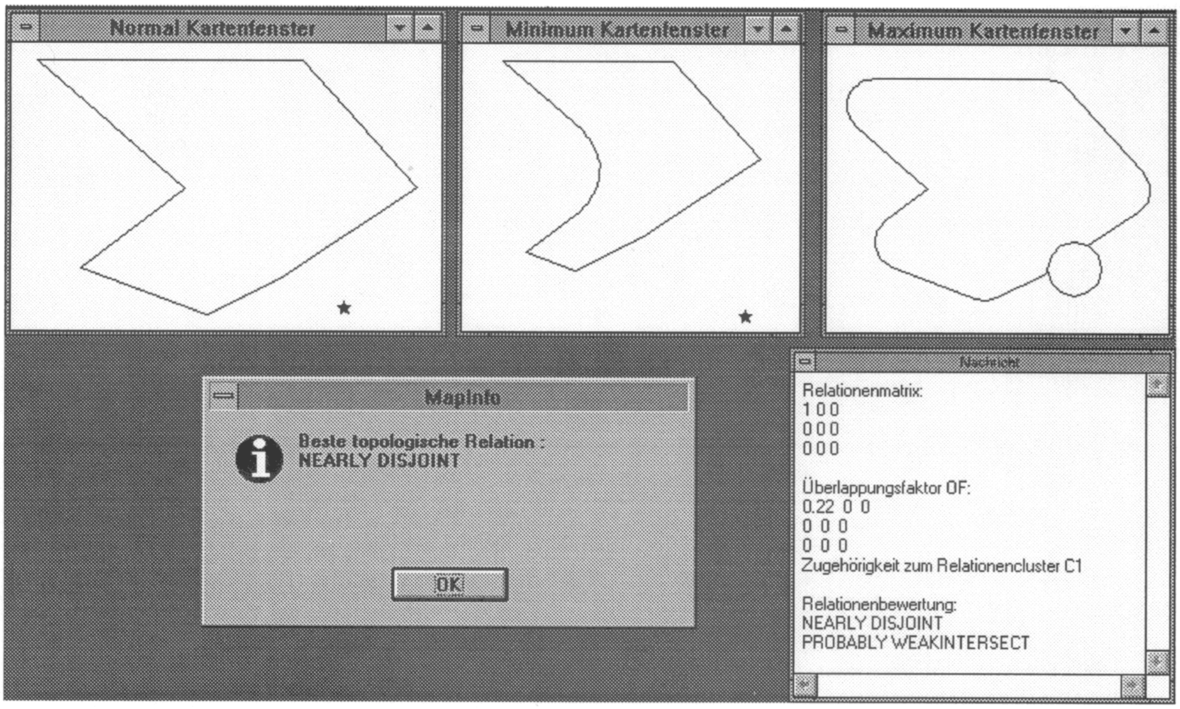


Abb. 9: Beispiel für die Bewertung von unsicheren Relationen mit *Disjoint* als beste Relation

