
Automatische Extraktion von Gebäuden aus Luftbildern mit Produktionsnetzen

U. Stilla

In: F.K. List (Hrsg.)
Photogrammetrie im Wandel: Lehre - Forschung - Anwendung
Publikationen der Deutschen Gesellschaft für Photogrammetrie und Fernerkundung
1998, Band 6, S. 347-355

17. Wissenschaftliche-Technische Jahrestagung der DGPF
September 3 - September 5, 1997
Frankfurt, Germany

Automatische Extraktion von Gebäuden aus Luftbildern mit Produktionsnetzen

UWE STILLA *

Kurzfassung: In diesem Beitrag wird ein modellbasiertes Verfahren für die automatische 3D-Analyse von man-made - Strukturen aus Luftbildern vorgestellt. Zu erkennende Objekte wie beispielsweise Gebäude werden durch Produktionen modelliert und durch ein Produktionsnetz graphisch repräsentiert. Es wird eine Einteilung verschiedener Modelle vorgenommen. Am Beispiel eines parametrischen Modells für ein einfaches Giebedach wird ein Produktionsnetz zur 3D-Erkennung dargestellt und die Verarbeitungsschritte der Bildanalyse illustriert. Zur automatischen Verfahrensprüfung und Bewertung von implementierten Produktionsnetzen parametrischer Modelle wird eine Testumgebung vorgeschlagen.

1 Einleitung

Die automatische Interpretation von komplexen Strukturen, wie sie beispielsweise in Luftbildern vorkommen, ist eine schwierige Aufgabe. Zur Erkennung von *man-made*-Objekten wie Gebäuden werden häufig modellbasierte Ansätze verfolgt (GRUEN ET AL., 1997). In einigen Systemen haben sich Aspekte und Methoden aus dem Bereich der KI durchgesetzt (ADE, 1997). Die wissensbasierten Systeme zeigen die typische Entkopplung von dem Wissen zur Systemsteuerung zu dem der Objektmodelle (QUINT, 1997; STILLA, 1995; WILLUHN, 1997). Generell bieten sich zur Erfassung und Darstellung von Objektmodellen verschiedene Formen der Wissensrepräsentation wie z.B. Graphen, semantische Netze, formale Grammatiken oder Produktionensysteme an (BUNKE, 1985). Die Wahl einer geeigneten Form zur Wissensrepräsentation und insbesondere die Vorgehensweise bei der Wissensnutzung wird häufig durch die Aufgabenstellung und die vorliegenden Randbedingungen bestimmt.

Im vorliegenden Beitrag wird eine modellbasierte Methode für die automatische 3D-Analyse von *man-made*-Strukturen aus Luftbildern vorgestellt. Dabei wird in Anlehnung an die Sichtweise von MARR (1982), „Vision als aktive Konstruktion einer symbolischen Beschreibung aus Bildern“ verstanden. Die Analyse erfolgt schrittweise, indem komplexe Strukturen aus weniger komplexen Strukturen zusammengesetzt werden, wobei Ergebnisse von zunehmendem Abstraktionsgrad entstehen. Das Wissen wird durch Produktionen formuliert und graphisch durch ein Produktionsnetz dargestellt. Die Auswertung erfolgt „bottom-up“, da nicht von externem Zusatzwissen ausgegangen wird, mit dem vor der Rekonstruktion eine Detektion möglich ist.

An einem einfachen Beispiel zur Gebäudeextraktion aus einem Stereo-Bildpaar wird die Problemlösung auf den Ebenen der Modellierung (Kapitel 2), der Wissensrepräsentation (Kapitel 3), der Implementierung (Kapitel 4) und der Verfahrensprüfung (Kapitel 5) skizziert.

*Dr.-Ing. U. Stilla, Forschungsinstitut für Informationsverarbeitung und Mustererkennung (FGAN-FIM), Eisenstockstr. 12, 76275 Ettlingen, GERMANY, ++49 7243 992-35, e-mail: usti@gate.fim.fgan.de

2 Modellierung

2.1 Einteilung der Objektmodelle

Im Bereich der Bildanalyse wird der Begriff *Modell* oftmals mit unterschiedlicher Bedeutung verwendet. Wir unterscheiden nach dem Freiheitsgrad der topologischen und geometrischen Modellkomponenten hier folgende Objektmodelle:

Spezifische Modelle beschreiben Objekte mit fester topologischen Struktur. Sie werden weiterhin in Hinblick auf ihre geometrischen Einschränkungen in *feste*, *formfeste* und *parametrische* Modelle unterschieden. **Feste Modelle** bilden ideale geometrische Repräsentationen physikalischer Objekte. Sie haben in der Szene eine feste Position und Orientierung. Typische Beispiele bilden Objektbeschreibungen aus Karten. **Formfeste Modelle** haben einen festen Satz von geometrischen Relationen - ihre globale Position und Orientierung ist jedoch variabel. **Parametrische Modelle** erlauben mehr Transformationen als *formfeste Modelle*, wobei die strukturelle Komplexität jedoch fest bleibt. Die geometrische Variation des Modells wird durch einen Satz von Parametern bestimmt.

