

# **Wissensbasierte Interpretation multitemporaler Luftbilder**

Vom Fachbereich Elektrotechnik und Informationstechnik  
der Universität Hannover

zur Erlangung des akademischen Grades

**Doktor-Ingenieur**

genehmigte

**Dissertation**

von

**Dipl.-Ing. Stefan Growe**

geboren am 24. September 1967 in Bad Oeynhausen

2000

Referent: Prof. Dr.-Ing. C.-E. Liedtke  
Korreferent: Prof. Dr.-Ing. C. Heipke  
Tag der Promotion: 12. Januar 2001

---

## Vorwort

Die vorliegende Dissertation entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung der Universität Hannover in der Zeit von September 1994 bis September 2000.

Herrn Prof. Dr.-Ing. C.-E. Liedtke danke ich für die Anregung zum Thema dieser Arbeit, die Übernahme des Hauptreferats und die hervorragenden Arbeitsbedingungen am Institut. Er schuf eine außerordentlich angenehme und freie Arbeitsatmosphäre und hat mit motivierenden Diskussionen das Entstehen dieser Arbeit maßgeblich gefördert.

Herrn Prof. Dr.-Ing. C. Heipke, dem Leiter des Instituts für Photogrammetrie und Ingenieurvermessungen der Universität Hannover, gilt mein Dank für die Übernahme des Koreferats und die fruchtbare Zusammenarbeit in den gemeinsamen, interdisziplinären Projekten.

Meine Zeit am Institut wird mir immer in positiver Erinnerung bleiben. Allen Kollegen, die durch ihre Anregungen, ihre ständige Hilfsbereitschaft und ihren Humor für ein konstruktives, kollegiales und angenehmes Arbeitsklima gesorgt haben, danke ich sehr herzlich. Mein besonderer Dank gilt all denen, die direkt oder indirekt zum Gelingen dieser Arbeit beigetragen haben. Besonders erwähnen möchte ich dabei meine Kollegen Dipl.-Inform. Jürgen Bückner, Dipl.-Inform. Hans Koch, Dipl.-Ing. Martin Pahl, Dipl.-Ing. Kian Pakzad und Dr.-Ing. Ralf Tönjes sowie meine Diplomanden bzw. studentische Hilfskraft Dipl.-Ing. Jan Bigalke, Dipl.-Ing. Torsten Schröder und cand. el. Martin Libor. Darüber hinaus bedanke ich mich bei allen Kollegen, die durch ihre (Neben-) Tätigkeit im Soft- und Hardwaremanagement stets für eine erstklassige Infrastruktur gesorgt haben. Für die kritische Durchsicht der Arbeit danke ich Dipl.-Inform. Jürgen Bückner und Dipl.-Ing. Jochen Wingbermhle.

Ein großes Dankeschön geht an meine Eltern, die mich schon früh zum selbständigen Arbeiten angeregt und mich während meines Werdegangs auf vielfältige Weise unterstützt haben. Für den moralischen Rückhalt bedanke ich mich außerdem bei meinem Bruder und seiner Familie sowie bei meinen Schwiegereltern.

Besonders in der Endphase dieser Arbeit blieb das Privatleben oft auf der Strecke. Meine Freunde und Bekannte bitte ich um Nachsicht, falls ich sie in dieser Zeit vernachlässigt haben sollte. Meiner Frau Dr. med. Ulrike Growe möchte ich ganz herzlich für ihr Verständnis, ihre Geduld und ihre liebevolle Unterstützung danken.

Hannover, im September 2000



---

## Gliederung

<b>1. Einleitung</b> .....	<b>1</b>
1.1 Einführung in die Thematik .....	1
1.2 Stand der Technik .....	2
1.3 Ziele und Aufbau der Arbeit .....	6
<b>2. Wissensbasierte Interpretation multisensorieller Luftbilder</b> .....	<b>8</b>
2.1 Definition grundlegender Begriffe .....	8
2.2 Systemarchitektur .....	11
2.3 Repräsentation des generischen Szenenmodells .....	13
2.3.1 Semantische Netze für die Wissensrepräsentation .....	14
2.3.2 Inhalt und Strukturierung der Wissensbasis .....	19
2.4 Steuerung der Bildinterpretation .....	22
2.4.1 Inferenzmaschine .....	23
2.4.2 Bildinterpretation als Suchprozeß .....	27
2.4.3 Possibilistisches Bewertungssystem .....	29
<b>3. Repräsentation und Nutzung von temporalen Wissensinhalten</b> .....	<b>34</b>
3.1 Systemanforderungen für die Auswertung multitemporaler Luftbilder ...	35
3.1.1 Wissensrepräsentation .....	36
3.1.2 Wissensnutzung .....	38
3.2 Ansätze zur Verarbeitung zeitlicher Zusammenhänge .....	39
3.2.1 Zustandsübergangsdiagramme und Markov-Ketten .....	41
3.2.2 Zeitliches Schlußfolgern in der KI .....	43
3.2.3 Netzplantechnik .....	46
3.3 Erweiterung des Bildinterpretationssystems .....	50
3.3.1 Repräsentation von Zustandsübergangsdiagrammen .....	51
3.3.2 Modellgetriebene Prädiktion von Folgezuständen .....	56
3.3.3 Strategie für die Interpretation multitemporaler Luftbilder .....	58
3.3.4 Zusammenfassung der temporalen Erweiterungen .....	62

---

<b>4. Ein probabilistisches Bewertungssystem .....</b>	<b>64</b>
4.1 Theorie der Bayes-Netze .....	67
4.1.1 Begriffe und Definitionen .....	68
4.1.2 Bayes-Netze als Wissensbasis .....	71
4.1.3 Propagierung von Evidenz in Bayes-Netzen .....	73
4.2 Bayes-Netze für die Bewertung von semantischen Netzen .....	79
4.2.1 Vom semantischen Netz zum Bayes-Netz .....	80
4.2.2 Bewertung eines Suchbaumknotens .....	87
4.3 Vergleich mit dem possibilistischen Bewertungssystem .....	90
<b>5. Anwendungsbeispiele .....</b>	<b>93</b>
5.1 Detektion eines Messegeländes .....	93
5.1.1 Datenmaterial und Wissensbasis .....	94
5.1.2 Bildverarbeitungsverfahren .....	96
5.1.3 Interpretationsprozeß .....	97
5.1.4 Vergleich der Bewertungssysteme .....	104
5.2 Erkennung baulicher Veränderungen in Industriegebieten .....	106
5.2.1 Datenmaterial und Wissensbasis .....	106
5.2.2 Interpretationsprozeß .....	108
5.2.3 Vergleich der Bewertungssysteme .....	112
5.3 Diskussion der Ergebnisse .....	113
<b>6. Zusammenfassung .....</b>	<b>115</b>
<b>Anhang .....</b>	<b>119</b>
A-1 Approximation von Polygonen .....	119
A-2 Detektion von Parkplätzen .....	122
A-3 Graphische Benutzerschnittstelle .....	124
<b>Literatur .....</b>	<b>126</b>
<b>Schlagwortverzeichnis .....</b>	<b>141</b>

