

# Gebäudeextraktion aus digitalen Bildern

CAROLA BRAUN, Bonn

## Zusammenfassung

Es wird ein Konzept für ein Verfahren vorgestellt, das aus perspektiven Einzelbildern die dreidimensionale Form und Orientierung von Gebäuden mit Hilfe objektspezifischer Bedingungen ableitet. Die vollständige Beschreibung der Gebäude kann durch Zuordnung der Ergebnisse mehrerer Einzelbildauswertungen im Raum erfolgen. Aus den Bilddaten werden Hypothesen über parallele und senkrechte Kanten und Flächen am Objekt und andere Eigenschaften von Gebäuden generiert. Mit Hilfe dieser Hypothesen, die als Bedingungen eingeführt werden, wird automatisch schrittweise auf numerische Werte für die gewählten Attribute der Objektelemente im Raum geschlossen. Reichen die automatisch ermittelten Bedingungen nicht zur Rekonstruktion des Objekts im Raum, können weitere Bedingungen am Objekt interaktiv vom Operateur vorgegeben werden. Flexibilität und Anwendbarkeit des Verfahrens werden an mehreren Beispielen demonstriert.

## Abstract

A concept is presented for deriving the three dimensional shape and orientation of buildings in single perspective projections by using additional constraints. Matching the results of individual image interpretations in space supplies a complete description of the object. Hypothesis about parallel, perpendicular and other relationships are generated. By introducing these hypothesis as constraints, it is possible to automatically reason about numerical values of attributes of the object in space. If the automatically derived constraints are not sufficient to reconstruct the object in space, further relationships can be assigned by an operator. Examples demonstrate the flexibility and applicability of the approach.

## 1. EINFÜHRUNG

In vielen Bereichen wie Stadtplanung, Umweltschutz und Kulturgeographie werden zunehmend digitale und dreidimensionale Stadtmodelle benötigt, um den gestiegenen Anforderungen, die Aktualität und Vollständigkeit der vorliegenden Daten betreffen, gerecht zu werden. Simulierte Stadtmodelle werden bei Planungsaufgaben, vor allem in großräumigen Gebieten, eingesetzt. Beispiele hierfür sind etwa städtische Klimaanalysen für Stadtplanungen (Beckröge/Junius 1991). Dabei werden die Gebäude bezüglich ihrer Genauigkeit und Auflösung meistens als topographische Objekte betrachtet. Dies reicht auch für viele Anwendungen aus. Die Datenerfassung für dreidimensionale Stadtmodelle ist aber selbst bei einer photogrammetrischen Auswertung vor allem wegen des anfallenden Mehrumfanges gegenüber zweidimensionalen Modellen teuer. Die dreidimensionale Gebäudeerfassung zur Erstellung von Stadtmodellen mit den im Kataster üblichen Genauigkeitsanforderungen ist um ein Vielfaches aufwendiger, selbst wenn kein Bezug zu den Eigentums Grenzen hergestellt wird.

Da bei photogrammetrischen Auswertungen die dritte Dimension ohnehin anfällt, ist eine Leistungssteigerung bei der Erfassung der dreidimensionalen Gestalt von Gebäuden nur durch eine Rechnerunterstützung möglich. Sie setzt ein rechnerintern verwertbares Modell von Gebäuden voraus und kann sich bei Vorliegen digitaler oder digitalisierter Bilder auf Methoden des Bildverstehens und der Mustererkennung stützen.

Im folgenden wird ein interaktiv unterstütztes automatisches Verfahren vorgestellt, das aus digitalen Einzelbildern die dreidimensionale Form und Orientierung von Gebäuden bestimmt. Eine vollständige Beschreibung des Objekts kann durch Zuordnung der Ergebnisse aus mehreren

Einzelbildinterpretationen im Raum erhalten werden (Diemer 1992). Es stellt einen Baustein für ein Verfahren dar, das bei minimalen interaktiven Vorgaben dreidimensionale Stadtmodelle mit der Genauigkeit topographischer Objekte erfaßt und zur Lösung von Aufgaben in verschiedenen Fachbereichen dienen soll.

Bilder enthalten wesentlich mehr Informationen als zur Rekonstruktion der Szene benötigt werden. Daher wird mit Methoden der Bildverarbeitung eine symbolische Beschreibung aus dem Bild extrahiert, die auf die Informationen reduziert ist, die zur Lösung der Aufgabe benötigt werden. Eine häufig verwendete und wichtige symbolische Bildbeschreibung stellt die Strichzeichnung dar, da sie genügend Informationen für eine geometrische Rekonstruktion enthält. Menschen können ohne Schwierigkeiten aus einer einzelnen Strichzeichnung die Dreidimensionalität der abgebildeten Objekte erkennen. Daher liegt es nahe zu versuchen, den menschlichen Interpretationsprozeß auf einem Rechner nachzubilden, indem diejenigen Annahmen, die Menschen unwillkürlich beim Betrachten eines Bildes über die Objekte treffen, automatisch aus den Bildern abgeleitet und zur Rekonstruktion verwendet werden. Damit wird das bei Gebäuden schwierige Zuordnungsproblem bei der Stereoanalyse mehrerer Bilder vermieden, dafür muß das grundsätzlich unterbestimmte Problem der Einzelbildauswertung durch Nutzung eines spezifischen Objektmodells gelöst werden.

Viele Gebäude können durch Polyeder approximiert werden. Sie enthalten darüberhinaus parallele und senkrechte Kanten und Flächen, die horizontal und vertikal sein können und Dreiecke. Wie sich zeigen wird, genügen diese Eigenschaften und Beziehungen in vielen Fällen zur Lösung des unterbestimmten Problems der Gebäuderekonstruktion.

Zur Interpretation von Strichzeichnungen existieren viele Verfahren. Die ersten Verfahren basierten auf qualitativen Bildstrukturen. Eine der bekanntesten stellt die Klassifikation der Objektkanten dar, die Huffman und Clowes 1971 unabhängig voneinander entwickelten.

Zur eindeutigen Rekonstruktion der Objektform im Raum werden jedoch quantitative Größen wie Normalenvektoren der Oberflächen benötigt. Verfahren zur Ableitung dieser Größen aus einem oder mehreren Bildern sind im Computer Vision Bereich bekannt als 'shape from X', wobei X die Quelle der symbolischen Bildbeschreibung bezeichnet, die Textur, Schatten oder Stereo/Motion sein kann (Weissensee 1990). Für Objekte, die, wie in unserem Kontext, durch mathematisch einfach beschreibbare Flächen repräsentiert werden können, werden vor allem Techniken des 'shape from angle' und 'shape from perspective' eingesetzt.