Generische Modelle sind wesentlich allgemeiner und beschreiben Objekte ohne eine feste topologische Struktur. Objekte, die durch generische Modelle beschrieben werden, können eine unterschiedliche Anzahl von Teilen besitzen. Als Beispiel sind Modelle für ein allgemeines Straßennetz oder eine Siedlung zu nennen.

2.2 Objektmodell DACH

In senkrecht aufgenommenen Luftbildern sind Gebäude hauptsächlich durch ihre Dächer zu erkennen. Für eine automatische Erkennung wird daher häufig das Daches modelliert. Für das folgende Beispiel wird von einzeln stehenden Häusern mit einfachen Giebeldächern ausgegangen. Die Abmessungen der Gebäude wie auch die Neigung der Dächer seien unbekannt. Ein Haus wird daher durch ein *parametrisches* Modell mit vier Parametern (l, w, h, γ) beschrieben (Abb. 1). Gibt man konkrete Werte des Parametersatzes an, wird ein *formfestes* Modell festgelegt; Parameterintervalle definieren eine Klasse von *formfesten Modellen*. Die angenommene Variationsbreite der Klasse S_{DACH} ist in Abb. 2 durch einige Repräsentanten dargestellt.

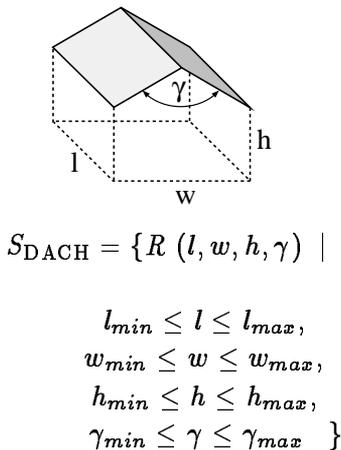


Abb. 1: *Parametrisches* Modell DACH

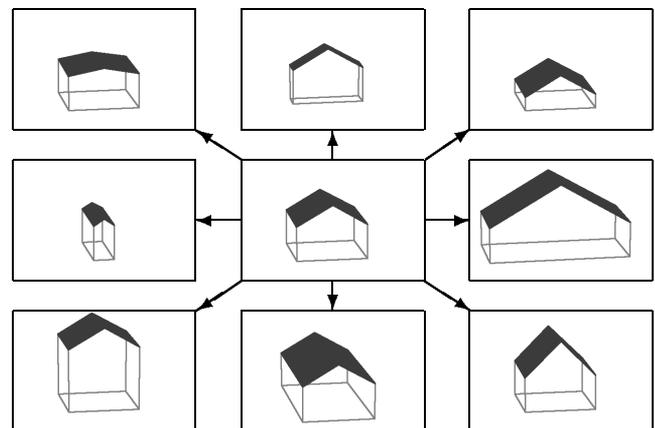


Abb. 2: *Formfeste* Modelle DACH variierend in Länge, Breite, Höhe und Dachwinkel

3 Wissensrepräsentation

3.1 Produktionen und Produktionsnetz

Allgemein läßt sich eine Produktionen als Anweisungen der Form „WENN *Bedingung* erfüllt, DANN *Aktion* ausführen“ angeben. Für die symbolische Analyse von Bildern und Szenen schreiben wir eine Produktion P_i wie in Abb. 3a dargestellt (STILLA, 1995).

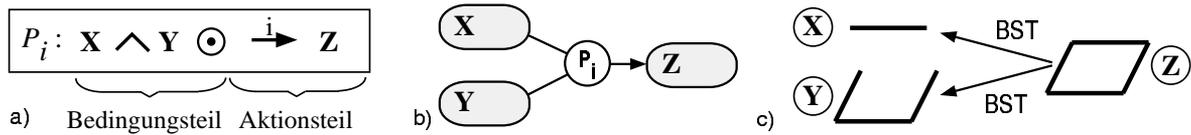


Abb. 3: a) Produktion, b) Produktionsnetz, c) graphisches Beispiel

Die Darstellung bedeutet: WENN ein Objekt vom Typ X und ein Objekt vom Typ Y ein spezielles Prädikat \odot erfüllen, DANN wird eine objektspezifische Generierungsfunktion $\overset{i}{\rightarrow}$ ausgeführt, die ein Objekt vom Typ Z produziert. Das Objekt vom Typ Z wird also aus den Objekten vom Typ X und Typ Y aufgebaut und „ersetzt“ damit sozusagen diese Eingangskonfiguration. In Abb. 3b ist für diese Produktion eine graphische Repräsentation durch ein Produktionsnetz dargestellt. Das Produktionsnetz zeigt den generellen Zusammenhang von Produktionen und Objekttypen (Konzepten). Die Zusammensetzung konkreter Objekte (Instanzen) wird durch einen Ableitungsgraphen dargestellt. Abb. 3c zeigt ein Beispiel für die Zusammensetzung eines parallelogramm-förmigen Objektes (Typ Z) aus einem linien-förmigen Objekt (Typ X) und einem u-förmigen Objekt (Typ Y). In der Regel läßt sich die Ersetzung durch die *Bestandteil*-Relation (BST) oder die inverse Relation *Teil_von* ausdrücken.