---

## Abkürzungen und Formelzeichen

### Notation der Formelzeichen (i. Zshg. mit Bayes-Netzen, s. Kapitel 4)

Normalschrift groß: $X$	diskrete Zufallsvariable
Normalschrift klein: $x$	diskrete Zufallsvariable $X$ mit zugewiesenem Wert $x$
Fettdruck groß: $X, Y, Z$	Menge von Zufallsvariablen
Fettdruck klein: $x, y, z$	Menge von Zufallsvariablen mit zugewiesenen Werten
Unterstrichen: $\underline{M}$	Matrix

### Verwendung von Indizes und Intervallen

$i, j, k$	allgemeine Indizes
$H, F, P, K$	Hypothese, fehlende, partielle und komplette Instanz
$inv$	inverse Richtung einer Relation
$cd$	kontextabhängige Relation
$obl$	obligatorische Relation
$opt$	optionale Relation
$[min\ max]$	Intervall einschließlich der Werte $min$ und $max$

### Abkürzungen

ATKIS	Amtliches Topographisch-Kartographisches Informationssystem
CPM	Critical Path Method (Verfahren der Netzplantechnik)
DAG	Gerichteter azyklischer Graph (engl.: Directed Acyclic Graph)
DHM	Digitales Höhenmodell
DLM	Digitales Landschaftsmodell
GIS	geographisches Informationssystem
IR	Infrarot
MPM	Metra-Potential-Method (Verfahren der Netzplantechnik)
PERT	Program Evaluation & Review Technique (Verf. d. Netzplantechnik)
RMK	Reihenmeßkamera
SAR	Radar mit synthetischer Apertur (engl.: Synthetic Aperture Radar)

### Kanten des semantischen Netzes

<i>attr-rel</i>	attributierte Relation zur Beschreibung topologischer Beziehungen
<i>cd(..)-of</i>	kontextabhängige Version eines Kantentyps
<i>con-of</i>	Konkretisierungskante zur Verbindung von Objekten unterschiedlicher konzeptioneller Ebenen
<i>data-of</i>	Relation zur Bindung von Daten
<i>instance-of</i>	Relation zur Verbindung von Instanzen mit ihren Konzepten
<i>is-a</i>	Spezialisierungskante zur Definition von Klassenhierarchien
<i>part-of</i>	Bestandteilkante zur Zerlegung in Objektteile
<i>temp-rel</i>	Temporale Relation zur Beschreibung zeitlicher Veränderungen

### Formelzeichen

$A_1, A_2$	überlappender bzw. nicht überlappender Bereich des Meßwerts $E$ eines Attributes mit dem korrespondierenden erwarteten Meßwert $H$
$BEL(x)$	Glaubwürdigkeit der Aussage $\{X=x\}$ bei gegebener Evidenz $e$
$d_{(min)(max)}^{S_i}$	(minimale/maximale) Dauer des Zustands $S_i$
$d_{Rest}^{S_i}$	maximale restliche Verweildauer im Zustand $S_i$
$d_{(min)(max)}^{T_{i,j}}$	(minimale/maximale) Dauer des Zustandsübergangs $T_{i,j}$
$D_X$	Wertebereich der diskreten Zufallsvariable $X$
$Cr(s)$	Glaubwürdigkeit (engl.: Credibility) der Aussage $s$
$e$	Evidenz
$e_X^-$	Evidenz des Teilbaums unterhalb des Bayes-Knoten $X$
$e_X^+$	Evidenz des Teilbaums oberhalb des Bayes-Knoten $X$
$E$	Meßwert eines Attributes
$\mathcal{F}_i$	charakteristische, im Bild detektierbare Merkmale eines Zustands
$f(\mathcal{N})$	Gesamtkosten des Pfades vom Startknoten zum Zielknoten über den Knoten $\mathcal{N}$ für $A^*$ -Algorithmus
$f_\varepsilon(\mathcal{N})$	Gesamtkosten des Pfades vom Startknoten zum Zielknoten über den Knoten $\mathcal{N}$ für $\varepsilon$ - $A^*$ -Algorithmus

---

$g(\mathcal{N})$	angefallene Kosten des Pfades vom Startknoten zum Knoten $\mathcal{N}$
$h(\mathcal{N})$	wahre Restkosten des Pfades vom Knoten $\mathcal{N}$ zum Zielknoten
$h^*(\mathcal{N})$	geschätzte Restkosten des Pfades vom Knoten $\mathcal{N}$ zum Zielknoten
$H$	Wertebereich bzw. erwarteter Meßwert eines Attributes
$I(n)$	Instanz des Konzeptes $n$
$I_F(n)$	fehlende Instanz, für die mindestens eine obligatorische Komponente fehlt
$I_H(n)$	Hypothese, die noch nicht verifiziert oder falsifiziert ist
$I_K(n)$	komplette Instanz mit allen obligatorischen kontextunabhängigen und kontextabhängigen Komponenten
$I_P(n)$	partielle Instanz mit allen obligatorischen kontextunabhängigen Komponenten
$\mathcal{M}_{x u}$	Matrix der bedingten Wahrscheinlichkeiten eines Bayes-Knoten $X$ mit seinen Elternknoten $U$
$\mathcal{N}$	Menge von Knoten, insbesondere Suchbaumknoten bestehend aus Konzepten und Instanzen inklusive ihrer Relationen
$Ne(s)$	Notwendigkeit (engl.: Necessity) der Aussage $s$
$P$	Wahrscheinlichkeit
$P(S_i)$	Auftretenswahrscheinlichkeit des Zustands $S_i$
$P(T_{i,j})=P(S_j S_i)$	Auftretenswahrscheinlichkeit des Zustandsübergangs $T_{i,j}$
$P(x)=P(X=x)$	diskrete a-priori-Wahrscheinlichkeitsverteilung der Zufallsvariable bzw. des Bayes-Knotens $X$
$P(x e)$	aktualisierte Wahrscheinlichkeit des Bayes-Knoten $X$ unter der Beobachtung der Evidenz $e$
$Pl(s)$	Plausibilität (engl.: Plausibility) der Aussage $s$
$po(s)$	Möglichkeitsverteilung
$Po(s)$	Möglichkeit (engl.: Possibility) der Aussage $s$
$r(n)$	Relation bezüglich Knoten $n$
$R(n)$	Regel, die auf den Knoten $n$ angewendet wird
$R_{AR:(inv)(cd)(opt)}$	Regel zur Generierung von Hypothesen über (optionale) kontext(un)abhängige attributierte Relationen (in inverser Richtung)