Ziel des von Herman und Kanade vorgestellten MOSAIC-Systems ist der Aufbau und die schrittweise Verbesserung der dreidimensionalen Beschreibung einer kompletten städtischen Szene (Herman/Kanade 1984, 1986). Gebäude mit Flachdächern werden als Prismen modelliert. Das System umfaßt die Einzelbild- und Stereoanalyse und die Konstruktion des Szenenmodells und erlaubt die fotorealistische 3D-Visualisierung. Mulgaonkar und Shapiro haben ein PROLOG-basierendes Reasoning System zur Interpretation von Strichzeichnungen vorgestellt, das etwa 100 Regeln zur Inversen Perspektive enthält (Mulgaonkar/Shapiro 1985) und eine dreidimensionale Rekonstruktion von geraden und kreisförmigen Linien erlaubt. Sugihara hat wohl die theoretisch umfassendste Abhandlung zur geometrischen Interpretation von Strichzeichnungen präsentiert (Sugihara 1986). Das Verfahren liefert eine qualitative Beschreibung der Szene, die i. a. bis auf vier Freiheitsgrade bestimmt ist, kann jedoch keine fehlerhaften oder unvollständigen Strichzeichnungen bearbeiten. Ein System zur Erkennung und Lokalisierung polyederförmiger Objekte unter Verwendung von Bildauswertemethoden ist von Müller entwickelt worden (Müller 1992). Es basiert auf den Methoden von Sugihara und ist anwendbar bei der automatischen Handhabung von Werkstücken mit Hilfe sichtgestützter Roboter. Eine wissensbasierte Methode zur Rekonstruktion von Gebäuden aus digitalen Luftbildern wurde von Gülch vorgestellt (Gülch 1992). Aus mehreren Luftbildern werden, manuell unterstützt, Bildmerkmale (Polygone) extrahiert und mit Hilfe eines regelbasierten Systems, das ein allgemeines Objektmodell enthält, interpretiert.

Alle genannten Systeme können als Prototypen angesehen werden, die die manuelle Datenerfassung vor allem im Hinblick auf Flexibilität, Zuverlässigkeit und Genauigkeit nicht erreichen. Die Haupteinschränkungen werden durch das vereinfachte Objektmodell bzw. durch die Anforderungen an die Bildsegmentierung verursacht. Sie lassen vermuten, daß mittelfristig eine vollautomatische Datenerfassung komplexer Objekte, wie es Gebäude darstellen, nicht erreichbar ist.

Das hier vorgestellte Verfahren der Einzelbildinterpretation versucht im Kontext einer interaktiven Datenerfassung die Stärken der bekannten Verfahren zu nutzen. Es enthält daher

- ein für die topographische Datenerfassung genügend allgemeines Gebäudemodell,
- eine Repräsentation der Objekte als Flächenmodell, die eine fotorealistische Darstellung der Ergebnisse erlaubt und
- eine Strategie, die in der Lage ist, unvollständig und fehlerhaft aus den Bildern abgeleitete Strichzeichnungen zu bearbeiten.

Das Verfahren ist so konzipiert, daß es sowohl Perspektiv- als auch Parallelprojektionen verarbeiten kann, damit auch Bilder sehr schmalwinkiger Kameras interpretiert werden können. Es ist keine Vorinformation über die Orientierung des Objekts oder die Äußere Orientierung der Kamera notwendig, wenn es lediglich um die Form der Objekte geht.

## 2. MODELLBILDUNG

Die Automatisierung der Bildanalyse stellt sich bei näherer Betrachtung eine äußerste komplexe Aufgabe dar. Zunächst ist in einem Bild wesentlich mehr Information enthalten als zur Lösung der jeweiligen Aufgabe benötigt werden, insbesondere Objekte und Objektdetails die für die jeweilige Aufgabenstellung irrelevant sind, wie Bäume oder Fenster. Andererseits ist die Information grundsätzlich wegen Verdeckungen und der Unzulänglichkeit der Bildverarbeitungsverfahren unvollständig. Schon weil die Tiefeninformation beim Abbildungsprozeß verloren geht, ist es unmöglich ohne Wissen über das Objekt die dreidimensionale Gestalt aus Einzelbildern zu rekonstruieren.

Es gilt daher Einschränkungen in Form eines Modells festzulegen, um die Komplexität der Aufgabe zu reduzieren. Dieses Modell enthält eine genügend vereinfachte aber für die zu lösende Aufgabe hinreichend angepaßte Repräsentation des realen Objekts, des Bildes und des Auswerteprozesses (vgl. Abb. 1):

Als Ausgangsdaten liegt ein digitales oder digitalisiertes analoges Bild der realen Welt vor. Mit Methoden der Bildverarbeitung wird auf der Basis des Objekt- und des Kameramodells aus dem digitalen Bild eine symbolische Bildbeschreibung gewonnen. Sie enthält die Informationen, die zur Lösung der Aufgabe benötigt werden.

Um den Vorgang der Bildaufnahme mathematisch behandeln zu können, wird das Modell der Zentralprojektion verwendet, das der photographischen Abbildung und dem natürlichen Sehprozeß am besten nachgebildet ist. Die Umkehrung des Modells der Zentralprojektion ist die Inverse Perspektive. Da die Zentralprojektion keine umkehrbar eindeutige Abbildung ist, kann ein räumliches Objekt in der Regel nicht aus einer Perspektive rekonstruiert werden.

Das Modell der zu rekonstruierenden Objekte, in unserem Fall Gebäude, ist rein geometrisch und nimmt an, daß die Oberfläche aus mathematisch beschreibbaren Flächen besteht. Der einfachste Fall liegt vor, wenn es sich dabei um ebene Flächen handelt. Dies trifft auf Gebäude, allg. auf Körper zu, die sich als Polyeder approximieren lassen. Wegen Verdeckungen muß die zu rekonstruierende Fläche aber nicht geschlossen sein, also einen Polyeder bilden. Es ist aber ein generisches Modell, da die Struktur frei ist und die Komponenten parametrisiert sind.



Hypothesen über diese Eigenschaften und Beziehungen können automatisch aus den Bildaten ermittelt werden, wenn bestimmte Voraussetzungen erfüllt sind, oder interaktiv am Objekt vorgegeben werden.

Die wichtigsten Beziehungen zur Bestimmung der Objektform sind Inzidenzen zwischen Punkten und Ebenen.

Dieses Wissen über das Objekt wird bei der Ableitung der Bildbeschreibung genutzt. Das Bildmodell besteht daher aus einer zusammenhängenden Menge geschlossener oder offener Polygone und den zugehörigen Inzidenzrelationen. Die Bildbeschreibung enthält alle geometrischen und topologischen Eigenschaften des Polygonnetzes.

Das Objekt- und das Bildmodell werden vom Verfahren der Einzelbildinterpretation zur Ableitung einer dreidimensionalen Beschreibung des Objekts genutzt.

### 3. PRINZIP DER INVERSEN PERSPEKTIVE

Die Aufgabe der inversen Perspektive besteht in der Umkehrung der perspektiven Abbildung zur Bestimmung der räumlichen Form des abgebildeten Objekts aus dem Bild.

Da wir es nur mit einem Bild zu tun haben, liegt es nahe für die Zentralprojektion und damit für die Inverse Perspektive die Repräsentation der geometrischen Basiselemente Punkt, Gerade und Ebene und für die Abbildung selbst eine geeignete Repräsentation zu verwenden.