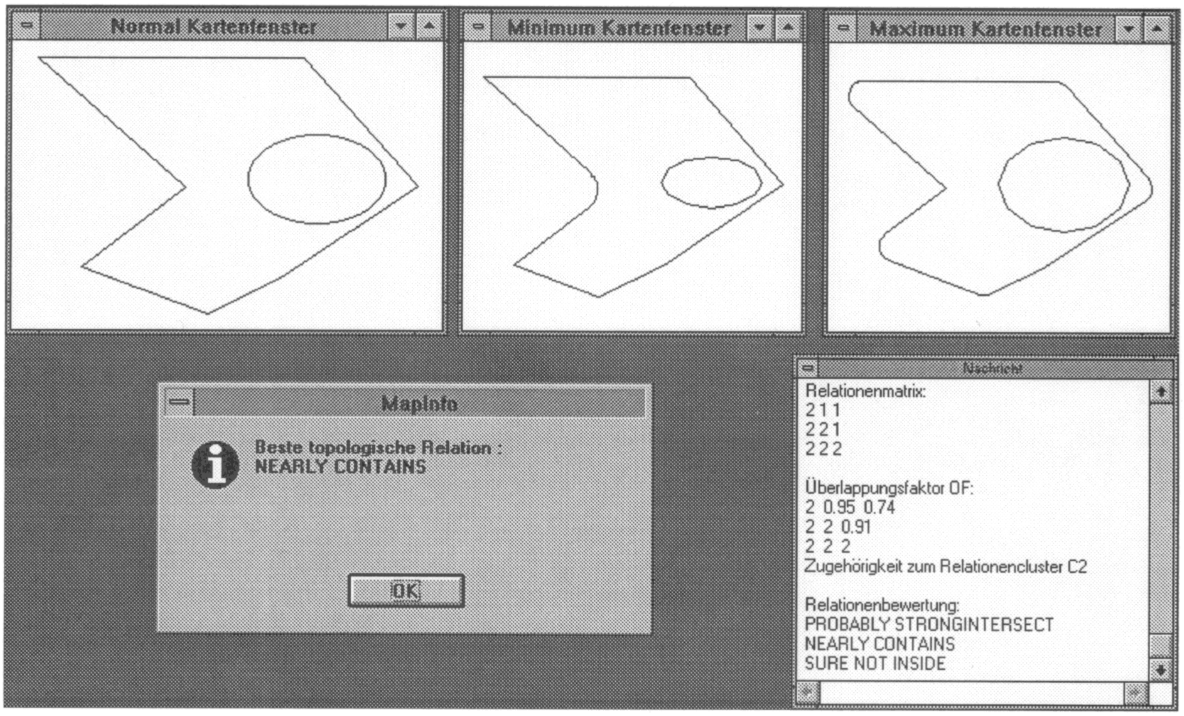


Abb. 10: Beispiel für die Bewertung von unsicheren Relationen mit *Contains* als beste Relation

Die unsichere Relation, für die die Häufigkeit n_{rel} maximal ist, kann als *beste Relation* bezeichnet werden. Wenn mehrere unsichere Relationen gleiche Häufigkeiten aufweisen, so sollte man sich immer für *Intersect* als beste Relation entscheiden, da diese am häufigsten auftritt.

Um nun auch eine Bewertung der labilen Relationen (*Meet*, *Covers*, *Covered By*, *Equal*) zu ermöglichen, sind die auftretenden Kombinationen von Relationen in der Relationenmatrix weitergehend zu analysieren. Dabei besitzt die Eigenschaft dieser Relationen, im Relationengraph zwischen mehreren stabilen Relationen zu liegen, besondere Bedeutung. Enthält die Relationenmatrix ein ausgeglichenes Verhältnis von Relationen (z.B. $n_{Intersect} = n_{Disjoint} = 4$), dann kann für die dazwischen liegende labile Relation (hier: *Meet*) ein Zutreffen angenommen werden. Da labile Relationen aber nur sehr kurzlebige Zustände darstellen, ist die Bewertung stets als *möglicherweise nicht* anzugeben. Dementsprechend werden labile Relationen nur dann als beste Relationen ausgewählt, wenn ihre Nachbarrelationen ebenfalls nur *möglicherweise nicht* zutreffen. Ein Beispiel für die Bewertung von labilen Relation ist in Abb. 11 gegeben.

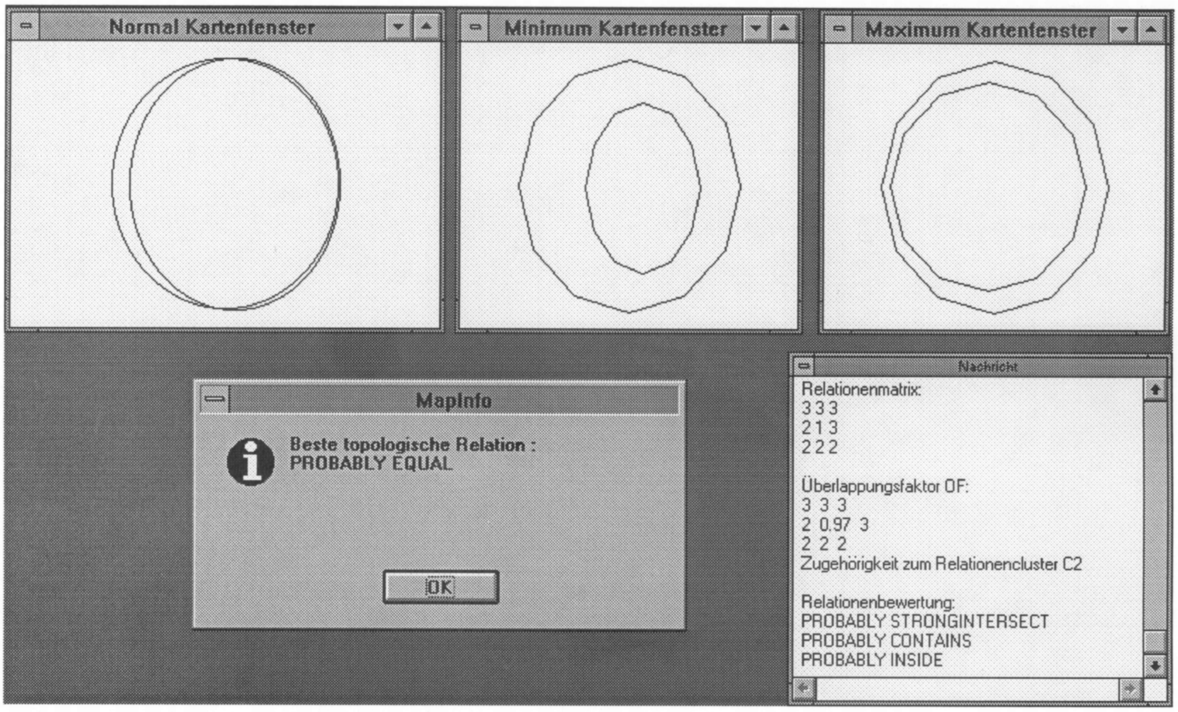


Abb. 11: Beispiel für die Bewertung von unsicheren Relationen mit *Equal* als beste Relation