---

$R_{Binde}$	Regel zum Verschmelzen von zwei Instanzen, um fehlende Relationen mit Hilfe der Relationen der zweiten Instanz zu ergänzen
$R_F$	Regel zur Generierung fehlender Instanzen
$R_K$	Regel zur Generierung kompletter Instanzen
$R_H$	Regel zur datengetriebenen Generierung von Hypothesen ausgehend von obligatorischen kontextunabhängigen Komponenten
$R_{H:inv(cd)(opt)}$	Regel zur modellgetriebenen Generierung von Hypothesen für obligatorische (optionale) kontext(un)abhängige Komponenten
$R_{data-of}$	Regel zur initialen Instanzierung aus den Daten
$R_P$	Regel zur Generierung partieller Instanzen
$R_{PMT}$	Regel zur Generierung partieller Instanzen bei multitemporaler Bildanalyse
$R_{Spez}$	Regel zur Spezialisierung von kompletten Instanzen
$R_{Temp}$	Regel zur modellgetriebenen Generierung zeitlicher Nachfolger für komplette Instanzen über die temporale Relation
$s$	Aussage
$S_i$	Zustand eines Objekts bzw. endlichen Automaten
$S_0$	Startzustand eines endlichen Automaten
$S_\infty$	Endzustand eines endlichen Automaten
$\mathcal{P}$	Menge aller Zustände
$t_{A_{min}}^{S_i}$	frühester Anfangszeitpunkt des Zustands $S_i$
$t_{A_{min}}^{S_i}$	spätestester Endzeitpunkt des Zustands $S_i$
$t_0^{S_i}$	(vordefinierter) möglicher Anfangszeitpunkt des Zustands $S_i$
$T_{i,j}$	Zustandsübergang vom Zustand $S_i$ zum Zustand $S_j$
$U$	Menge von diskreten Zufallsvariablen
$U^i$	Elternknoten des Bayes-Knoten $X$
$w$	Gewicht (für Knoten bzw. Attribute des semantischen Netzes)
$+x/\neg x$	Zustände der binären Zufallsvariablen $X$

---

$X$	diskrete Zufallsvariable, Element der Menge $U$
$Y^j$	Kindknoten des Bayes-Knoten $X$
$a$	Normierungsvektor zur Berechnung der Glaubwürdigkeit $BEL(x)$
$\varepsilon$	maximale Abweichung der Pfadkosten von denen des optimalen Pfades
$\lambda(x)$	diagnostischer Support der Aussage $\{X=x\}$
$\pi(x)$	kausaler Support der Aussage $\{X=x\}$
$\Sigma$	Eingabealphabet eines endlichen Automaten
$\Theta(s)$	Unwissenheit bzgl. der Aussage $s$

### Graphische Symbolik

	Konzept $X$ des semantischen Netzes
	Hypothese $X$ des semantischen Netzes
	Instanz $X$ des semantischen Netzes
	Knoten $X$ des Bayes-Netzes

## Kurzfassung

Die vorliegende Arbeit behandelt die automatische Auswertung multitemporaler, d. h. zeitlich versetzt aufgenommener Luftbilder. Das vorgestellte Analysesystem weist den Bildobjekten für jeden Beobachtungszeitpunkt eine Bedeutung zu und erstellt so eine symbolische Beschreibung der gesamten Szene. Diese Szenenbeschreibung gestattet die Dokumentation von Landschaftsveränderungen auf einer semantischen Ebene sowie die Erkennung komplexer Objekte, die sich durch die zeitliche Abfolge charakteristischer Ereignisse auszeichnen. Die Arbeit liefert somit einen Beitrag zur rechnergestützten Landschaftsüberwachung und zur automatischen Fortführung von Karten.

Während der Luftbildauswertung wird Vorwissen über die Struktur, Geometrie und Topologie der erwarteten Objekte genutzt, wodurch sich die Interpretationsleistung steigern läßt. Diesem Prinzip folgend wird für die Analyse multitemporaler Luftbilder erstmalig Wissen über zeitliche Zusammenhänge verwandt, um die Erkenntnisse aus vorhergehenden Bildern nutzbar zu machen. Mit Hilfe einer temporalen Schlußfolgerung werden mögliche Landschaftsveränderungen prädiziert und unmögliche Veränderungen a priori ausgeschlossen.

Die Wissensinhalte werden explizit und maschinell verwertbar in Form eines semantischen Netzes repräsentiert. Die Modellierung zeitlicher Veränderungen erfolgt durch Zustandsübergangsdiagramme, die zusätzlich Informationen über die minimale und maximale Dauer und die Auftretenswahrscheinlichkeit eines Zustands bzw. Zustandsübergangs beinhalten. Bei der Prädiktion von Landschaftsveränderungen entstehen in der Regel konkurrierende Alternativen, die es nacheinander zu verifizieren gilt. Ein neuartiges, auf Bayes-Netzen beruhendes Bewertungssystem sorgt dafür, daß wahrscheinlichere Alternativen besser beurteilt und deshalb bevorzugt untersucht werden. Bei der Bewertung wird sowohl die diagnostische Unterstützung in Form von Beobachtungen in den Bilddaten als auch die kausale Unterstützung durch a priori bekannte Wahrscheinlichkeiten berücksichtigt.

Das vorgeschlagene Bildinterpretationssystem wird exemplarisch zur Detektion baulicher Veränderungen in einem Industriegebiet eingesetzt. Darüber hinaus wird demonstriert, wie sich ein Messegelände anhand typischer Auf- und Abbauaktivitäten aus einer Bildreihe erkennen läßt. Für beide Beispiele zeigt sich, daß das neue Bewertungsverfahren zu einer deutlichen Steigerung der Analyseeffizienz führt.

Stichworte: Wissensbasiertes System, Semantisches Netz, Bayes-Netz, multitemporal, Änderungsdetektion, Fernerkundung, Bildinterpretation

## Abstract

The presented work deals with the automatic interpretation of multitemporal aerial images, i.e. images from different acquisition times. The proposed analysis system assigns a semantic to the image objects for each observation time and thus builds a symbolic description of the entire scene. This scene description allows the documentation of landscape changes at a semantic level and the detection of complex objects characterized by a typical sequence of events over time. Hence the work contributes to the computer-aided monitoring of landscapes and the automatic update of maps.

During image interpretation a priori knowledge about the structure, the geometry and the topology of the expected objects is used, which increases the performance of the analysis. Following this principle knowledge about temporal aspects is exploited for the interpretation of multitemporal aerial images, in order to make usable the results from the preceding images. Using a temporal inference mechanism possible landscape changes are predicted and impossible changes are excluded right from the beginning.