Die 2D Bildbeschreibung und 3D Objektbeschreibung des rekonstruierten Objekts werden in einem kartesischen  $xyz$ -Koordinatensystem  $S_K$  dargestellt. Der Ursprung des Systems liegt im Projektionszentrum und die negative  $z$ -Achse in Richtung der optischen Achse der Kamera (Abb. 2). Zur Lösung der Inversen Perspektive werden alle geometrischen Grundelemente im dreidimensionalen Raum repräsentiert. Durch Positionieren der Bildebene in die Szene, bezogen auf das Kamerasystem, wird auch das Bild als Objekt betrachtet. Die Bildebene sei identisch mit der Ebene  $z = -c$ , mit  $c$  als Kammerkonstante.

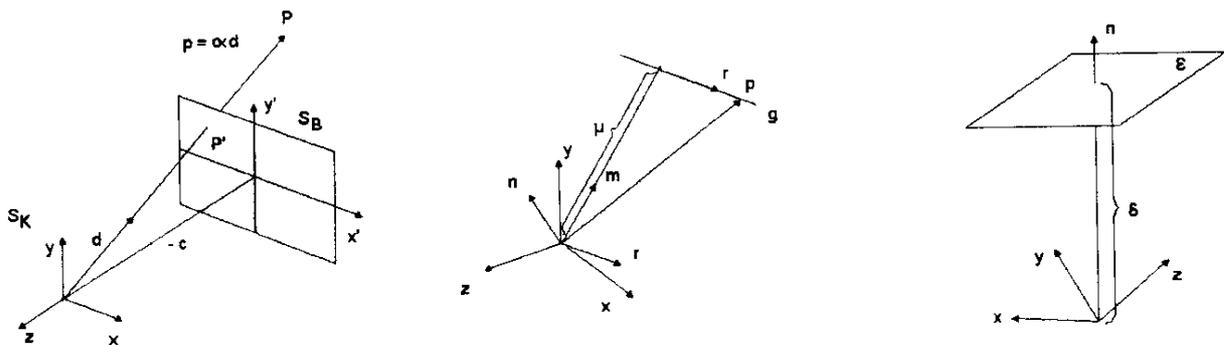


Abbildung 2: Definition des Kamerasystems und Repräsentation der Punkte, Geraden und Ebenen im Raum.

In der Bildebene ist ein  $(x',y')$ -Bildkoordinatensystem parallel zum  $(x,y)$ -System der Kamera definiert. Jeder Punkt auf der Bildebene wird durch einen Richtungsvektor  $\vec{d}' = (d'_1, d'_2, d'_3)^T$  der Länge 1 repräsentiert, der die Orientierung des Strahls von  $O$  zum entsprechenden Punkt  $P'$  bezeichnet. Der Richtungsvektor  $\vec{d}'$  läßt sich aus den Bildkoordinaten  $(x',y')$  und der Kammerkonstante  $c$  berechnen.

Punkte  $P$  im Objektraum werden in Polarkoordinaten durch einen Richtungsvektor und den Abstand repräsentiert. Damit sind drei unabhängige Parameter zur Bestimmung des Projektionsstrahls

$$P: \vec{p} = \alpha \vec{d}; \quad \vec{d}, \alpha \text{ fest}$$

erforderlich.

Entsprechend werden Geraden  $g$  mit der Normalen  $\vec{n} = (n_1, n_2, n_3)^T$  der projizierenden Ebene, dem Abstand  $\mu$  dieser Ebene vom Ursprung und dem Richtungsvektor  $\vec{r} = (r_1, r_2, r_3)^T$  der Geraden im Raum repräsentiert (Abb. 2):

$$g: \vec{p} = \mu (\vec{n} \times \vec{r}) + \beta \vec{r}; \quad \vec{n}, \vec{r}, \mu \text{ fest}, \quad \vec{n} \times \vec{r} = \mathbf{0}$$

Die Ebene lautet in Hessescher Normalform bei gegebener Normalen  $\vec{n} = (n_1, n_2, n_3)^T$  und Abstand  $\delta$ :

$$e: \vec{p}^T \vec{n} = \delta; \quad \vec{n}, \delta \text{ fest}$$

Die Perspektivische Abbildung lautet für Punkte und Geraden

$$P \rightarrow P': \vec{d}' = \vec{d},$$

$$g \rightarrow g': \vec{n}' = \vec{n}.$$

und postuliert die Invarianz von Projektionsstrahlen und -ebenen.

Die 2D Bildbeschreibung enthält die Information über die extrahierten Polygonzüge aus dem Bild, womit die Polygonpunkte  $P'$  mit den Richtungen  $\vec{d}'$ , die Polygonlinien  $g'$  mit ihren Anfangs- und Endpunkten und die Inzidenzrelationen zwischen Punkten und Flächen  $e'$  vorliegen.

Aufgabe der Inversen Perspektive ist es, aus der 2D Bildbeschreibung die unbekannt Attribute der Objektelemente im Raum zu bestimmen. Diese Attribute sind für einen Punkt der Abstand  $\alpha$ , für eine Gerade der Richtungsvektor  $\vec{r}$ , die Normale  $\vec{n}$  der projizierenden Ebene und der Abstand  $\mu$  und für die Ebene der Abstand  $\delta$  und die Normale  $\vec{n}$ .

Es gibt eine Reihe einfacher Schlußfolgerungsregeln zur Generierung von Hypothesen über Eigenschaften und Beziehungen der Elemente im Raum, etwa:

- Aus den Inzidenzrelationen im Bild kann auf die entsprechenden Inzidenzrelationen im Raum geschlossen werden.
- Wenn ein Knotenpunkt das Bild eines Dreibeins darstellt und bekannt ist, ob es sich um eine vor- oder zurückspringende Ecke handelt, können die Richtungsvektoren der abgebildeten Objektkanten im Raum berechnet werden (Kanatani 1990, Shomar et.al. 1992).
- Stellt eine Gruppe von Bildgeraden das Bild einer Parallelenschar dar, kann man aus dem Fluchtpunkt automatisch auf die Orientierung der zugehörigen parallelen Objektkanten im Raum geschlossen werden (Abb. 3) (Maggee/Aggarval 1984, Barnard 1982, Haralick 1989).

Aufgrund der Unterbestimmtheit der Aufgabe sind die mit Hilfe dieser Regeln neu ermittelten Eigenschaften und Beziehungen der geometrischen Elemente Hypothesen, da sie mit den Daten nur eines Bildes nicht verifiziert werden können. Sie können jedoch so lange zur Berechnung weiterer Parameter des Objekts genutzt werden, bis sie durch widersprüchliche Hypothesen verworfen werden.