3.2 Produktionsnetz DACH

Ein Produktionsnetz für das parametrische Objektmodell DACH ist in Abb. 4 dargestellt. Es wird angenommen, daß die Teile des Daches als Rechtecke in der Szene und als Parallelogramme im Bild erscheinen. Ein Objekt PARALLELOGRAMM wird zerlegt in die 2D-Objekte U_STRUKTUR, WINKEL und LINIE. Beginnend mit den Primitivobjekten werden aus Paaren der Objekte LINIE Objekte WINKEL konstruiert (P_1). Wenn zwei Objekte WINKEL die Struktur eines offenen Parallelogramms formen, werden sie zu einem Objekt U_STRUKTUR zusammengefügt (P_2). Ein Objekt PARALLELOGRAMM kann zusammengesetzt werden, wenn Objekte U_STRUKTUR und LINIE kompatibel sind (P_3).

In der 3D-Analyse versucht man in unterschiedlichen Bildern Paare von 2D-Objekten (U_STRUKTUR oder PARALLELOGRAMM) zu finden, die die Projektion derselben Oberfläche darstellen. Korrespondieren 2D-Objekte von verschiedenen Bildern, so wird daraus

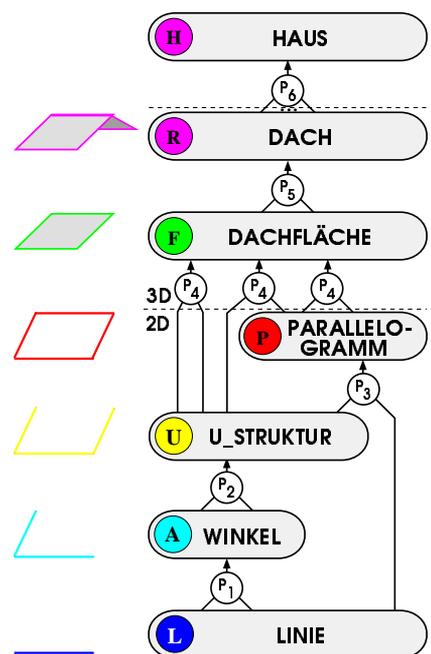


Abb. 4: Produktionsnetz DACH

das 3D-Objekt DACHFLÄCHE generiert (P_4). Sind Objekte DACHFLÄCHE so orientiert, daß ihre Oberflächennormalen einen Winkel γ einschließen, der in einem bestimmten Winkelintervall $[\gamma_{min}, \gamma_{max}]$ liegt und sind die Eckpunkte benachbart, so wird das Objekt DACH generiert (P_5).

Eine mögliche Konstruktion für den Aufbau einer Instanz DACH aus 14 Instanzen LINIE ist in Abb. 5 durch den Ableitungsgraphen dargestellt. Ordnet man den Kanten die Bedeutung *Bestandteil* oder *Teil-von* zu, entspricht die Beschreibung einem semantischen Netz. Um gegenüber Störungen tolerant zu sein, lassen sich die Objekte DACHFLÄCHE bei fehlenden Objekten LINIE auch direkt aus den Objekten U_STRUKTUR konstruieren (Abb. 4, P_4). Dadurch können aus einem Produktionsnetz verschiedenartige Ableitungsgraphen entstehen.

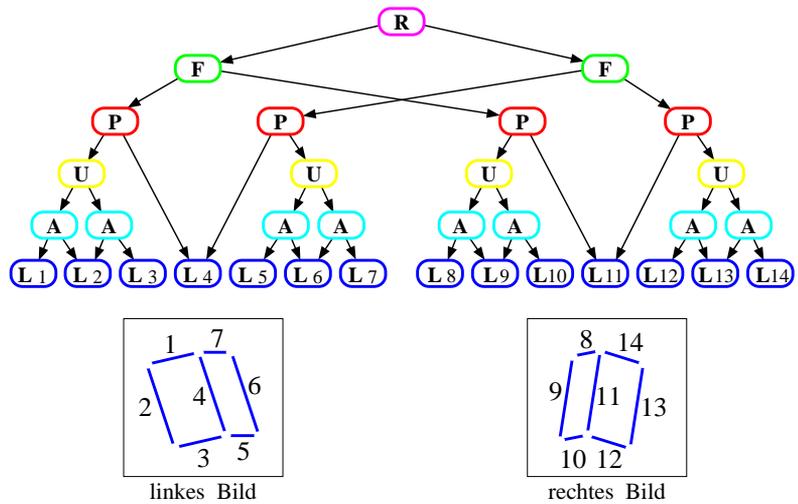


Abb. 5: Konstruktion der Instanz DACH (R)

Nach der 3D-Analyse werden Cluster der Objekte DACH untersucht. Das beste Objekt jedes Clusters wird ausgewählt und als Objekt HAUS gespeichert (P_6). Diese Objekte können wiederum mit Straßen und anderen Häusern in globalere Zusammenhänge wie z.B. Hausreihen eingebunden werden (STILLA & MICHALSEN, 1997).