The knowledge is represented explicitly and machine-usable within a semantic network. Temporal changes are modelled by state transition diagrams, which contain additional information about the minimum and maximum duration and the probability of states and state transitions. The prediction of landscape changes yield usually multiple competing alternatives, which have to be verified consecutively. A new judgement approach based on Bayesian networks ensures, that more probable alternatives are judged better and therefore are investigated first. The approach considers both, the diagnostic support derived from the observations in the image data and the causal support represented by the prior known probabilities.

The proposed image interpretation system is applied exemplarily to the detection of architectural changes in an industrial area. Furthermore the recognition of a fairground is demonstrated based on the detection of typical construction and dismantling activities in a multitemporal image series. For both examples the efficiency of the analysis is increased significantly using the new judgement approach.

Keywords: knowledge-based system, semantic net, Bayesian network,  
multitemporal, change detection, remote sensing, image interpretation



---

# 1. Einleitung

## 1.1 Einführung in die Thematik

Die Auswertung von Luftbildern blickt auf eine jahrzehntelange Geschichte zurück und wird traditionell für die Kartierung der Erdoberfläche genutzt. Mit der Verfügbarkeit von Satelliten und der Möglichkeit, Luftbilder in immer kürzeren Zeitabständen aufnehmen zu können, entstanden in den letzten Jahren neue Anwendungsgebiete von zunehmender wirtschaftlicher und politischer Bedeutung. So wird beispielsweise mit Hilfe von Luft- und Satellitenbildern die Landnutzung [74][128] und deren Änderung [32][70] großflächig erfaßt und die Einhaltung von Abrüstungsverträgen [33][162] kontrolliert. Aus Fernerkundungsdaten berechnete 3D-Landschaftsmodelle werden zu Planungs- und Simulationszwecken z. B. in der Stadt- und Landschaftsarchitektur [73][148] sowie der Telekommunikation [91][138] eingesetzt. Ein Hauptanwendungsgebiet der Luft- und Satellitenbilddauswertung ist die Erstellung und Fortführung von topographischen Karten [75][129].

Die traditionellen analogen Kartenwerke werden in zunehmendem Maße durch digitale und somit maschinell verarbeitbare Geo-Daten ersetzt, die in geographischen Datenbanken, sogenannten Geoinformationssystemen (GIS), verwaltet werden. Derartige Datenbestände bilden beispielsweise die Basis für Routenplaner und Navigationssysteme. Die genannten Anwendungen haben jedoch nur dann einen Nutzen, wenn die verwendeten Geo-Daten nicht veraltet oder unvollständig sind. Der hohe Qualitätsanspruch an die Datenbestände wirft das Problem ihrer Aktualisierung und Vervollständigung auf [63]. Als Informationsquellen kommen neben den Ergebnissen der terrestrischen Vermessung und des topographischen Meldedienstes in erster Linie Luft- und Satellitenbilder in Frage [7][38][57]. Die Bereitstellung hochaktueller Bilddaten stellt prinzipiell kein Problem dar. Sehr zeit- und kostenintensiv ist jedoch die manuelle Datenerfassung aus den Luftbildern. Aus diesem Grund wird vielerorts an Verfahren zur *automatischen Luftbilddauswertung* geforscht. Von besonderem Interesse sind neben der Klassifikation der Landnutzung die Erkennung künstlicher Objekte wie Straßen [11][159] oder Gebäude [22][39][50][58][85].

Die heutigen automatischen Bildanalyseverfahren reichen in der Regel nicht an die Interpretationsleistung eines menschlichen Auswerters heran. Dieser bewertet die komplexen Bildinhalte vor dem Hintergrund seines umfangreichen Wissens über die Landschaftsobjekte. Er verknüpft die Beobachtungen im Bild mit seinen Erwartungen bezüglich Form, Farbe, Struktur und Topologie der Objekte und bezieht auch Kontextinformationen wie Bildmaßstab und Aufnahmezeitpunkt in seinen Analysepro-

zeß mit ein. Anwendungen, wie z. B. die Erkennung von Landnutzungsänderungen, erfordern die vergleichende Untersuchung mehrerer, zeitlich nacheinander aufgenommener Bilder. Bei der Auswertung dieser *multitemporalen Luftbilder* ist ein Mensch in der Lage, relevante Veränderungen von irrelevanten zu unterscheiden, wie sie beispielsweise durch die Variation von Beleuchtung, Perspektive und Jahreszeit entstehen. Die Zeitabstände von einem Bild zum nächsten wird ein Auswerter nutzen, um seine Vermutungen bezüglich der Landschaftsveränderungen auf Plausibilität zu prüfen. Aufgrund seines Wissens über zeitliche Zusammenhänge kann er die Einzelbeobachtungen in den Bildern zu einem dynamischen Verhaltensmuster zusammenfügen und intelligente Schlußfolgerungen ziehen. All die genannten Fähigkeiten in einem System zur automatischen Luftbildauswertung zu vereinen, ist ein ehrgeiziges Ziel.

Die vorliegende Arbeit hat die automatische Auswertung multitemporaler Luftbilder zum Thema. Es sollen relevante Objekte im Bild erkannt und durch die Zuordnung einer symbolischen Bedeutung wie "Straße", "Wald" oder "Gebäude" klassifiziert werden. Dieser Vorgang wird als *Bildinterpretation* bezeichnet. In der Regel sind die Ergebnisse der heutigen Detektions- und Klassifikationsverfahren aufgrund der Komplexität der abgebildeten Bildinhalte fehlerhaft und unvollständig. Aus diesem Grund ist es sinnvoll, - einem menschlichen Auswerter ähnlich - Zusatzwissen über die erwarteten Objekte in den Interpretationsprozeß einzubringen. Dabei tritt das Problem auf, wie dieses Wissen zu strukturieren und im Rechner zu repräsentieren ist, so daß es für die automatische Bildinterpretation nutzbar wird.

Speziell für die Auswertung multitemporaler Luftbilder wird Wissen über zeitliche Zusammenhänge benötigt. Erst dann läßt sich die Information über die Zeitabstände zwischen den Bildern gewinnbringend nutzen. Mit Hilfe dieses temporalen Wissens können die Landschaftsveränderungen besser erkannt und interpretiert werden.

Um der großen Bandbreite möglicher Auswerteaufgaben gerecht werden zu können, ist eine Systemarchitektur mit größtmöglicher Flexibilität anzustreben. Dieses läßt sich durch die Trennung des anwendungsspezifischen Wissens von der allgemeingültigen Systemsteuerung erreichen, was in der Literatur als wissensbasierter Ansatz bezeichnet wird. Ein derartiger Ansatz wird in dieser Arbeit verfolgt.