5 Zusammenfassung

Der vorliegende Beitrag beschäftigt sich mit der Integration der geometrischen Genauigkeit in ein GIS. Die Integration erfordert die Definition eines Genauigkeitsmodells. Hierfür wurde das Minimum-Maximum-Modell formuliert, das die Variation in der Geometrie durch Festlegung von minimalen und maximalen Ausdehnungen eines Objektes eingrenzt. Neben verschiedenen Aspekten, die bei der Umsetzung des Modells in einem System zu beachten sind, konzentrierten sich die Ausführungen auf die Berücksichtigung der Genauigkeiten innerhalb von Analysen. Drei in GIS häufig verwendete Funktionen standen dabei im Mittelpunkt: der Punkt-im-Polygon-Test, die geometrische Verschneidung und die Festlegung von topologischen Relationen zwischen den Objekten. Für alle drei Funktionen wurden Verfahren entwikkelt, die eine Fortpflanzung der Genauigkeit auf das Ergebnis erlauben und so eine Bewertung der Analyseergebnisse durch den Nutzer ermöglichen. Da generell aber alle Funktionen über eine solche Fortpflanzungskomponente verfügen sollten, bleiben umfangreiche Aufgaben für zukünftige Arbeiten erhalten.

6 Literatur

Bill, R. 1996: Grundlagen der Geo-Informationssysteme. Band 2: Analysen, Anwendungen und neue Entwicklungen. Wichmann, Heidelberg.

Blakemore, M. (1984): Generalisation and Error in Spatial Data Bases. Cartographica, Vol. 21, 131-139, 1984.

Caspary, W. (1992): Genauigkeit als Qualitätsmerkmal digitaler Datenbestände. In: Grünreich, D., Buziek, G. (Hrsg.) (1992), Gewinnung von Basisdaten für Geo-Informationssysteme. DVW-Schriftenreihe, Heft 4, 157-166.

Chrisman, N.R. (1982): A Theory of Cartographic Error and ist Measurement in Digital Databases. Auto-Carto 5, Crystal City, 159-168.

Clementini, E., Di Felice, P. (1996): An Algebraic Model for Spatial Objects with Indeterminate Boundaries. In: Burrough & Frank (1996), Geographic Objects with Indeterminate Boundaries. ESF-GISDATA, Vol. 2, 155-169, Taylor & Francis, 1996.

Egenhofer, M.J., Herring, J.R. (1990): A Mathematical Framework for the Definition of Topological Relationships. Proceedings 4th International Symposium on Spatial Data Handling, International Geographical Union, Zürich, 803-813.

Egenhofer, M.J., Al-Taha, K.K. (1992): Reasoning about Gradual Changes of Topological Relationships. In: Frank, A.U., Campari, I., Formentini, U. (Hrsg.): Theories and Models of Spatio-Temporal Reasoning in Geographic Space. LNCS639, Springer, Berlin, 196-219.

Glemser, M. (1994): Behandlung der Genauigkeit räumlicher Daten in Geo-Informationssystemen. In: Die benutzte Erde. Alfred-Wegener-Stiftung (Hrsg.), Ernst&Sohn, Berlin.

Glemser, M. (1996): Integration der geometrischen Datenqualität in GIS-Funktionen. In: Proceedings des Workshops: Datenqualität und Metainformation in Geo-Informationssystemen, Universität Rostock.

Krauß, S. (1998): Qualitative Beschreibung von unsicheren topologischen Relationen innerhalb des Minimum-Maximum-Modells. Diplomarbeit am Institut für Photogrammetrie der Universität Stuttgart, http://www.ifp.uni-stuttgart.de/education/diplomarbeiten.

Schneider, C. (1998): Integration der geometrischen Genauigkeit in MapInfo nach dem Minimum-Maximum-Modell. Diplomarbeit am Institut für Photogrammetrie der Universität Stuttgart (unveröffentlicht).

Winter, S. (1996): Unsichere topologische Beziehungen zwischen ungenauen Flächen. Deutsche Geodätische Kommission, Reihe C, München, 1996.