4 Implementierung

Das Produktionsnetz wird in einer Blackboard-Architektur (NII, 1986) in der Systemumgebung BPI (LÜTJEN, 1986) implementiert. Das BPI-System besteht aus einer globalen Datenbasis (Blackboard), einer Menge von Verarbeitungsmoduln (Wissensquellen) und einer Systemsteuerung (Auswahlmodul). Jede Produktion ist durch ein Verarbeitungsmodul realisiert, das spezielle Relationen von bestimmten Typen von Objekten prüft, und bei Erfüllung ein neues Objekt generiert. Das System arbeitet kumulierend, d.h. daß die „ersetzte“ Eingangskonfiguration aus der Datenbasis nicht gelöscht wird. Einmal erzeugte Teilergebnisse stehen so zur Verfolgung verschiedener Hypothesen während der gesamten Analyse zur Verfügung. Das klassische *Backtracking* in Suchbäumen entfällt. Alle ausgeführten Zusammensetzungen der Objekte (Instanzen) werden durch eine Verzeigerung protokolliert, können zurückverfolgt und durch einen Ableitungsgraphen (siehe STILLA, 1995) dargestellt werden.

4.1 Vorverarbeitung

In der Vorverarbeitungsstufe wird aus den abgetasteten Luftbildern eine erste symbolische Beschreibung erzeugt. Prinzipiell stehen dazu eine Reihe von Verfahren zur Verfügung. Zur Beschreibung von *man-made*-Strukturen verwenden wir hier ein Mehrschwellen-Verfahren, das kurze Geradenstücke liefert. Abb. 6 zeigt für zwei Bildausschnitte die stufenweise Verarbeitung.