## **1.2 Stand der Technik**

### **Auswertung multitemporaler Luft- und Satellitenbilder**

Multitemporale Luft- und Satellitenbilder werden zumeist zum Zwecke der Änderungsdetektion (engl.: Change Detection) ausgewertet. Dabei wird versucht, aus

---

mindestens zwei Bildern unterschiedlichen Aufnahmezeitpunkts relevante Landschaftsveränderungen zu extrahieren. Die Relevanz der Veränderungen ist anwendungsabhängig und fokussiert teilweise auf künstliche Objekte wie urbane Gebiete [9][140][160] und Verkehrswege [40] und teilweise auf Veränderungen der Vegetation [23][52][68] und deren Auswirkungen z. B. auf die Hydrologie [145]. Nur wenige Autoren versuchen in ihren Arbeiten, den Bildinhalt und die erkannten Änderungen automatisch zu interpretieren [14][24][26][32][86]. Die Mehrzahl der Veröffentlichungen beschränkt sich darauf, die Veränderungen im Bild nur zu detektieren und deren Interpretation einem menschlichen Auswerter zu überlassen. Verfahrenstechnisch lassen sich dabei die *Bild-zu-Bild-Ansätze* und der *nachträgliche Klassifikationsvergleich* (engl.: Post Classification Comparison) unterscheiden [30][139]:

Die *Bild-zu-Bild-Verfahren* vergleichen die Bilddaten direkt miteinander. Es werden Differenz (engl.: Image Differencing, Vegetation Index Differencing) [93][120] oder Verhältnis (engl.: Image Ratioing) [161] korrespondierender Bildpunkte gebildet oder der Fehler eines mittels linearer Regression prädierten Bildpunktwertes berechnet (engl.: Image Regression) [68][120]. Weitere Ansätze analysieren die vektorielle Änderung eines Bildpunkts im Spektralraum (engl.: Change Vector Analysis) [84][106] oder wenden eine Hauptachsentransformation (engl.: Principal Component Analysis) auf die Daten an [9][93][106]. Für jeden bildpunktweisen Vergleich von multitemporalen Daten ist neben einer exakten geometrischen Registrierung eine radiometrische Korrektur und Kalibrierung erforderlich, um die Helligkeitsunterschiede aufgrund variierender atmosphärischer Bedingungen, Beleuchtung oder Bodenfeuchtigkeit kompensieren zu können. Desweiteren können alle Bild-zu-Bild-Verfahren lediglich detektieren, *daß* eine Veränderung stattgefunden hat. Sie liefern keine Aussage darüber, *welcher Art* diese Änderung ist, was für eine automatische Bildinterpretation, wie sie hier angestrebt wird, zwingend erforderlich ist.

Beim *nachträglichen Klassifikationsvergleich* [41][93] werden die multitemporalen Bilder zunächst getrennt voneinander klassifiziert, d. h. jeder Bildpunkt wird genau einer Klasse aus einer vorgegebenen Menge relevanter Objektklassen zugeordnet. In einem zweiten Schritt werden die Klassifikationsergebnisse miteinander verglichen, um Veränderungen zu detektieren. Im Gegensatz zu den erstgenannten Verfahren erfolgt die vergleichende Untersuchung nicht auf Bildpunktebene, sondern auf einer semantischen Ebene. Die Ergebnisse liefern somit nicht nur Auskunft über den Ort der Landschaftsveränderung, sondern auch darüber, welcher Klasse die veränderte Region vorher und nachher angehörte. Da die Klassifikation für jedes Bild separat erfolgt, kann die aufwendige radiometrische Kalibrierung der Daten entfallen. Nachteilig ist jedoch,

daß die Resultate der Änderungsdetektion entscheidend von der Qualität der Klassifikationsergebnisse abhängt [147]. Diese läßt sich verbessern, indem Zusatzwissen z. B. über Objektform und -struktur in den Klassifikationsprozeß eingebracht wird [31][51][160]. Darüber hinaus wäre es vorteilhaft, die bereits bekannten Erkenntnisse aus den vorhergehenden Bildern zu nutzen, um veränderte und unveränderte Landschaftsteile vorherzusagen und anschließend die Prognosen im aktuellen Bild zu verifizieren. Für eine derartige modellgetriebene Änderungsdetektion ist zusätzliches Wissen über mögliche Landschaftsänderungen erforderlich.

Neben der Detektion von Veränderung können multitemporale Datensätze auch dazu genutzt werden, eine zuverlässigere oder detailliertere Gesamtinterpretation des Beobachtungsgebiets zu erstellen [98]. Bruzzone et al. [15] und Solberg et al. [141] verbessern das Klassifikationsergebnis beispielsweise dadurch, daß multitemporale Bilder unter Berücksichtigung von Klassenübergangswahrscheinlichkeiten gleichzeitig analysiert werden. Andere Autoren verwenden statistische Modelle [142][67] oder neuronale Netze [9][15][25] zur Fusion multitemporaler Bilder. Explizit formuliertes, symbolisches Wissen über Objektzustände und deren zeitliche Veränderungen wird bislang nicht benutzt. Damit wäre es möglich, komplexe Objektklassen zu beschreiben, die sich durch ein charakteristisches zeitliches Verhalten auszeichnen. So ließen sich z. B. speziell genutzte Flächen wie Messegelände oder Festplätze an Auf- und Abbauaktivitäten oder landwirtschaftliche Nutzflächen an Pflüge- und Erntetätigkeiten aus zeitlich aufeinanderfolgenden Luftbildern erkennen. Zur maschinellen Repräsentation und der automatischen Nutzung des dafür relevanten temporalen Wissens erscheint ein wissensbasierter Bildinterpretationsansatz besonders geeignet.

## **Wissensbasierte Bildinterpretation**

Als *wissensbasierte Bildinterpretation* wird der Prozeß des Bildverstehens bezeichnet, bei dem signifikante Objekte im Bild erkannt und mit einer symbolischen Bedeutung versehen werden.

Die erste Generation wissensbasierter Systeme zur Luftbildinterpretation stammt aus den achtziger Jahren. Die bekanntesten Vertreter sind VISIONS [53], ACRONYM [17], SPAM [96], MESSIE [21] und SIGMA [94]. Ihre Anwendungsgebiete liegen in der Erkennung von Einzelobjekten wie Flugzeugen, Straßen oder Gebäuden bzw. der Interpretation komplexer Flughafenszenen aus einem oder mehreren Luftbildern.

Allen wissensbasierten Systemen ist die explizite Modellierung des Vorwissens gemein. Sie unterscheiden sich jedoch in der gewählten Wissensrepräsentationsform. Während SPAM Regeln nutzt, verwenden die anderen genannten Systeme hierarchische Strukturen wie Graphen oder semantische Netze. Im BPI-System [92][146]

---

wird eine Bestandteilshierarchie von anwendungsspezifischen Produktionsregeln als Wissensbasis eingesetzt. Demgegenüber schlägt Mees [97] vor, strategisches Wissen in UND/ODER-Bäumen, globales Wissen in Fuzzy-Produktionsregeln und lokales Wissen in generischen Prototypen mit spezialisierten Bildverarbeitungsoperatoren zu repräsentieren. Das System ERNEST [101][124] und seine Modifikationen zur kartengestützten Luft- [113][114] und Satellitenbilddauswertung [83] verwenden semantische Netze zur Beschreibung des Vorwissens. ERNEST definiert eine allgemeine Netzwerksprache, so daß die Strukturierung und die Inhalte der Wissensbasis frei wählbar sind. Ein ähnlicher Ansatz wird im System AIDA [89][151][152] verfolgt, das über eine auf die Bildinterpretation zugeschnittene Netzwerksprache verfügt.

