
Wissenserwerb für Produktionsnetze zur 3D-Erkennung von Fahrzeugen

E. Michaelsen, U. Wankmüller, U. Stilla

In: P. Levi, R.-J. Ahlers, F. May, M. Schanz (Hrsg.)
Mustererkennung 1998
pp. 507 - 514, Berlin: Springer

20. DAGM-Symposium
September 29 - Oktober 01, 1998
Stuttgart, Germany

Wissenserwerb für Produktionsnetze zur 3D-Erkennung von Fahrzeugen

E. Michaelsen, U. Wankmüller und U. Stilla

Forschungsinstitut für Informationsverarbeitung und Mustererkennung (FGAN-FIM)
Eisenstockstr. 12, 76275 Ettlingen
{mich,wan,usti}@gate.fim.fgan.de

Kurzfassung. Zur dreidimensionalen Detektion und Klassifikation von Fahrzeugen aus Bildfolgen wird ein modellbasiertes Erkennungsverfahren eingesetzt. Die verwendeten Modelle werden durch Produktionsnetze beschrieben. Der Wissenserwerb erfolgt aus CAD-Modellen. Zur Erstellung eines neuen Produktionsnetzes werden interaktiv Hierarchiegraphen erstellt und aus den CAD-Daten standardisierte Produktionen parametrisiert.

1 Einleitung

Für die 3D-Erkennung von *man-made objects* aus Luftbildern werden häufig modellbasierte Verfahren vorgeschlagen [6]. Zur Modellierung sind verschiedene Formen der Wissensrepräsentation bekannt [13], wie z.B. Produktionen, Frames oder semantische Netze. Beim FIM werden Produktionen zur Erkennung von Brücken [10, 2], Straßenkreuzungen [15] und Gebäuden [16, 17, 18] in Luftbildern verwendet. Wegen der großen Variationsmöglichkeiten dieser Objekte finden zur Klassifikation parametrische Modelle Verwendung.

Bei der Modellierung von Fahrzeugen können zur Beschreibung formfeste Modelle [17] verwendet werden, da die spezielle Geometrie eines Fahrzeugtyps als bekannt vorausgesetzt werden kann. Ein Beispiel für die Modellierung und 3D-Erkennung eines Fahrzeugtyps in einer Bildfolge ist in [11, 12] dargestellt.

Soll eine breite Palette verschiedener Fahrzeugtypen detektiert werden, so ist jeweils die spezielle Geometrie zu erfassen und durch entsprechende Produktionen zu modellieren. Da geometrische Beschreibungen von Fahrzeugen typischerweise in Form von CAD-Modellen vorliegen, liegt es nahe, dieses Wissen zur Automatisierung der Modellierung zu verwenden.

CAD-Modelle wurden häufig für Analysen auf der Basis von Aspekt-Graphen verwendet, wie z.B. in [3, 8, 9]. Im vorliegenden Beitrag wird die Wissenserwerbskomponente eines Analysesystems beschrieben, mit der Produktionen aus CAD-Modellen von Fahrzeugen parametrisiert werden.

2 Objekterkennung mit Produktionsnetzen

Symbolische Verfahren der Mustererkennung arbeiten häufig auf Strukturen,

die durch Zeichenketten, Bäume, Felder oder Graphen beschrieben werden. Im Gegensatz dazu erfolgt die Objekterkennung mit Produktionsnetzen auf ungeordneten Mengen. Die Modellbeschreibung basiert auf der Theorie der *Koordinaten Grammatiken* [14, 11, 12].

Die Datenbasis, auf der die Produktionen arbeiten, besteht aus Objekten (Instanzen), die jeweils einen symbolischen Namen (Objekttyp) und einen Attributvektor beinhalten. Im Attributvektor einer Instanz sind z.B. Bild- oder Szenenkoordinaten, Orientierungen, Oberflächennormalen, Abmaße usw. abgespeichert. Ein Tupel solcher Instanzen bildet eine Konfiguration.