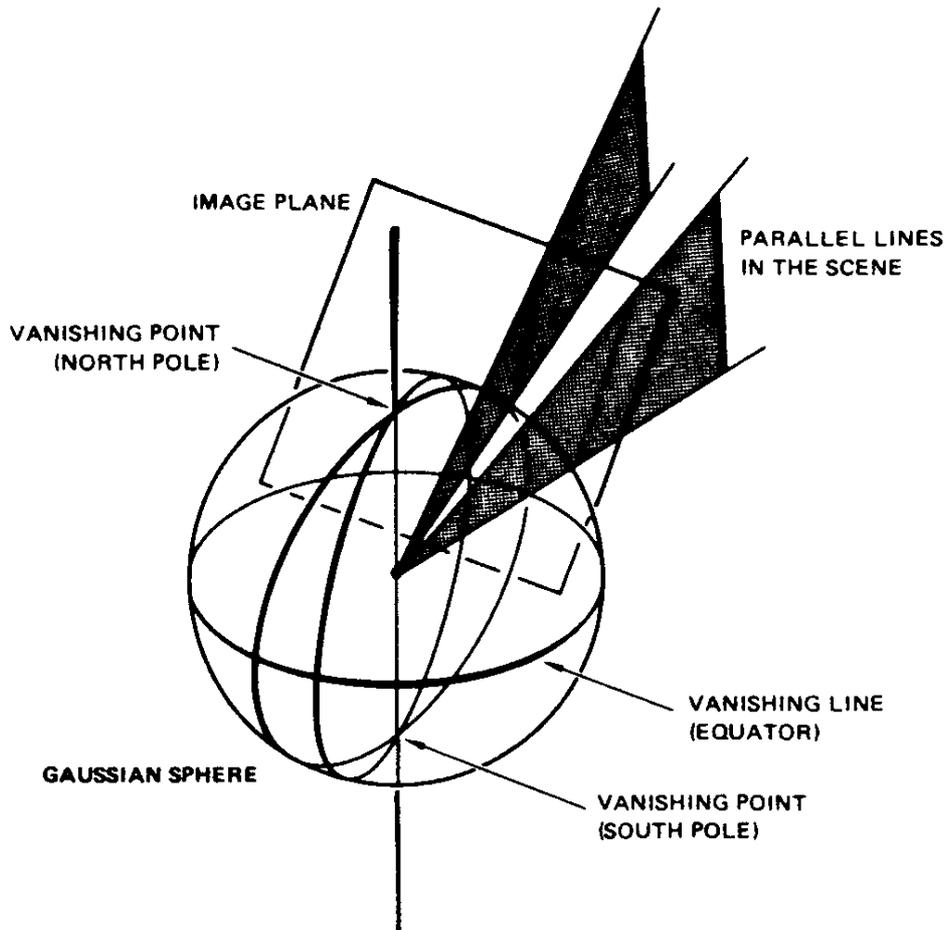


Abbildung 3: Die Richtungsvektoren zu Fluchtpunkten entsprechen der Orientierung der im Raum zugehörigen parallelen Geraden (aus Barnard 1982).

## 4. DAS VERFAHREN DER EINZELBILDINTERPRETATION

### 4.1 Ableitung der symbolischen Bildbeschreibung

Es gibt verschiedene Möglichkeiten für die Ableitung der 2D Bildbeschreibung aus einem Bild. Die Automatisierbarkeit ist abhängig von der Bildqualität, den Kontrastverhältnissen und dem Rauschen im Bild, von Objektverdeckungen und Schatten.

In Bildern mit schlechten Kontrastverhältnissen oder/und Verdeckungen wird die 2D Bildbeschreibung derzeit vollständig manuell erfaßt. Bei Bildern besserer Qualität erfolgt die automatische Ableitung der 2D Bildbeschreibung mit einem am Institut entwickelten Verfahren zur Kantenextraktion (Schickler 1992). Knotenpunktsbereiche werden derzeit noch ausgespart, da für ihre Rekonstruktion spezielles Modellwissen erforderlich ist. Durch näherungsweise Digitalisierung der abgebildeten Objektumrisse liegt die Beschreibung als Liste geschlossene Polygonzüge vor. Die Einpassung dieser Modellkanten auf die extrahierten Bildkanten liefert geschlossene Polygonzüge mit einer Genauigkeit im Subpixelbereich. Sind Polygonzüge offen, sind die Inzidenzen manuell vorzugeben.

Die vollautomatische Ableitung der Bildbeschreibung ist in Bildern mit guten Kontrastverhältnissen und mit künstlichen Objekten möglich, z.B. von Hand gezeichnete Skizzen oder bei optimalen Beleuchtungsverhältnissen aufgenommene Maschinenteile für Industrieanwendungen.

## 4.2 Verfahrensablauf

Der Ablauf des Verfahrens der Einzelbildinterpretation kann in mehrere Schritte unterteilt werden (Abb. 4).