Alle genannten Systeme legen in der Wissensbasis Informationen über die 2D- oder 3D-Geometrie der relevanten Objekte ab, um sie im Bild gezielt detektieren zu können. Nur wenige Autoren trennen sensorabhängige und -unabhängige Wissensinhalte. Sie setzen voraus, daß lediglich Bilder eines einzelnen Sensor- oder Kamertyps zu analysieren sind. Durch die Fusion der Informationen aus unterschiedlichen Spektralbereichen kann die Interpretationsleistung jedoch erhöht werden. Das System MESSIE [21] repräsentiert explizit Wissen über Sensoren sowie über Geometrie und Material der Objekte, so daß die Bildanalyse sensorspezifisch erfolgen kann. Darüber hinausgehend werden in AIDA sensorspezifische Abbildungsmodelle für die Transformation zwischen Welt- und Bildkoordinaten genutzt. Karteninformationen oder Daten eines Geoinformationssystems werden in SPAM, MOSES [113], bei Kunz [83] und in AIDA ausgewertet.

Keines der beschriebenen Systeme ist in der Lage, multitemporale Bilder zu interpretieren. Das Potential, aus derartigen Bildreihen ein verbessertes Interpretationsergebnis zu erzielen, wurde bislang nicht genutzt. Lediglich für den Bereich der Änderungsdetektion werden Konzepte vorgestellt, um mit Hilfe eines wissensbasierten Ansatzes, Landschaftsveränderungen automatisch zu erkennen und zu deuten. Dai und Khorram [25][26] streben eine Detektion von veränderten Bildinhalten mittels neuronaler Netze an. Die erkannten Änderungen sollen in einem separaten Modul unter Zuhilfenahme von GIS- und anderen Zusatzdaten klassifiziert werden. Der vorgestellte Ansatz hat allerdings bislang nur konzeptionellen Charakter. Dreschler-Fischer et al. [32][86] schlagen die Nutzung einer Beschreibungslogik vor, mit der mögliche zeitliche Veränderungen eines Objekts, wie z. B. die Erweiterung einer Landebahn, repräsentiert werden können. Das System ist in der Lage, in einem Bild Veränderungen gegenüber einem aus einer Karte abgeleiteten Landschaftsmodell zu detektieren und ggf. zu

erklären. Die Auswertung einer gesamten Zeitreihe von Bildern ist allerdings nicht vorgesehen.

Die vorgestellten Systeme zur wissensbasierten Bildinterpretation eignen sich zumeist nur für die Auswertung einzelner Bilder. Wenige Ansätze verarbeiten mehrere Bilder gleichzeitig, wobei eventuelle Zeitabstände zwischen den Aufnahmen vernachlässigt werden. Für die Auswertung multitemporaler Luftbilder fehlen Ansätze, Wissen über zeitliche Zusammenhänge zu repräsentieren und dieses zusammen mit der Information über die Aufnahmezeitpunkte für die Bildinterpretation zu nutzen. Die vorliegende Arbeit soll hierzu einen Beitrag liefern.

### **1.3 Ziele und Aufbau der Arbeit**

In dieser Arbeit wird die automatische Interpretation multitemporaler Luftbilder unter Verwendung von temporalen Wissensinhalten angestrebt. Zur flexiblen Anpassung an verschiedenartige Auswerteaufgaben wird ein wissensbasierter Ansatz auf Basis des Bildinterpretationssystems AIDA favorisiert. Die Möglichkeiten der Wissensrepräsentation müssen dahingehend erweitert werden, daß das Vorwissen über Objekte und deren zeitliche Veränderungen anwendungsunabhängig beschrieben werden kann. Darüber hinaus ist darauf zu achten, daß sich das zeitliche Vorwissen maschinell für die Bildinterpretation nutzen läßt.

Die wissensbasierte Auswertung der multitemporalen Luftbilder soll eine verbesserte Änderungsdetektion ermöglichen. Im Gegensatz zu den bisherigen Verfahren des Klassifikationsvergleichs, soll die Analyse der einzelnen Luftbilder nicht unabhängig voneinander erfolgen. Das Ziel der Arbeit ist vielmehr, bei der Interpretation der Bildreihe die Erkenntnisse aus den vorhergehenden Bildern zu nutzen. Die bekannten Ergebnisse sollen zusammen mit dem Wissen über mögliche zeitliche Landschaftsveränderungen für eine Prädiktion herangezogen werden. Dabei können unmögliche Landschaftsveränderungen a priori ausgeschlossen werden, wodurch eine effizientere und qualitativ bessere Bildinterpretation zu erwarten ist.

Desweiteren soll das Potential multitemporaler Luftbilder genutzt werden, spezielle Objektklassen zu detektieren, die aufgrund einer zeitlichen Abfolge charakteristischer Ereignisse erkannt werden können. Dadurch erhält man eine detailliertere Beschreibung des Beobachtungsgebiets im Vergleich zu einer monotemporalen Bildauswertung.

In der Regel sind für ein Objekt vielfältige, mehr oder weniger wahrscheinliche zeitliche Veränderungen denkbar, so daß bei einer Prädiktion der möglichen Objektänderungen alternative Lösungen entstehen. Um eine möglichst effiziente Bildinterpretation zu gewährleisten, ist es plausibel, die wahrscheinlichste Alternative als erste zu

---

untersuchen und erst dann, wenn diese nicht verifiziert werden kann, die unwahrscheinlicheren Lösungen zu bearbeiten. Das Bildinterpretationssystem soll dahingehend erweitert werden.

Die Arbeit ist wie folgt gegliedert:

**Kapitel 2** beschreibt das in dieser Arbeit zugrundegelegte Konzept der wissensbasierten Luftbildinterpretation. Nach der Erläuterung grundlegender Begriffe wird die Architektur sowie die bisherige Form der Wissensrepräsentation und Wissensnutzung des Interpretationssystems AIDA vorgestellt.

In **Kapitel 3** werden zunächst die Anforderungen an ein System zur multitemporalen Bildanalyse herausgearbeitet. Einige Ansätze aus der Literatur aufgreifend wird schließlich ein allgemeingültiger Formalismus zur Repräsentation zeitlicher Zusammenhänge entworfen und in das System AIDA integriert, mit dem ein temporales Schlußfolgern ermöglicht wird.