Produktion. Jede *Produktion* besteht aus einem *Bedingungsteil* und einem *Aktionsteil*. Der Bedingungsteil enthält Objekte und ein Prädikat, welches die zugehörigen Attribute erfüllen müssen. Der Aktionsteil enthält eine Attributberechnungsfunktion und erzeugt ein neues Objekt. Nehmen wir als Beispiel eine einfache Produktion der Form:

$$(\textit{Linie}, \textit{Linie}) \pi \xrightarrow{\phi} (\textit{Winkel})$$

Das Objekt *Linie* besitzt als Attribute Ortskoordinaten und eine Orientierung. Das Prädikat π beinhaltet die Relationen 'benachbart' und 'nicht parallel' und bildet zusammen mit dem Paar $(\textit{Linie}, \textit{Linie})$ den Bedingungsteil. Die Funktion ϕ bestimmt den Schnittpunkt der zu den Linien gehörenden Geraden. Dieser wird als Attribut für die entstehende Instanz des Objektes *Winkel* benötigt.

Produktionsnetz. Das Produktionsnetz zeigt den generellen Zusammenhang von Produktionen und Objekttypen (Konzepten). Ein Produktionsnetz ist ein gerichteter Graph, der ähnlich wie ein Petri-Netz [17] zwei verschiedene Arten von Knoten besitzt - Konzepte und Produktionen. Eine Kante führt von einem Konzept zu einer Produktion, wenn der Objekttyp im Bedingungsteil steht. Ist es mehrfach darin enthalten, wird die Kante entsprechend mehrfach aufgeführt. Eine Kante führt von einer Produktion zu einem Konzept, wenn der Objekttyp im ausführbaren Teil steht. Beispiele für Produktionsnetze finden sich in [15, 16, 17, 18].

Implementierung. Produktionsnetze werden in einer Blackboard-Architektur in der Systemumgebung BPI [10] implementiert. Jede Produktion ist durch ein Verarbeitungsmodul realisiert, das den Bedingungsteil auf einer Konfiguration prüft und bei Erfüllung eine neue Instanz generiert. Das System arbeitet akkumulierend, 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 [12]. Die Datenbasis wird in einem Assoziativspeicher gehalten, der einen effizienten Zugriff auf Mengen von Instanzen erlaubt [16].

3 Vom Modellwissen zum Produktionsnetz

Die interaktive Konstruktion von Produktionsnetzen erfolgt unter Zuhilfenahme von CAD-Modellen der zu erkennenden Fahrzeuge. Abb. 1a zeigt als Beispiel ein CAD-Modell eines Lastwagens in vier verschiedenen Ansichten.

CAD-Modell. Durch Anbindung eines Moduls an das CAD-System AUTOCAD wurde eine Möglichkeit geschaffen, durch Markieren von Modellteilen interaktiv eine Untergliederung des Modells in eine Hierarchie von Teilen vorzunehmen. Neben der Teil_von-Beziehung muß hierbei vom Anwender auf jeder Hierarchieebene festgelegt werden, ob eine Unterteilung in notwendige oder hinreichende Teile erfolgt. Als terminale Teile werden Flächen im Raum (Viereck, Rechteck, Dreieck und Kreis) verwendet, die verschiedene Attribute besitzen. Bei einem Rechteck werden z.B. Länge, Breite, Schwerpunkt, Normalenrichtung und Richtung der Längsachse festgehalten. Die Untergliederung bzw. Gruppierung der Teile wird in einem sogenannten Hierarchiegraphen gespeichert.

Hierarchiegraph. Der Hierarchiegraph enthält somit zwei Typen von Kanten, nämlich notwendige Teil_von-Beziehungen und hinreichende Teil_von-Beziehungen. Notwendige Teilobjekte sind paarweise verknüpft. Ist für die Erkennung eines Objekts die Verknüpfung von mehr als zwei notwendigen Teilobjekten erforderlich, erfolgt dies durch die Einfügung von Zwischenstufen in der Hierarchie. Für manche Teile, die als Details im CAD-Modell vorkommen, ist eine Erkennung in Luftbildern unwahrscheinlich. Sind diese Teile jedoch zur Visualisierung nützlich, können sie durch eine spezielle Markierung in den Hierarchiegraphen aufgenommen werden.