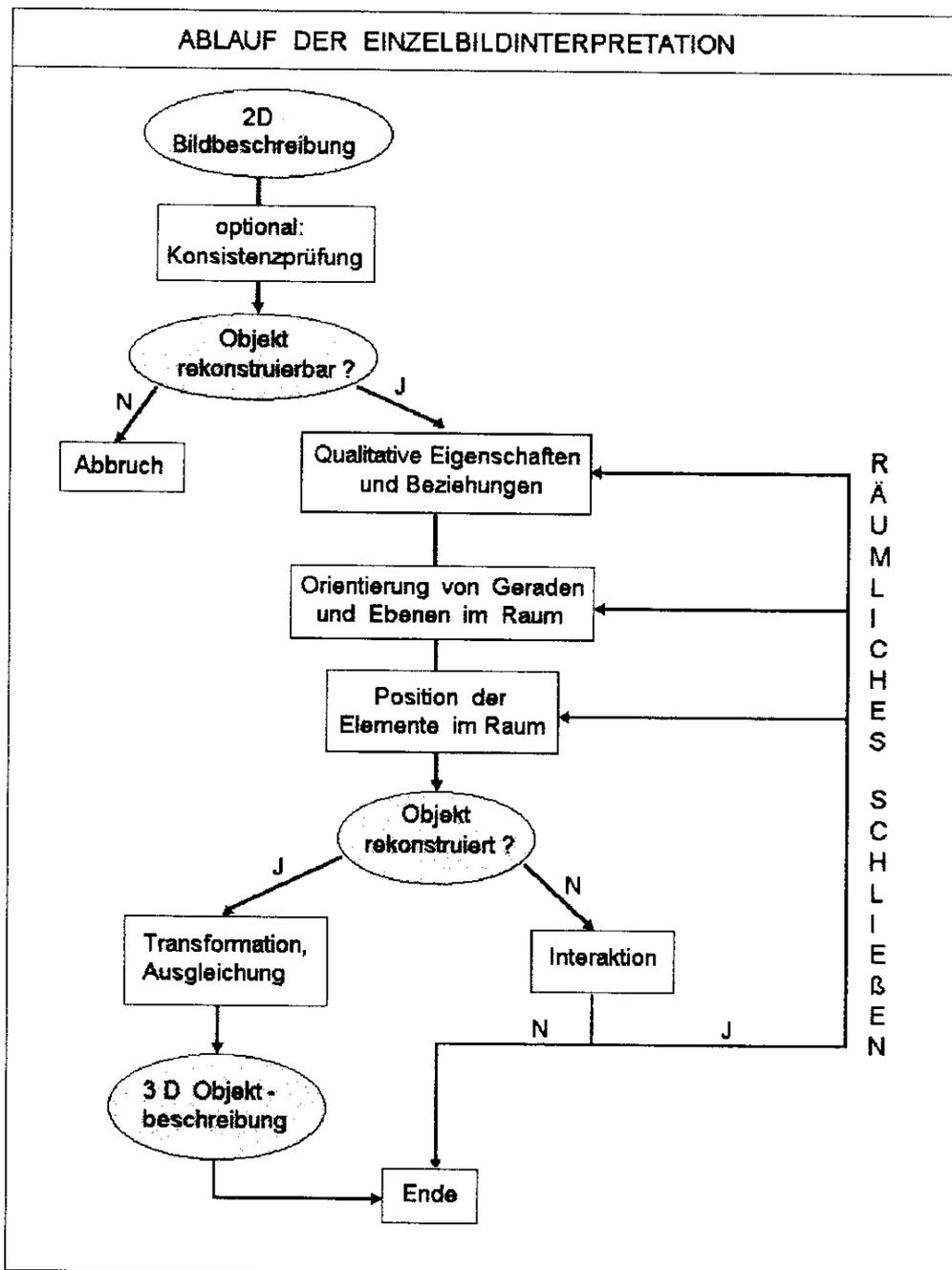


Abbildung 4: Ablauf des Verfahrens.

Als Ausgangsdaten liegt die aus dem digitalen Bild abgeleitete 2D Bildbeschreibung vor. Abb 6. zeigt die 2D Bildbeschreibung eines Hauses als Strichzeichnung visualisiert.

Ziel der Einzelbildinterpretation ist die Ableitung der 3D Objektbeschreibung, die durch das Tripel  $O = (P, K, F)$  repräsentiert wird.  $P$ ,  $K$  und  $F$  enthalten Teilmengen mit den für Punkte, Kanten und Flächen gewählten Attributen.

$$P_i = \{ d_1 d_2 d_3 \alpha \}$$

$$K_j = \{ \mu n_1 n_2 n_3 \beta r_1 r_2 r_3 \}$$

$$F_k = \{ n_1 n_2 n_3 \delta \}$$

Der Richtungsvektor  $\vec{d}$  zum Punkt  $P$  und die Normale  $\vec{n}$  der Projektions- oder Sehebene der Geraden, die in  $K_j$  gespeichert ist, lassen sich aus der 2D Bildbeschreibung direkt berechnen. Zu berechnen bleiben die numerischen Werte für die Orientierungen  $\vec{d}$ ,  $\vec{r}$  und  $\vec{n}$  sowie die Maßstabsfaktoren  $\alpha$ ,  $\mu$ ,  $\beta$  und  $\delta$ . Die Objektrekonstruktion ist abgeschlossen, wenn alle Attribute in den Teilmengen berechnet sind.

1. Handelt es sich um geschlossene Polygonzüge und ein triedrisches Objekt, wird die Konsistenzprüfung durchgeführt.
2. Im nächsten Schritt werden die qualitativen Eigenschaften und Beziehungen der Elemente aus dem Bild ins Objekt übertragen, insbesondere z.B. Inzidenzen und Identitäten.

Die folgenden beiden Schritte werden solange wiederholt bis keine weitere Schlußfolgerungen mehr möglich sind.

3. Orientierung von Geraden und Ebenen
4. Position der Punkte im Raum

Numerische Werte für die Attribute der Elemente im Raum können berechnet werden, wenn Hypothesen über Eigenschaften und Beziehungen der Objektelemente automatisch oder interaktiv generiert wurden. Das Verfahren zur automatischen Lokalisierung von Fluchtpunkten liefert für das Beispiel Gruppen von im Raum parallelen Geraden, die in Abb. 5a als dicke Linien markiert sind.

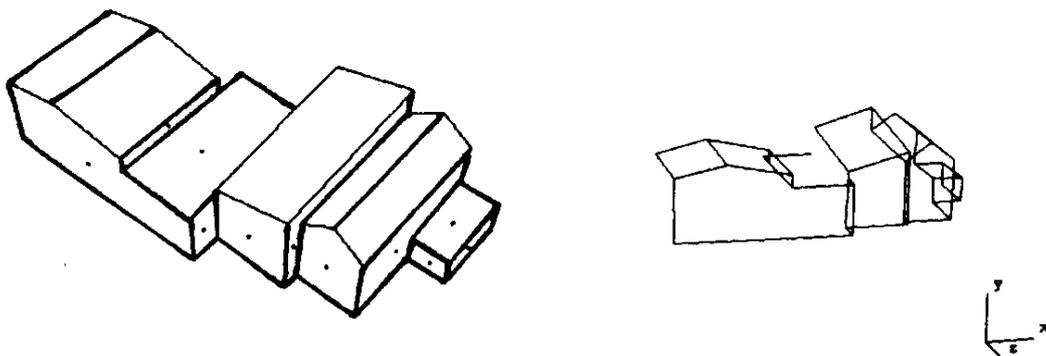


Abbildung 5: Häuserreihe a) Strichzeichnung b) Visualisierung des rekonstruierten Objekts.

Können keine weiteren Hypothesen automatisch generiert werden, testet das Verfahren, ob die bis zu diesem Zeitpunkt vorliegenden Orientierungen der Geraden und Ebenen ausreichen, die Position der Elemente im Raum zu berechnen. Es werden sukzessiv die Ebenen des Objekts und die zugehörigen Punkte und Geraden im Raum berechnet. Dieser Prozeß wird so lange wiederholt, bis in einem Durchgang keine neuen Attribute berechnet werden können.

Im 1. Durchgang können für das Beispiel die markierten Flächen berechnet werden. Im 2. Durchgang liegen auch für die anderen Flächen genügend Zwischenergebnisse vor, um für ihre Attribute numerische Werte zu berechnen.

Reichen die ermittelten Eigenschaften und Beziehungen nicht zur Beseitigung der am Objekt vorliegenden Freiheitsgrade, müssen weitere Eigenschaften und Beziehungen am Objekt interaktiv vorgegeben werden. Dieser Prozeß wird so lange wiederholt, bis auf die räumliche Struktur des Objekts geschlossen werden kann oder ein Ende des Verfahrens gewünscht wird. Ist eine Objektdistanz und/oder die Äußere Orientierung des Bildes bekannt, wird die 3D Objektbeschreibung in das Objektsystem transformiert. Die Ausgleichung der 3D Objektbeschreibung liefert eine konsistente Form für das Objekt unter Berücksichtigung aller ermittelten Informationen.

Die 3D Objektbeschreibung als Ergebnis des Verfahrens kann als Strichzeichnung von einem anderen Blickpunkt visualisiert werden (Abb. 5).