**Kapitel 4** hat die effiziente Steuerung des Interpretationsprozesses zum Thema. Es wird ein neuartiges Verfahren für die Bewertung alternativer Bildinterpretationen vorgeschlagen. Dieses Verfahren ermöglicht es, unter ansonsten vergleichbaren Lösungen die a priori wahrscheinlichste für die weitere Analyse auszuwählen.

**Kapitel 5** beinhaltet zwei Anwendungsbeispiele für das in Kapitel 3 vorgestellte System zur Interpretation multitemporaler Luftbilder. Anhand einer geeigneten Bildreihe wird demonstriert, wie ein Messegelände anhand typischer Auf- und Abbauaktivitäten automatisch erkannt werden kann. Das zweite Beispiel behandelt das Thema der Änderungsdetektion. Für beide Anwendungen wird aufgezeigt, daß sich das in Kapitel 4 entwickelte Bewertungsverfahren effizienzsteigernd gegenüber dem bisherigen Ansatz auswirkt.

Die Arbeit schließt mit einer Zusammenfassung in **Kapitel 6**.

## 2. Wissensbasierte Interpretation multisensorieller Luftbilder

In der vorliegenden Arbeit soll ein System entworfen werden, das eine Reihe von Luftbildern aus unterschiedlichen Aufnahmezeitpunkten automatisch auswertet und dabei temporales Wissen für Anwendungen wie die Änderungsdetektion oder die Erkennung von komplexen Objekten nutzt. Explizit formulierbares, temporales Wissen, wie z. B. *“Laubbäume tragen im Winter keine Blätter.”* bezieht sich stets auf Begriffe mit symbolischer Bedeutung wie der *Laubbaum* im genannten Beispiel. Über rein geometrische Begriffe, wie z. B. einen *Kreis*, existiert kein allgemeingültiges temporales Wissen. Will man zeitliche Wissensinhalte bei einer Luftbildauswertung nutzen, so ist es demnach zunächst erforderlich, den abgebildeten Objekten eine symbolische Bedeutung zuzuweisen, so daß z. B. ein detektierter *Kreis* die Bedeutung *Laubbaum* erhält. Erst dann kann das temporale Wissen auf die erkannten Bildinhalte übertragen werden. Dieser Vorgang der Bedeutungszuweisung wird *Interpretation* eines Bildes genannt.

Mit AIDA [151][152] wurde bereits ein System zur Interpretation von Luftbildern vorgestellt. Es gestattet die Analyse und dreidimensionale Rekonstruktion einer durch Luftbilder aufgenommenen Landschaftsszene. Die vorliegende Arbeit erweitert den dort vorgestellten Ansatz um die Nutzung temporalen Wissens. Zur Erläuterung der notwendigen Grundlagen werden im folgenden zunächst einige Begriffe definiert und anschließend das Konzept der wissensbasierten Bildinterpretation vorgestellt. Dabei wird auf die Darstellung der 3D-Rekonstruktion von Landschaften verzichtet und nur auf den Interpretationsteil eingegangen.

### 2.1 Definition grundlegender Begriffe

Zum besseren Verständnis werden an dieser Stelle die Begriffsdefinitionen aus [152] weitestgehend übernommen. Demnach besteht eine *Landschaftsszene* aus dreidimensionalen Körpern, die auf der Erdoberfläche räumlich angeordnet sind. Ein oder mehrere *Sensoren* erfassen die Szene aus der Vogelperspektive und liefern ein zweidimensionales Bild der dreidimensionalen Welt.

In der Fernerkundung werden unterschiedliche bildgebende Sensortypen eingesetzt. Neben den vorhandenen Satelliten (z. B. Landsat TM, SPOT, ERS-1/2) werden fluggestützte Plattformen zur Erdbeobachtung genutzt. Großbildkameras (z. B. Reihenmeßkamera der Firma Zeiss) liefern je nach eingelegtem Filmmaterial Bilder aus dem visuellen oder dem nahen Infrarot-Spektrum. Thermische Infrarot-Sensoren

---

zeichnen die Wärmestrahlung der Erdoberfläche auf. Darüber hinaus existieren Multi- und Hyperspektral-Sensoren, die in mehreren Spektralbereichen gleichzeitig arbeiten. Neben diesen passiven Sensoren kommen aktive Sensoren zum Einsatz, die selbst ein Signal aussenden und das von der Erdoberfläche reflektierte Signal aufzeichnen. Ein Beispiel ist der im Mikrowellenbereich arbeitende SAR-Sensor (engl.: Synthetic Aperture Radar). Durch Auswertung der Signallaufzeit und -amplitude lassen sich die Reflexionseigenschaften der Erdoberfläche in eine bildhafte Darstellung umrechnen. Der Laserscanner hingegen tastet die Landschaft mit einem Laserstrahl ab und gewinnt aus der gemessenen Signallaufzeit Höheninformationen zu jedem Geländepunkt. Die Meßwerte werden in ein regelmäßiges Raster umgerechnet und liefern so ein digitales Höhenmodell (DHM) der Landschaft.

Alle Bildprodukte von fluggestützten Sensoren werden in der vorliegenden Arbeit als *Luftbild* bezeichnet und somit vom *Satellitenbild* abgegrenzt. Werden zur Analyse einer Landschaft mehrere Bilder verschiedener Sensortypen verarbeitet, handelt es sich um *multisensorielle Luftbilder*. So ist es z. B. denkbar, daß sowohl ein Bild einer Reihenmeßkamera (RMK) als auch ein SAR-Bild derselben Region während der Analyse ausgewertet werden. Bilder mit unterschiedlichen Aufnahmezeitpunkten werden als *multitemporale Luftbilder* bezeichnet. Eine Zeitreihe von RMK-Bildern ist demnach multitemporal, jedoch nicht multisensoriell. Im Gegensatz dazu handelt es sich bei RMK- und SAR-Bildern einer Region um multisensorielle und multitemporale Luftbilder, falls sie zu unterschiedlichen Zeiten aufgenommen wurden.

Ziel der rechnergestützten *Bildinterpretation* ist es, automatisch eine *symbolische Beschreibung* der abgebildeten Szene (hier: Landschaftsszene) zu erstellen (Bild 2.1). Dabei soll den Objekten der Szene eine Bedeutung zugeordnet werden. Da es nicht möglich ist, diese Bedeutungszuordnung bildpunktweise vorzunehmen, erfolgt zunächst eine *Segmentierung* der Bilder. Geeignete Bildverarbeitungsoperatoren gruppieren die Bildpunkte zu *Bildprimitiven* wie Regionen oder Konturen. Als Gruppierungskriterium werden z. B. die Grau- oder Farbwerte eines Bildpunktes und seiner Nachbarschaft herangezogen. Die Bildprimitive lassen sich durch Form- und Farbmerkmale beschreiben, haben in der Regel jedoch keine Struktur oder Semantik. Die Gruppierung zu komplexeren Strukturen und die Bedeutungszuweisung erfolgt während der *Interpretation* der Bildprimitive. Durch Vergleich von gespeichertem Modellwissen mit den Eigenschaften der Bildprimitive und den aus ihnen gebildeten Strukturen wird eine Zuordnung zu semantischen Begriffen möglich, die schließlich zu einer symbolischen Beschreibung der gesamten abgebildeten Szene führt.