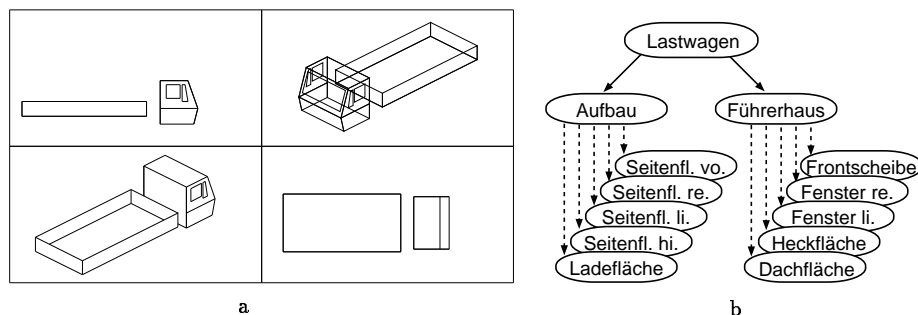


Abb. 1: Beispiel Lastwagen. a) CAD-Modell, b) Hierarchiegraph

Den Hierarchiegraphen des modellierten Lastwagens zeigt Abb. 1b. Bei diesem Beispiel wurde das Objekt *Lastwagen* in die beiden Teilobjekte *Aufbau* und *Führerhaus* unterteilt. Diese wurden wiederum in solche Einzelteile zerlegt, die durch Flächen im Raum beschrieben werden können. Für die Erkennung des *Lastwagens* erscheint beispielsweise die Präsenz von *Aufbau* und *Führerhaus* erforderlich (durchgezogene Pfeile - notwendig). Für die Präsenz des

Aufbaus oder des *Führerhauses* genügt dagegen jeweils ein einziges der darunter liegenden Teile (gestrichelte Pfeile - hinreichend), etwa die *Ladefläche* bzw. die *Dachfläche*.

Ausgehend von einem Referenzpunkt und einer Referenzrichtung im Raum werden die Lage und die Orientierung der Teilobjekte relativ zu den jeweiligen Vorgängern in der Hierarchie aus dem CAD-Modell bestimmt. Die relativen Transformationen werden den Kanten des Hierarchiegraphen als Attribute zugewiesen.

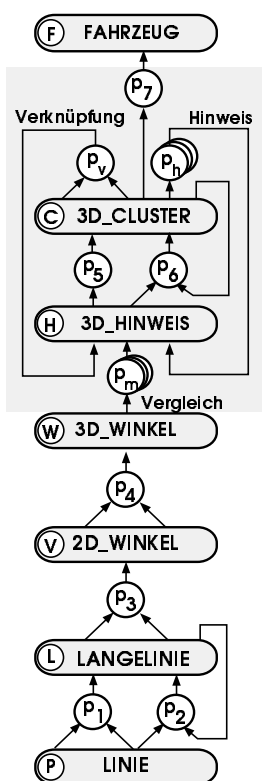


Abb. 2: Produktionsnetz

Produktionsnetz. Die in dem Produktionsnetz von Abb. 2 dargestellten Produktionen können im wesentlichen in zwei unterschiedliche Typen eingeteilt werden. Die Produktionen p_1 bis p_4 enthalten allgemeines Wissen für die 2D-Analyse im Einzelbild und für den Übergang in die 3D-Szene [17, 18]. Die verwendeten Parameter hängen von der Auflösung der Bilddaten ab. Die anderen Produktionen verarbeiten spezielles Wissen, wobei die parametrisierten Produktionen p_m , p_h , und p_v durch den Hierarchiegraphen spezialisiert werden.