### 4.3 Beispiele

Bei dem aus einem Luftbild interaktiv abgeleiteten einfachen Gebäude werden 3 Geraden gruppiert, die sich in einem Fluchtpunkt schneiden (Abb. 6a). Da das Bild als Luftbild gekennzeichnet wurde und dieser Fluchtpunkt sich nahe des Hauptpunktes befindet, wird die Hypothese generiert, daß es sich um den Nadirpunkt handelt und die Geraden im Raum vertikal sind. Weitere Fluchtpunkte werden nicht gefunden. Mit den im Raum orientierten vertikalen Objektkanten ist es zunächst nicht möglich, das Objekt im Raum zu rekonstruieren.

Es erfolgt die rechnerische Interpretation eines interaktiv vorgegebenen Dreibeins (Abb. 6b). Da die Richtung einer der drei Geraden im Raum bereits bekannt ist, ist die Lösung nach Angabe der Spitze des Dreibeins eindeutig. Die Richtungsvektoren der drei Objektkanten reichen nun zur Rekonstruktion des abgebildeten Objekts im Raum (Abb. 6c).

In den folgenden Beispielen handelt es sich um die von analogen Bildvorlagen digitalisierten S/W Aufnahmen von Gebäuden (Abb. 7,8). Die Orientierung eines Dreibeins liefert genügend Information zur Rekonstruktion der Objekte im Raum.

Die 3D Objektbeschreibung wird jeweils von einem geänderten Aufnahmestandort und bei geänderten Kameraparametern in eine neue Bildebene projiziert und als Strichzeichnung visualisiert. Die Projektion der Grauwerte aus dem Originalbild in die Strichzeichnung liefert eine fotorealistische Darstellung des rekonstruierten Objekts.

## 5. ZUSAMMENFASSUNG UND AUSBLICK

Wir haben ein Verfahren zur Rekonstruktion von Gebäuden aus digitalen Bildern vorgestellt, das mit Hilfe objektspezifischer Bedingungen auf die räumliche Struktur des Objekts schließt. An Beispielen wurde gezeigt, daß die Interpretation einzelner Bilder zur Rekonstruktion räumlicher polyedrischer Objekte möglich ist. Das verwendete Modell ist genügend allgemein, um eine große Klasse von Gebäuden abzudecken.

In Gruppierungsprozessen werden Eigenschaften und Beziehungen der Objektelemente automatisch ermittelt. Dies betrifft insbesondere Kollinearitäten und Fluchtpunkte, die zur Stabilisierung und zur Orientierungsbestimmung verwendet werden. Der zur Bestimmung der geometrischen Elemente erforderliche Schlußfolgerungsprozeß ist vollständig und effizient automatisiert. Es sind derzeit Grauton-Bilder, Handskizzen und Perspektiv- oder Parallelprojektionen auswertbar.

Als Ergebnis liegt eine dreidimensionale Beschreibung des Objekts in Form von Punkten, Kanten und Flächen im Raum vor.

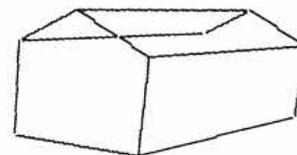
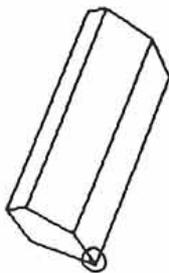
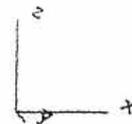
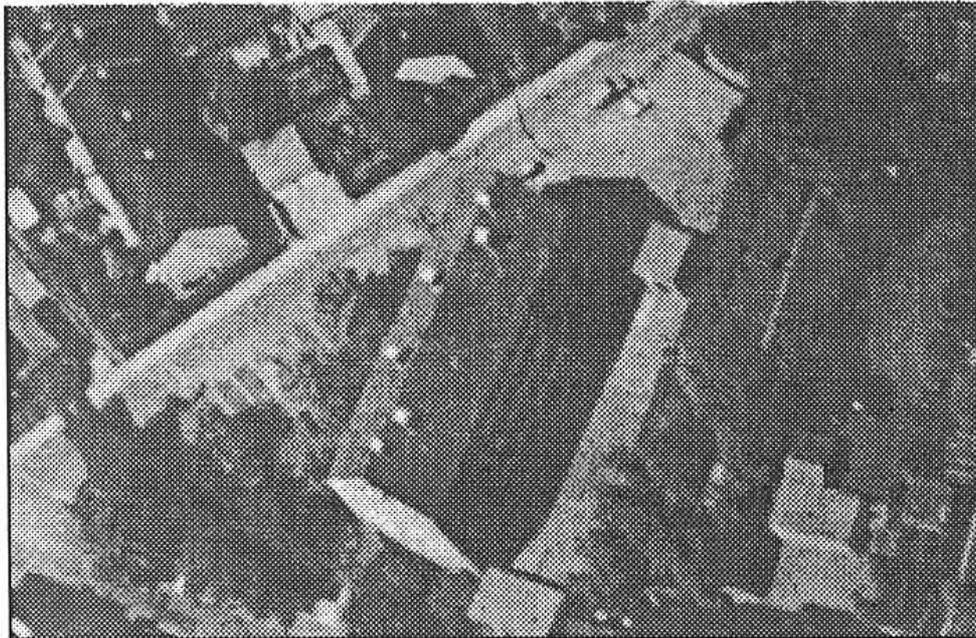


Abbildung 6: Ausschnitt eines Luftbildes mit Gebäude a) S/W Bild b) Strichzeichnung und als Dreibein markierter Punkt c) Ergebnis der Einzelbildinterpretation.

Als Ergebnis erhält man zunächst eine topologisch verknüpfte und geometrisch attributierte Beschreibung der Objekte. Sie stellen Näherungswerte für eine anschließende Ausgleichung dar, in der alle vorliegenden Beobachtungen und Bedingungen zur Beseitigung der verbleibenden Inkonsistenzen verwendet werden. Schließlich kann die zugrundeliegende Intensitätsinformation der Bilder für eine fotorealistische Darstellung genutzt werden.

Die automationsunterstützte Ableitung der dreidimensionalen Gestalt von Gebäuden und Gebäudekomplexen kann zur automatischen Erfassung und Aktualisierung dreidimensionaler Stadtmodellen in Informationssystemen beitragen. Mit der fotorealistischen Darstellung ist die Verbindung zu den Techniken der sogenannten 'Virtual Reality' hergestellt, etwa in Form von Videofilmen geplanter Objekte, die zur Erhöhung der Transparenz und der Bürgernähe bei Planungsentscheidungen beitragen können.