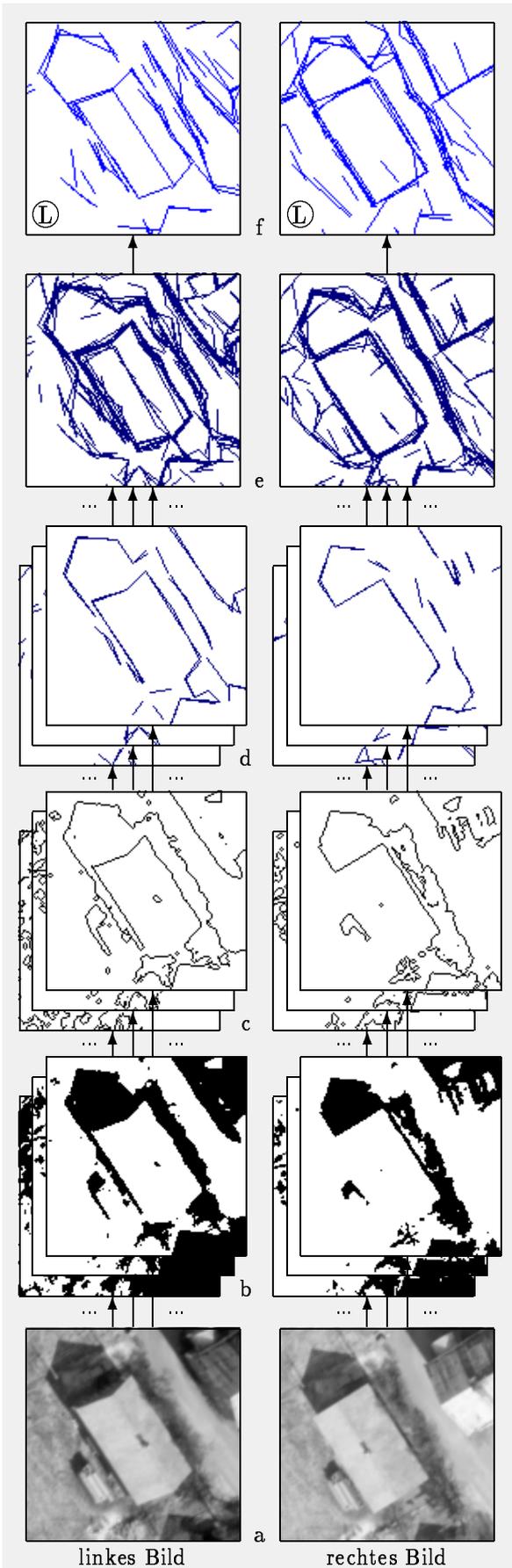


Abb. 6: Stufen der Vorverarbeitung

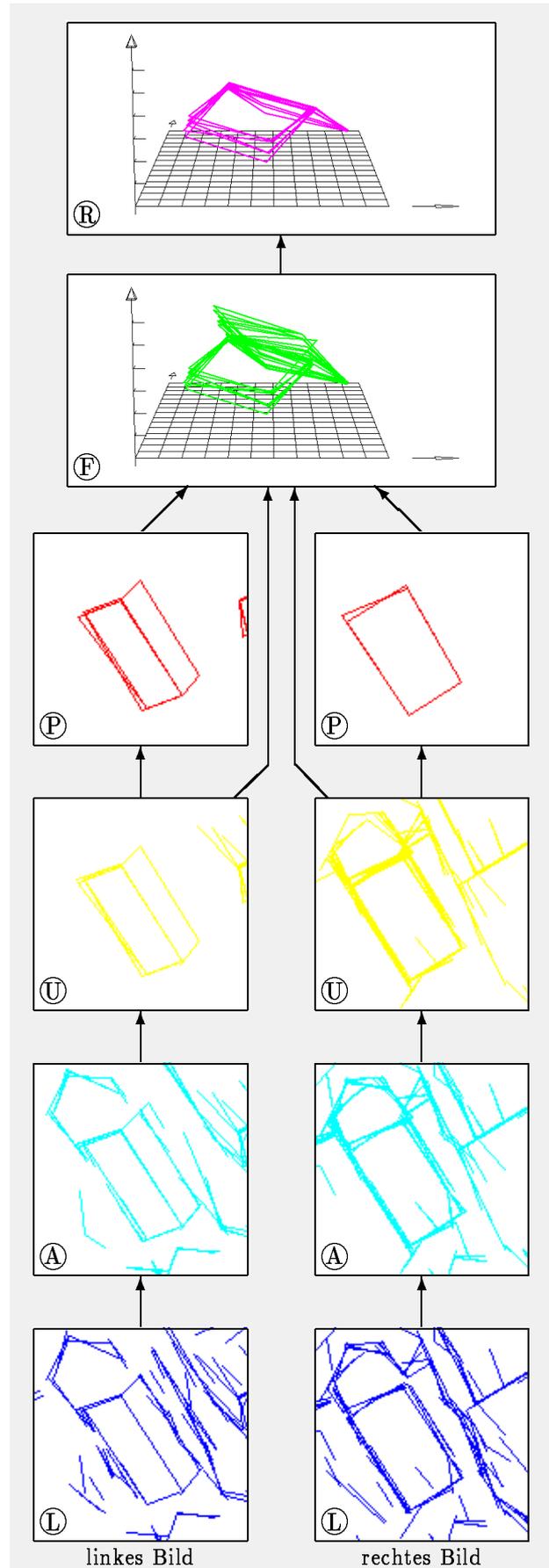


Abb. 7: Symbolische Analyse (Objektebenen)

Zunächst wird das Grauwertbild durch n_t Schwellwerte in eine Sequenz von Binärbildern überführt (Mehrschwellen-Segmentierung, Abb. 6b). In jedem Binärbild der Sequenz werden dann die Konturlinien detektiert (Konturdetektion, Abb. 6c). Ein dynamischer Split-Algorithmus approximiert die Konturlinien durch kurze Geradenstücke (Konturapproximation, Abb. 6d). Ohne Berücksichtigung der topologischen Relationen werden die Geradenstücke aller Bilder einer Sequenz als Menge zusammengefaßt (Linienammlung, Abb. 6e). Ein Gruppierungsprozeß faßt schließlich kurze kollineare Geradenstücke zu längeren Geradenstücken zusammen (Linienverlängerung, Abb. 6f). Diese Objekte LINIE (L) bilden die Primitivobjekte der nachfolgenden symbolischen Analyse.

4.2 Symbolische Analyse

Der Analyseablauf erfolgt datengetrieben durch ein Hypothese-Test Schema (STILLA, 1995). Einem Objekt werden dabei eine oder mehrere Hypothesen „Teil-von einem komplexeren Objekt“ zugeordnet. Ein Objekt U_STRUKTUR erhält z.B. die Hypothese „Teil-von-PARALLELOGRAMM“, die aus dem Produktionsnetz abgeleitet wird (Abb. 4). Ein solches Objekt-Hypothesen-Paar $X|Z$ *triggert* ein spezielles Verarbeitungsmodul, um die Hypothese Z zu verifizieren. Für das Objekt-Hypothesen-Paar U_STRUKTUR|PARALLELOGRAMM ist zu prüfen, ob ein passendes Objekt LINIE vorliegt. Dabei wird die in der Produktion angegebene Relation \odot in „invertierter“ Form genutzt. Statt die geometrische Relation über die Attribute aller Objektpaare von möglichen Partnern zu prüfen, werden ausgehend von der *triggernden* Instanz (z.B. U_STRUKTUR) die für die Bestätigung notwendigen Attributwerte eines Partners (z.B. LINIE) errechnet. Unter Berücksichtigung von Toleranzen wird durch einen assoziativen Zugriff auf die Datenbasis geprüft, ob Partner mit diesen Attributwerten existieren.