Ausgehend von den Primitivobjekten LINIE werden in jedem Einzelbild durch sukzessive Linienverlängerung mit p_1 und p_2 die Objekte LANGELINIE erzeugt. Produktion p_3 liefert aus diesen Objekten die Objekte 2D_WINKEL. Korrespondieren zwei Objekte 2D_WINKEL in einem Bildpaar, so wird durch p_4 ein räumliches Objekt 3D_WINKEL generiert.

Eine spezialisierte Modellvergleichs-Produktion p_m prüft für ein Objekt 3D_WINKEL, ob es mit einem speziellen terminalen Modellteil kompatibel ist. Wenn dies der Fall ist, wird ein Objekt 3D_HINWEIS generiert, wobei die Bezeichnung des Modellteils als Attribut gespeichert wird.

Die räumliche Lage des Objektes wird durch Minimierung der Fehlerquadratrate bestimmt. Läßt das primitive Modellteil durch seine Symmetrie und Orientierung zu den Freiheitsgraden mehrfache Lageinterpretationen zu, wie z.B. bei einer horizontalen Rechteckfläche, so werden entsprechend viele unterschiedliche Objekte 3D_HINWEIS generiert.

Beispielsweise prüft die Produktion für das Modellteil *Dachfläche*, ob das Objekt 3D_WINKEL (Abb. 3a) eine senkrechte Normale besitzt, die beiden Schenkel ungefähr einen rechten Winkel bilden und mit der Länge und Breite der Rechteckfläche ungefähr übereinstimmen. Ein erzeugtes Objekt

3D_HINWEIS (Abb. 3b), das die Hypothese Dachfläche repräsentiert, besitzt als Attribute die räumliche Position des Schwerpunktes und die Ausrichtung der Fläche. Da das Objekt 3D_WINKEL für die *Dachfläche* eine um 180 Grad gedrehte Lageinterpretation zuläßt, wird ein zweites Objekt 3D_HINWEIS (Abb. 3c) generiert.

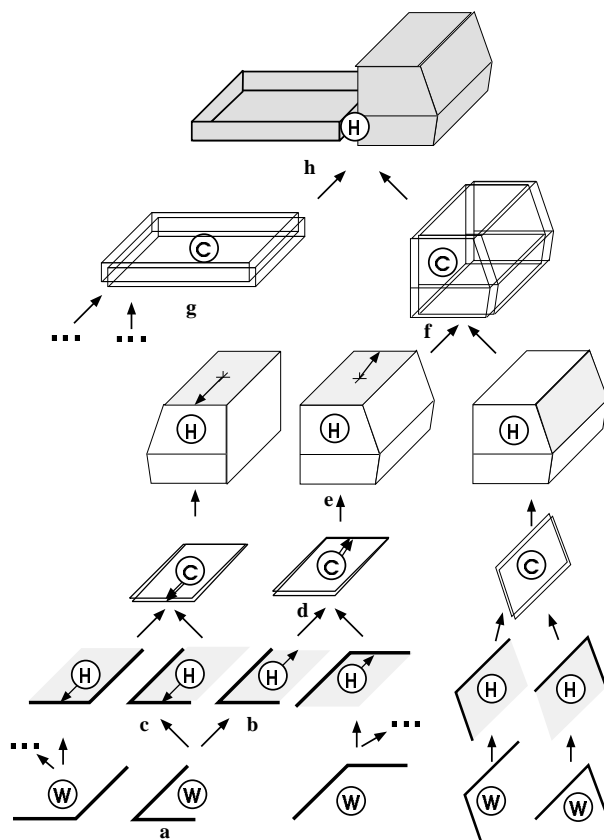


Abb. 3: Schrittweiser Aufbau

Die Produktion p_5 initialisiert einen Clusterbildungsprozeß, bei dem durch p_6 die Objekte 3D_HINWEIS gleicher Modellteile in ähnlicher Position und Ausrichtung durch Repräsentanten (3D_CLUSTER) zusammengefaßt werden (Abb. 3d). Die Attribute bestimmen sich durch Mittelwertbildung.