Eine vollautomatische Erfassung wird wohl mittelfristig nicht erreichbar sein, vor allem weil die in den verschiedenen Anwendungsbereichen anfallenden Aufgaben, wenn überhaupt, nur schwierig

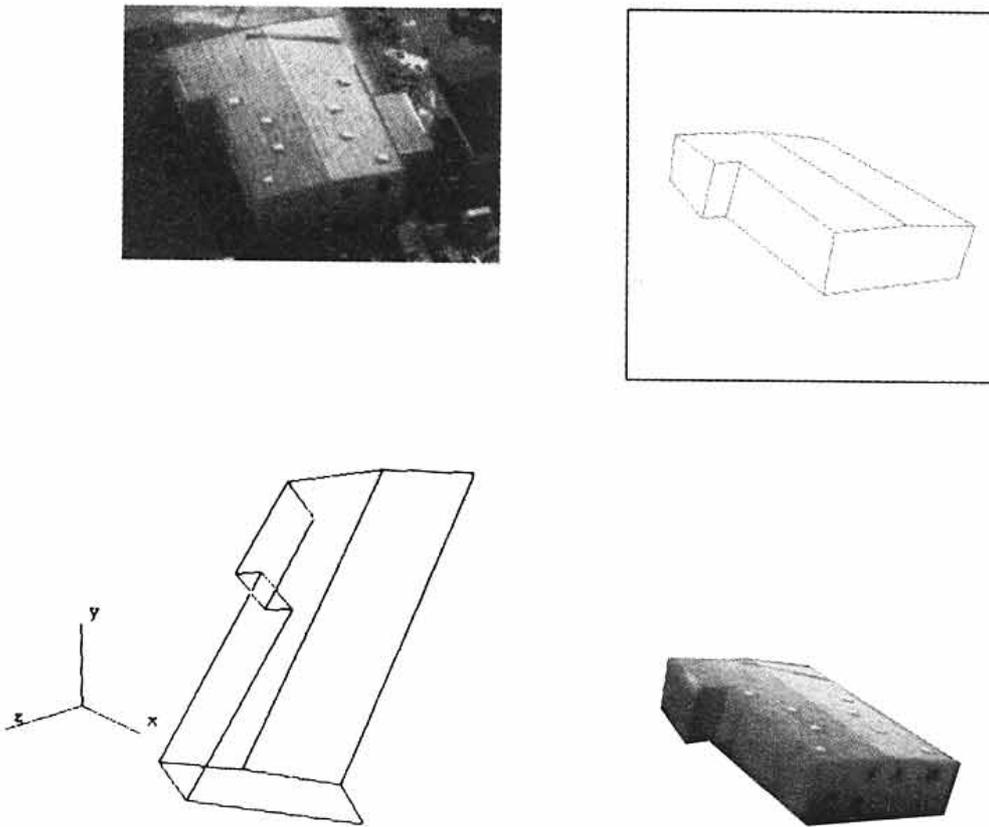


Abbildung 7: Fabrikgebäude a) S/W Bild b) Strichzeichnung c) Drahtmodell des im Raum rekonstruierten Objekts d) Fotorealistische Darstellung.

formalisierbar sind. Die Arbeit des Interpreten wird aber durch die u. U. weitreichende Teilautomatisierung von Bildanalyseprozessen wesentlich entlastet.

Das vorgestellte System läßt sich in verschiedener Hinsicht erweitern:

- Das Verfahren sollte um Methoden zur Interpretation abgebildeter Zylinder- und Kegelflächen ergänzt werden. Damit ließen sich z.B. Torbögen, Türme oder Erker erfassen.
- Symmetrien und andere Regelmäßigkeiten der Struktur sollten zur Stabilisierung und zur weiteren Reduktion des interaktiven Aufwandes verwendet werden.
- Zentrales Problem ist derzeit noch die Bildsegmentierung. Sie könnte bei Vorliegen eines genügend genau spezifizierten Objektmodells wesentlich verbessert werden, etwa bei der Interpretation der Regionen, die durch gerade und kegeschnittförmige Segmente begrenzt werden.
- Verfahren der qualitativen Auswertung von Handskizzen könnten zur weiteren Vereinfachung der Interaktion beitragen und mit quantitativen Verfahren, wie dem hier vorgestellten, verknüpft werden (vgl. Dickinson 1993).
- Zur vollständigen Erfassung von 3D-Objekten ist die Integration der Interpretationen mehrerer Einzelbilder notwendig. Die Zuordnung und Verschmelzung mehrerer partieller Objektbeschreibungen ist bei bekannter äußerer Orientierung vergleichsweise einfach und zuverlässig (vgl. Diemer 1993).
- Alternativ kann das Ergebnis einer Einzelbildinterpretation die Lösung des Zuordnungsproblems bei einer Stereoanalyse wesentlich vereinfachen. Denn die bereits genäherte vorliegende Objektinformation kann zur Suche in den anderen Bildern verwendet werden.

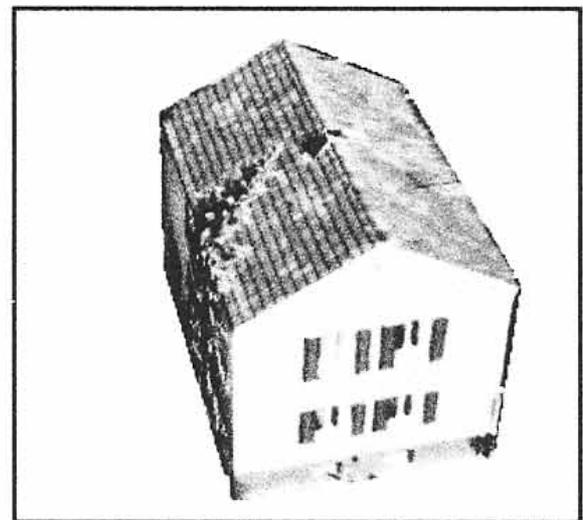
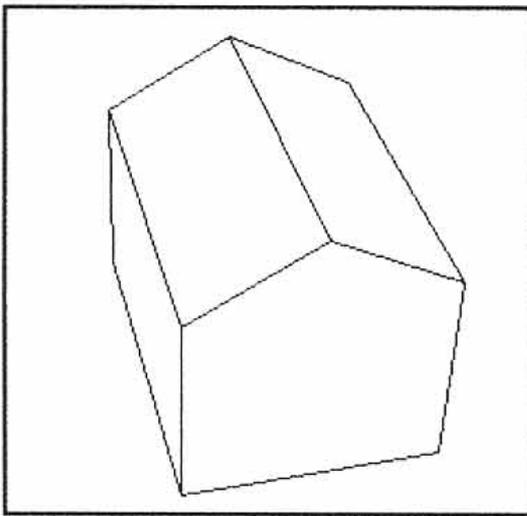
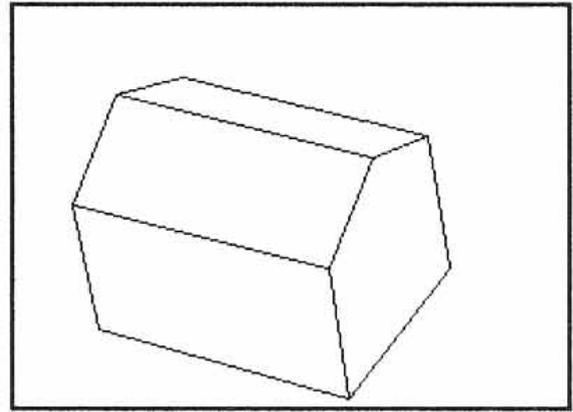
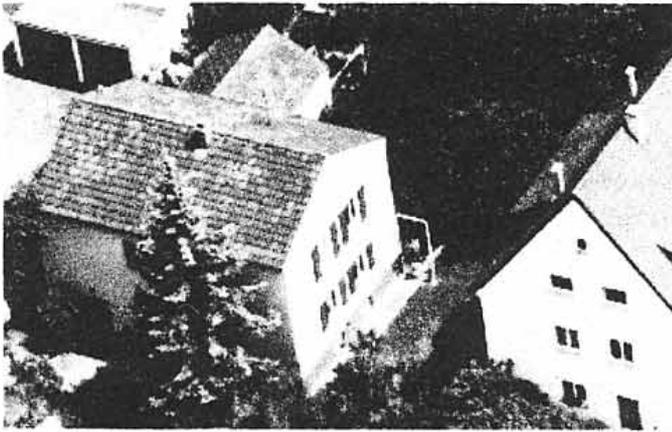