Abb. 7 zeigt für einen Szenenausschnitt in graphischer Form alle erzeugten Instanzen in den unterschiedlichen hierarchischen Stufen entsprechend dem Produktionsnetz aus Abb. 4. Verfolgt man die Bildauschnitte (L) bis (P), so erkennt man, daß aus der Menge der Primitivobjekte LINIE parallelogramm-förmige Strukturen herausgefiltert werden und aus diesen im 3D-Bereich Dächer rekonstruiert werden.

5 Verfahrensprüfung

Bei der Untersuchung von Verfahren unterscheiden wir abhängig von den verwendeten Bilddaten die Verfahrensprüfung auf der Basis von a) synthetischen Objekten, b) Modellobjekten und c) realen Objekten.

5.1 Verfahrensprüfung mit Bilddaten von synthetischen Objekten

Um die Eigenschaften eines Produktionsnetzes für ein parametrisches 3D-Gebäudemodell automatisch zu überprüfen, wurde eine Testumgebung entworfen (Abb. 8), in der Bilddaten von synthetischen Objekten generiert werden (STILLA ET AL., 1996).

Ausgehend von dem *parametrischen* Gebäudemodell werden durch zufällige Auswahl von Parameterwerten aus den spezifizierten Parameterintervallen *formfeste* Modelle erzeugt. Mit der Auswahl der Werte für Orientierung und Position werden diese *formfesten* Modelle dann als *feste* Modelle in der Modellszene plaziert. Statt einer zufäl-

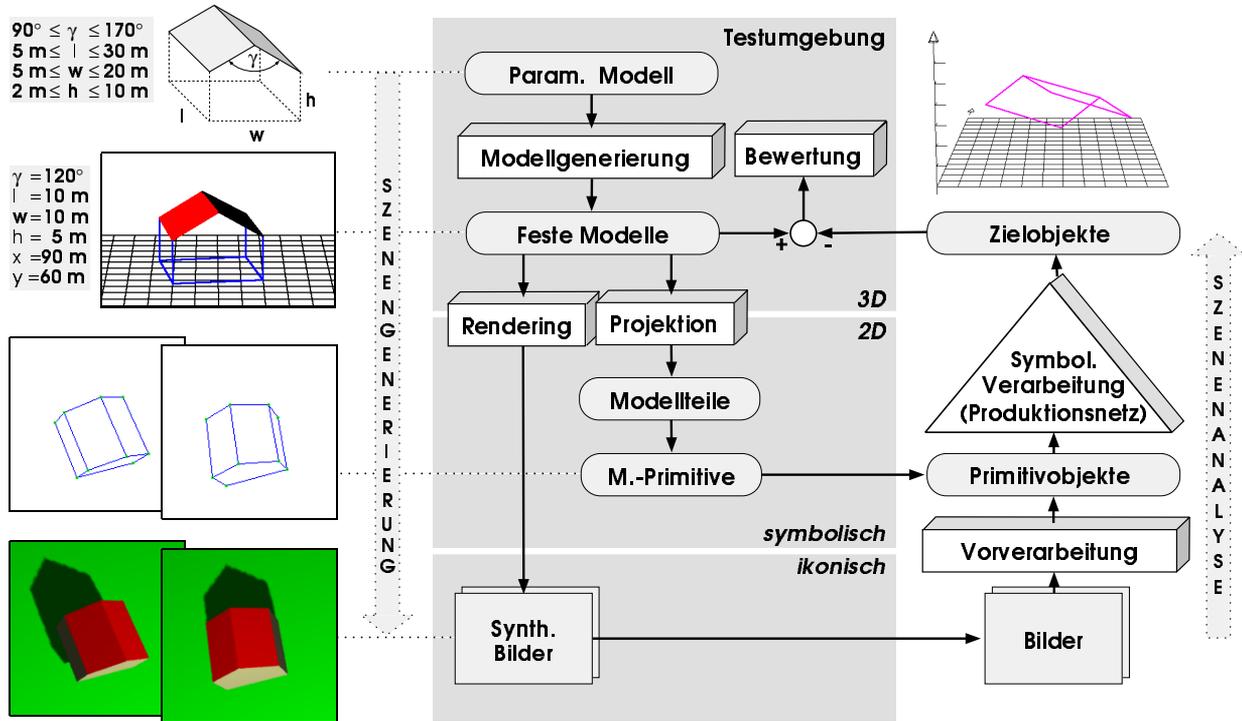


Abb. 8: Testumgebung für das Produktionsnetz des parametrischen Modells DACH

ligen Auswahl kann auch eine systematische Auswahl der Parameterwerte erfolgen, indem man z.B. Kombinationen von Grenzwerten wählt und die Positionierung und Orientierung in bestimmten Strukturen vornimmt (z.B. Reihen von Häusern mit gleicher Giebelorientierung). Unter Annahme von aufnahmespezifischen Parametern (Kameraorientierung, Bildgröße, Bildauflösung usw.) wird das Szenenmodell in Form einer *symbolischen* oder *ikonischen* Repräsentation in den Bildebene projiziert.