Während die Hinweis-Produktionen p_h Evidenz entlang der hinreichenden Kanten des Hierarchiegraphen transportieren, transportieren die Verknüpfungs-Produktionen p_v Evidenz entlang der notwendigen Kanten.

Beispielsweise wird durch eine Produktion p_h aus einem Objekt 3D_CLUSTER zu einer *Dachfläche* (Abb. 3d) ein Objekt 3D_HINWEIS zu ei-

nem *Führerhaus* (Abb. 3e) abgeleitet, ohne daß zusätzliche Bedingungen erfüllt sein müssen. Die Attribute zur Lage werden durch eine entsprechende Transformation berechnet, die aus dem Hierarchiegraphen stammt. Aus einem Objekt 3D_CLUSTER zum *Führerhaus* (Abb. 3f) und einem Objekt 3D_CLUSTER zum *Aufbau* (Abb. 3g) wird durch p_v ein Objekt 3D_HINWEIS zum *Lastwagen* (Abb. 3h) abgeleitet, wenn die räumlichen Positionen und Ausrichtungen der Objekte kompatibel sind.

Die Produktion p_7 erzeugt die Zielobjekte FAHRZEUG, wenn die Objekte 3D_CLUSTER zu einem *Fahrzeug* eine Mindestbewertung überschreiten. In die Bewertung geht z.B. die gefundene Anzahl der terminalen Modellteile ein.

4 Ergebnisse

Zur Verfahrensprüfung wurden Bildfolgen von Fahrzeugen in Schrägsichten mit einer Mittelformatkamera aufgenommen. Abb. 4a zeigt ein Beispiel, wobei die abgetasteten Bilderausschnitte eine Auflösung von jeweils 1000 x 1600 Bildpunkten besitzen. Das stehende Fahrzeug ist nicht in allen Bildern vollständig zu sehen. Die Kameraparameter wurden durch vier Paßpunkte bestimmt, die außerhalb der dargestellten Bildausschnitte liegen. Abb. 4b zeigt für die mittleren beiden Bilder das Ergebnis der Merkmalsextraktion. Die Generierung der Primitivobjekte erfolgte durch Segmentierung mit einem Mehrschwellenverfahren und einer nachfolgenden Konturapproximation mit einem dynamischen Split-Algorithmus (siehe [17]). Abb. 4c zeigt ein Resultat, das mit dem Hierarchiegraphen aus Abb. 1b und dem entsprechenden Produktionsnetz aus Abb. 2 erzielt wurde. Dafür wurde die beste resultierende Instanz in ein Bild der Folge rückprojiziert.

Die 3D-Klassifikation wurde mit drei verschiedenen Fahrzeugtypen untersucht, die in bodenbasierten Seitenansichten, Schrägsichten und aus senkrechten Blickrichtungen mit Kleinbild-, Mittelformat- und Luftbildkameras aufgenommen wurden.

5 Diskussion

Die Berücksichtigung hinreichender und notwendiger Teile erfolgt für die Analyse häufig durch Fallunterscheidungen. Bei semantischen Netzen wird dies z.B. in Form von Modalitäten [13] oder bei aspekt-basierten Systemen durch Untergraphen [4] erfaßt. Der vorgestellte Ansatz ermöglicht die Vermeidung dieser Fallunterscheidungen. Stattdessen werden auf höheren Hierarchieebenen die zu einem Objekt gehörenden hinreichenden und notwendigen Teile akkumuliert (Produktion p_6). Die Evidenz wird in der Szene, also im problemadäquaten Raum, zusammengefaßt, unabhängig davon welche Teilkonturen welcher Flächen in welchen Bildern gefunden werden. Durch diese Clusterung erhält der vorgestellte Ansatz auch seine Robustheit.