Abbildung 8: Satteldachhaus a) S/W Bild b) Skizze des Objekts c) Visualisierung der 3D Objektbeschreibung d) Fotorealistische Darstellung des rekonstruierten Objekts.

- Ein Fortschritt würde durch die überwiegend zweidimensionale Vorinformation aus dem Kataster erreicht. Diese Informationen würden das automatische Formulieren von Hypothesen erleichtern, da etwa die Position rechteckiger Dreibeine sehr sicher präzisiert werden kann.
- Schließlich sollte die 3D-Struktur der rekonstruierten Objekte mit dem Ziel einer volumenartigen Repräsentation interpretiert werden. Die aus dem CAD-Bereich bekannte Repräsentation des 'Constructive Solid Modelling' (CSG), bei der komplexe Objekte aus Teilkörpern zusammengesetzt werden, würde die Entwicklung von wirklich dreidimensionalen Stadtmodellen wesentlich unterstützen (vgl. Förstner 1993).

Das vorgestellte Verfahren wurde allgemein für polyedrische Objekte, nicht gezielt für die Gebäudeextraktion entwickelt, so daß sich nach entsprechender Adaption auch Aufgaben in anderen Anwendungsbereichen damit lösen lassen. Beispiele wären die sichtgestützte Rekonstruktion von Maschinenteilen bei der industriellen Fertigung oder die Interpretation gescannter Handzeichnungen im Kontext der Mensch-Maschine-Kommunikation (Fischler/Leclerc 1992) zur vereinfachten Eingabe der 3D-Struktur von polyedrischen Objekten.

Die Forschung im Bereich der Extraktion von topographischen Objekten, und damit auch von Siedlungsstrukturen, erhält durch die von der amerikanischen Regierung geförderten Projekte 'Research and Development for Image Understanding Systems' (RADIUS, Mundy et.al. 1992) und 'Image Understanding Environment' (IUE, Binford et.al. 1992) und durch das von der Deutschen Forschungsgemeinschaft geförderte Projekt 'Semantische Modellierung und Extraktion räumlicher Objekte aus Bildern und Karten' (Förstner 1993) neuen Auftrieb, so daß mit einer vergleichsweise raschen Entwicklung zu rechnen ist, die in wenigen Jahren die Praxis erreichen wird.

## 6. LITERATUR

- Barnard St. (1982): Interpreting Perspective Images, Technical Note 271, SRI Project 1009, Menlo Park, CA, 1982
- Beckröge W., Junius H. (1991): Klimatologische Begutachtung bestimmter stadtplanerischer Untersuchungsbereiche im Dortmunder Südwesten unter Einsatz eines Geo-Informationssystems, VR 53/3+4, Juni 1991
- Binford/Boult/Mundy/Hanson/Haralick/Ramesh/Kohl/Lawton/Morgan/Price/Strat (1992): The Image Understanding Environments Program, Proc. IU WS, San Diego, CA, 1992
- Clowes M. B. (1971): On Seeing Things, Artificial Intelligence 2, 79-116, 1971
- Diemer (1993): Verschmelzung partieller 3D-Modelle von Polyedern, Diplomarbeit, Institut für Photogrammetrie, Bonn, 1993
- Fischler M., Leclerc Y. (1992): Recovering 3-D Wire Frames from Line Drawings, Image Understanding Workshop, Jan. 1992
- Förstner W. (1993): Mustererkennung und 3D-Geoinformationssysteme, 3. internat. Anwenderforum für GIS, Duisburg 1993
- Förstner W. (1993): Semantische Modellierung und Extraktion räumlicher Objekte aus Bildern und Karten - Ein gemeinsames Forschungsvorhaben von Photogrammetrie und Kartographie, Zeitschrift für Photogrammetrie und Fernerkundung, 5/1993
- Gülch E. (1992): A Knowledge Based Approach to Reconstruct Buildings in Digital Aerial Imagery, Int. Archives of ISPRS, Com. III, Washington D.C., 1992
- Haralick R. M. (1989): Monocular Vision Using Inverse Perspective Projection Geometry, Analytic Relations, Conf. on Computer Vision and Pattern Recognition, San Diego, 1989, pp. 370ff
- Herman M., Kanade T. (1986): The 3-D Mosaic Scene Understanding System, From Pixels to Predicates, Ablex Publishing Corporation, New Jersey, 1986
- Huffmann D. A. (1971): Impossible Objects as Nonsense Sentences, Machine Intelligence, Vol. 6, ed. by B. Meltzer, D. Michie (Edinburgh University Press, Edinburgh, U. K. 1971), pp. 295-323
- Kanatani K. (1990): Group-Theoretical Methods in Image Understanding, Springer, Berlin 1990
- Maggee M. J., Aggarwal J.K. (1984): Determining Vanishing Points from Perspective Images, Computer Vision Graphics and Image Processing, 26, 256-267, 1984
- Müller C. (1992): Verwendung von Bildauswertungsmethoden zur Erkennung und Lagebestimmung von generischen polyedrischen Objekten im Raum, infix-Verlag, St. Augustin 1992
- Mulgaonkar P.G., Shapiro L. G. (1985): Hypothesis-Based Geometric Reasoning about Perspective Images, Proc. 3rd Workshop on Computer Vision, Bellaire, Michigan, 1985, pp. 11-18
- Mundy/Welty/Quam/Strat/Bremner/Horwedel/Hackett/Hoogs (1992): The RADIUS Common Development Environment, Proc. IU WS, San Diego, CA, 1992
- Schickler W. (1992): Merkmalsextraktion für Meßaufgaben in der digitalen Photogrammetrie, Zeitschrift für Photogrammetrie und Fernerkundung, 4, 1992

- Shomar W., Seetharaman G., Young T. (1992): An Expert System for Recovering 3D Shape and Orientation from Single View, in: Computer Vision and Image Processing, ed. by Shapiro, Rosenfeld, Academic Press, Inc., 1992, pp.459-515
- Sugihara K. (1986): Machine Interpretation of Line Drawings, The MIT Press, Cambridge, MA, 1986