Eine *symbolische* Repräsentation eignet sich zur Prüfung des Produktionsnetzes ohne die Vorverarbeitung (innere Schleife in Abb. 8). Die Modelloberflächen werden in die 2D-Modellteile projiziert und durch einen Hidden-Line Algorithmus nachverarbeitet. Die Konturen der Modellteile (z.B. PARALLELOGRAMM) werden in Modellprimitive (z.B. LINIE) zerlegt, die dem implementierten Produktionsnetz als Primitivobjekte angeboten werden.

Eine *ikonische* Repräsentation eignet sich zur Prüfung des Analysesystems einschließlich der Vorverarbeitung (äußere Schleife in Abb. 8). Durch einen Rendering-Prozeß, der zusätzliche Szeneninformation zur Beleuchtung und Reflexion benötigt, werden synthetische Bilder erzeugt. Diese Bilder dienen als Eingangsdaten für das Analysesystem.

Da durch den Generierungsprozeß das Szenario bekannt ist, können die Ergebnisse der Analyse (Zielobjekte) automatisch mit den *festen* Modellen verglichen werden. Mit der symbolischen Übereinstimmung wird die semantische Korrektheit überprüft und aus den numerischen Abweichungen eine Bewertung abgeleitet. Weiterhin lassen sich auch Grenzleistungen bestimmen, indem für die 3D-Rekonstruktion systematisch kritische Situationen (z.B. Verkleinerung der Triangulationsbasis) simuliert werden.

Für Untersuchungen mit komplexeren Objektstrukturen in der Szenengenerierung-Szenenanalyse-Schleife bieten sich als *festen* Modelle auch Stadtmodelle an, wie sie z.B. in GÜLCH (1997), GRUBER ET AL. (1997) oder GRUEN (1997) dargestellt sind.

5.2 Verfahrensprüfung mit Bilddaten von Modellobjekten

Mit zunehmenden Anspruch an die photo-realistische Erscheinung synthetischer Objekte steigt der Aufwand für die Modellierung und Bildgenerierung erheblich an. Andererseits ist die systematische Aufnahme von Bilddaten realer Objekte durch eine Befliegung sehr kostenintensiv. Zusätzlich muß bei unbekannt Szenen für die automatische Verfahrensprüfung erst eine Szenbeschreibung (Ground-Truth) aufgenommen werden. Eine Alternative dazu bilden für bestimmte Untersuchungen Bilddaten von Modellobjekten (siehe z.B. LIN ET AL., 1995; HUERTAS ET AL., 1995; HECKER, 1996). Sie können witterungsunabhängig, kostengünstig und unter reproduzierbaren Bedingungen aufgenommen werden.

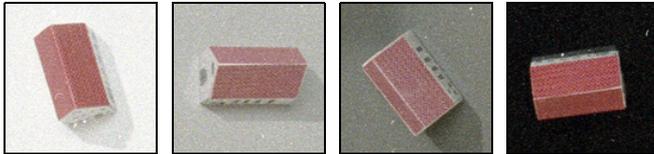


Abb. 9: Modelle auf unterschiedlichem Hintergrund



Abb. 10: Aufnahmeeinrichtung

5.3 Verfahrensprüfung mit Bilddaten von realen Objekten

Neben den Untersuchungen mit Bilddaten von synthetischen Objekten oder Modellobjekten sind Tests mit Bilddaten von realen Objekten unerlässlich. Zur Prüfung des in Kap. 3.2 dargestellten Produktionsnetzes wurde der ISPRS-Testdatensatz FLAT verwendet (Abb. 11). Die Ergebnisse wurden in STILLA & JURKIEWICZ (1996) diskutiert und in SESTER ET AL. (1996) ausgewertet.