Im allgemeinen sind CAD-Modelle nicht für eine Erkennungsaufgabe erstellt worden. So sind bestimmte Bestandteile in Luftbildern nicht detektierbar (z.B. der Fahrzeugunterbau) oder aufgrund der Bildauflösung schlecht detektierbar

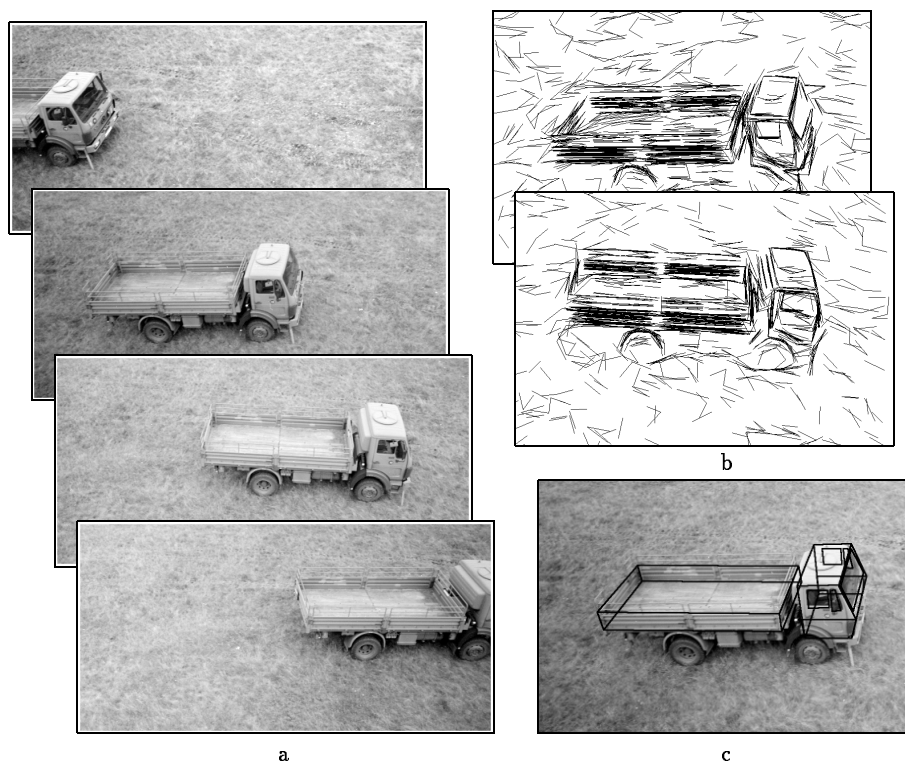


Abb. 4: Anwendungsbeispiel Lastwagen

(z.B. Fahrzeugdetails). Eine für die Erkennung nutzbare Untergliederung ist für komplexere Modelle in der Regel automatisch nicht ableitbar. Diese Aufgabe erfordert daher eine Interaktion mit dem Benutzer. Durch den teilautomatische Wissenserwerb können die Modelle jedoch rascher an neue Aufgaben angepaßt und erprobt werden.

6 Zusammenfassung

Der vorliegende Beitrag stellt einen Ansatz zum Wissenserwerb für die 3D-Objekterkennung mit Hilfe von Produktionsnetzen vor. Der Wissenserwerb erfolgt interaktiv mit Hilfe von CAD-Modellen. Um den Wissenserwerb hinreichend einfach und flexibel zu halten, wurde die Gestalt der Produktionsnetze gegenüber den früheren Untersuchungen eingeschränkt und standardisiert. Die vorhandenen Komponenten des Systems zur Beobachtung der entstehenden Statistiken, zur Fehlersuche, zur Erklärung und zur graphischen Ausgabe der Ergebnisse werden weiter verwendet.

Literaturverzeichnis

- [1] Füger H.; Lütjen K.; Michaelsen E.; Schwan K.: *Strukturorientierte 3D-Szenenanalyse in Bildfolgen*. In: Großkopf R. (Hrsg): *Mustererkennung 1990*, Springer, Berlin, 1990, 659-666.
- [2] Füger H.; Stein G.; Stilla U.: *Multi-populations evolution strategies for structural image analysis*, IEEE Conference on Evolutionary Computation (ICEC'94), Orlando, Vol I, 1994, 229-234.
- [3] Glauser T.; Bunke H.: *Generierung von Entscheidungsbäumen aus CAD-Modellen für Erkennungsaufgaben*. In: Burkhardt H.; Höhne K. H.; Neumann B. (Hrsg): *Mustererkennung 1989*, Springer, Berlin, 1990, 334-340.
- [4] Gmür E.; Bunke H.: *3-D object recognition based on subgraph matching in polynomial time*. In: Mohr R.; Pavlidis Th.; Sanfeliu A.: *Structural Pattern Analysis*, World Scientific, Singapore, 1989, 131-147.
- [5] Grimson L.; Eric W.: *Object recognition by computer: The role of geometric constraints*, MIT Press, Cambridge, Mass., 1990.
- [6] Gruen A., Baltsavias E.P., Henricsson O.: *Automatic extraction of man-made objects from aerial and space images (II)* Birkhäuser, Basel, 1997
- [7] Guoqing Z.: *Recognizing primitives using aspect-interpretation model matching in a both CAD and LP-based measurement system*, ISPRS Congress, 1996, Vol. 31, B3, 1018-1027
- [8] Ikeuchi K.: *Generating an interpretation tree from CAD model for 3D-object recognition in bin-picking tasks*, International Journal of Computer Vision, 1 1987, 145-165
- [9] Lanser S., Munkelt O., Zierl C.: *Robust video-based object recognition using CAD models*. In: Rembold U, et al. (Hrsg.): *Intelligent autonomous systems*, IOS Press, Amsterdam, 1995, 529-536.
- [10] Lütjen K.: *BPI: Ein Blackboard-basiertes Produktionssystem für die automatische Bildauswertung*. In: Hartmann G. (Hrsg.): *Mustererkennung 1986*, Springer, Berlin, 1986, 164-168.
- [11] Michaelsen E, Stilla U.: *Remarks on the notation of coordinate grammars*. In: Amin A, Dori D, Pudil P, Freeman H (Eds.): *3D Advances in Pattern Recognition*, Springer: Berlin, 1998, 421-428.
- [12] Michaelsen E.: *Über Koordinaten Grammatiken zur Bildverarbeitung und Szenenanalyse*, Diss., Techn. Fak., Univ. Erlangen, in Vorbereitung 1998.
- [13] Niemann H.: *Pattern analysis and understanding*, Springer, Berlin, 1990.
- [14] Milgram D. L.; Rosenfeld A.: *A note on 'grammars with coordinates'*. In: Nake F.; Rosenfeld A.: *Graphic languages*, North Holland, Amsterdam, 1972, 187-194.
- [15] Stilla U.: *Map-aided structural analysis of aerial images*. ISPRS Journal of Photogrammetry and Remote Sensing, Vol 50, 1995, 3-10.
- [16] Stilla U.; Michaelsen E.; Lütjen K.: *Structural 3D-analysis of aerial images with a blackboard-based production system*. In: Gruen A.; Kuebler O.; Agouris P.: *Automatic extraction of man-made objects from aerial and space images*, Birkhäuser, Basel, 1995, 53-62.
- [17] Stilla U.; Michaelsen E.; Lütjen K.: *Automatic extraction of buildings from aerial images*. In: Leberl F.; Kalliany R.; Gruber M.: *Mapping buildings, roads and other man-made structures from images*, IAPR TC-7 Workshop, Oldenbourg, Wien, 1996, 229-244.
- [18] Stilla U.; Michaelsen E.: *Semantic modeling of man-made objects by production nets*. In: Gruen A.; Baltsavias E. P.; Henricsson O.: *Automatic extraction of man-made objects from aerial and space images II*, Birkhäuser, Basel, 1997, 43-52.