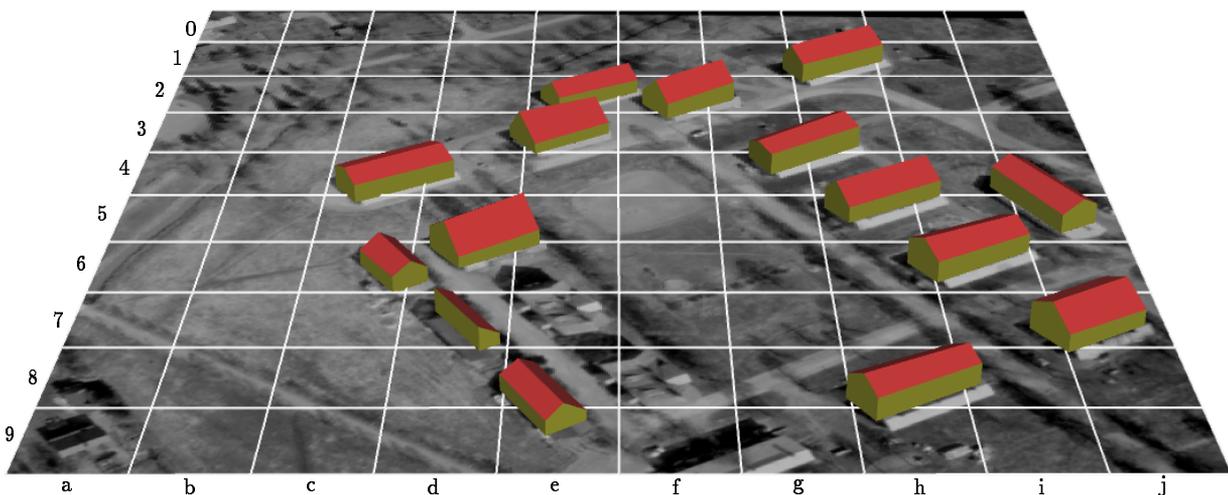


Abb. 11: 3D-Visualisierung der rekonstruierten Gebäude aus dem ISPRS-Testdatensatz FLAT

6 Literaturverzeichnis

- ADE F (1997) The role of artificial intelligence in the reconstruction of man-made objects from aerial images. In: Gruen A, Baltsavias EP, Henricsson O (eds), 23-32
- BUNKE H (1985) Modellgesteuerte Bildanalyse. Stuttgart: Teubner
- GÜLCH E (1997) Application of semi-automatic building acquisition. In: Gruen A, Baltsavias EP, Henricsson O (eds), 129-138
- GRUBER M, KOFLER M & LEBERL F (1997) Managing large 3D urban database contents supporting phototexture and level of detail. In: Gruen A, Baltsavias EP, Henricsson O (eds), 377-386
- GRUEN A & DAN H (1997) TOBAGO - a topology builder for the automated generation of building models. In: Gruen A, Baltsavias EP, Henricsson O (eds), 149-160
- GRUEN A, BALTSAVIAS EP & HENRICSSON O (1997) Automatic extraction of man-made objects from aerial and space images (II). Basel: Birkhäuser
- HECKER P (1996) Analysis of aerial images for ILS inspection. In: Jähne B et. al. (eds) Mustererkennung 1996. Berlin: Springer
- HUERTAS A, BEJANIN M & NEVATIA R (1995) Model registration and validation. In: Gruen A, Kuebler O & Agouris P (eds), Automatic extraction of man-made objects from aerial and space images. Basel: Birkhäuser
- LIN C, HUERTAS A & NEVATIA R (1995) Detection of buildings from monocular images. In: Gruen A, Kuebler O & Agouris P (eds), Automatic extraction of man-made objects from aerial and space images. Basel: Birkhäuser
- LÜTJEN K (1986) BPI: Ein Blackboard-basiertes Produktionssystem für die automatische Bildauswertung. In: Hartmann G (ed) Mustererkennung 1986, 8. DAGM-Symposium. Berlin: Springer, 164-168
- Marr D (1982) Vision. San Francisco: Freeman
- NII HP (1986) Blackboard systems. AI Magazine, 7: 38-53, 82-106
- QUINT F (1997) MOSES: A structural approach to aerial image understanding. In: Gruen A, Baltsavias EP, Henricsson O (eds), 323-332
- SESTER M, SCHNEIDER W & FRITSCH D (1996) Results of the test on image understanding of ISPRS working group III/3. ISPRS, XVIIIth Congress, International Archives of Photogrammetry and Remote Sensing, Vol. 31, Part B3, 768-773
- STILLA U (1995) Map-aided structural analysis of aerial images. ISPRS Journal of Photogrammetry and Remote Sensing, 50(4): 3-10
- STILLA U & JURKIEWICZ K (1996) Structural 3D-analysis of urban scenes from aerial images. ISPRS, XVIIIth Congress, International Archives of Photogrammetry and Remote Sensing, Vol. 31, Part B3, 832-838
- STILLA U, MICHAELSEN E & LÜTJEN K (1996) Automatic extraction of buildings from aerial images. In: Leberl F, Kalliany R, Gruber M (eds) Mapping buildings, roads and other man-made structures from images. Wien: Oldenbourg, 53-62
- STILLA U & MICHAELSEN E (1997) Semantic Modelling of man-made objects by production nets. In: Gruen A, Baltsavias EP, Henricsson O (eds) Automatic extraction of man-made objects from aerial and space images (II). Basel: Birkhäuser, 43-52
- WILLUHN W (1997) Integrating models for object recognition. Technical Report 175, CV-Group, Communication Technology Lab, ETH Zürich