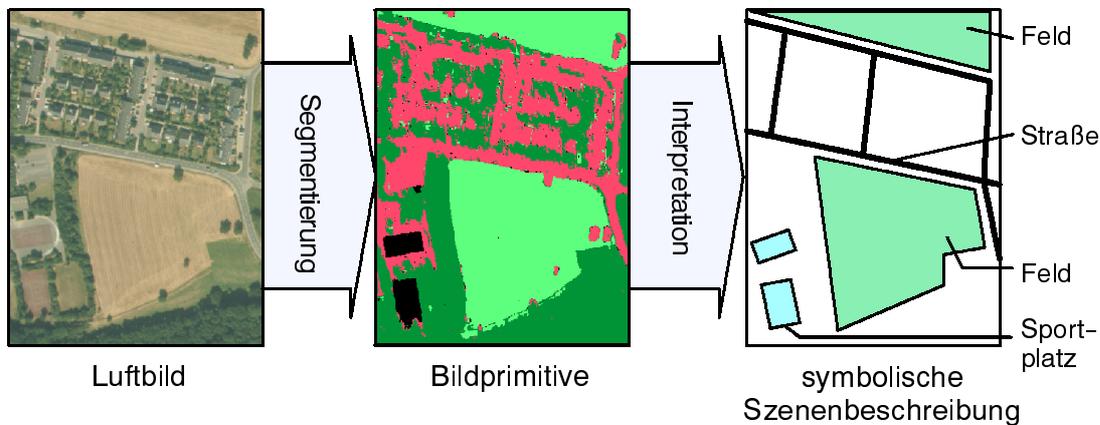


Bild 2.1: Erzeugung einer symbolischen Szenenbeschreibung durch Segmentierung und Interpretation eines Luftbilds

Entscheidend für die Qualität des Interpretationsergebnisses ist der Umfang und die Güte des abgelegten Modellwissens. Umfaßt das Wissen z. B. lediglich ein Modell für das Objekt *Straße*, so können nur Straßen, aber keine Wälder im Luftbild erkannt werden. Da einerseits nicht die ganze Welt im Rechner modellierbar ist und andererseits für eine bestimmte Fragestellung nicht immer alle Objektklassen relevant sind, ist es naheliegend, sich einzuschränken und das Modellwissen anwendungsspezifisch zu definieren. Wollte man z. B. aus Luftbildern das Straßennetz extrahieren, so ist es nicht erforderlich, ein Modell für das Objekt *See* zu definieren. Man wird sich auf Modelle der Objekte *Straße* und *Straßennetz* beschränken.

Die meisten Objekte in der Natur lassen sich in ihrem Erscheinungsbild nicht exakt beschreiben. Straßen haben nicht immer dieselbe Breite, Seen sind unterschiedlich groß. Trotzdem ist der Mensch in der Lage, jeden See und jede Straße zu erkennen, da er über Erfahrungswissen verfügt und sich aus seinen zahlreichen Beobachtungen ein *generisches Modell* abgeleitet hat. Ein rechnergestütztes Interpretationssystem sollte analog vorgehen und ein *generisches Szenenmodell* nutzen. So läßt sich die Straßenbreite z. B. durch einen zulässigen Wertebereich definieren, der die auftretenden Toleranzen berücksichtigt. Weiterhin lassen sich sinnvolle Annahmen über die maximale Krümmung einer Straße machen. Neben diesem *objektspezifischen Wissen* läßt sich *strukturelles Wissen* nutzen. Darunter versteht man Informationen über den Aufbau eines komplexen Objektes aus mehreren Teilobjekten. So besteht ein Straßennetz aus einer Menge von Straßen, die an Straßenkreuzungen miteinander verbunden sind. *Topologisches Wissen* beinhaltet Annahmen über Beziehungen zwischen benachbarten Objekten wie z. B. die Tatsache, daß Häuser in der Regel in der Nähe von Straßen

---

stehen, die als Zufahrtsweg dienen. Schließlich läßt sich *temporales Wissen* in eine Interpretation einbringen. Darunter fällt zum einen das Wissen über das (jahres-) zeitlich veränderte Erscheinungsbild bestimmter Objektklassen (z. B. die Blattfarbe von Laubbäumen im Herbst oder die höhere Wahrscheinlichkeit von Schnee im Winter etc.). Zum anderen lassen sich Erfahrungen über zeitliche Vorgänge formulieren. So ist es z. B. äußerst unwahrscheinlich, daß an einer Stelle, wo ein Haus detektiert wurde, kurze Zeit später ein Wald entstanden ist. Je umfangreicher das eingebrachte Wissen während der Interpretation ist, desto leichter lassen sich Unzulänglichkeiten der Bildverarbeitung ausgleichen und Widersprüche bei der Bedeutungszuweisung auflösen.

Neben den klassischen Verfahren der *strukturellen Mustererkennung* wie Relaxation, dynamische Programmierung oder Graphsuchverfahren werden in zunehmendem Maße Techniken der *künstlichen Intelligenz* zur Lösung des Zuordnungsproblems eingesetzt. Die dort vorgeschlagenen Schlußfolgerungsverfahren ermöglichen es, flexibel auf Konflikte und Widersprüche zu reagieren, die unweigerlich bei der Interpretation komplexer Szenen auftreten. Besonders geeignet für sich ändernde Aufgabenstellungen sind die *wissensbasierten Systeme* [54][90]. Sie zeichnen sich dadurch aus, daß das (anwendungsabhängige) Wissen explizit in einer Wissensbasis formuliert ist. Eine anwendungsunabhängige Steuerung nutzt das Wissen für den Analyseprozeß. Durch Austausch der Wissensbasis läßt sich das System leicht an andere Anwendungen anpassen. Besonderes Augenmerk ist auf die Frage zu lenken, wie das Wissen über Szeneninhalte und Analysestrategien repräsentiert und genutzt werden kann.

Die vorliegende Arbeit stellt eine Weiterentwicklung des Systems AIDA dar. Aus diesem Grund diskutieren die folgenden Abschnitte zunächst die Architektur und die Form der Wissensrepräsentation und Wissensnutzung des Systems, wie es Tönjes in [152] vorschlägt.

## 2.2 Systemarchitektur

Bild 2.2 zeigt die Architektur des wissensbasierten Bildinterpretationssystems AIDA. Multisensorielle Luftbilder werden automatisch ausgewertet, um eine symbolische Beschreibung der abgebildeten Landschaftsszene zu erhalten. In [152] wird diese Szenenbeschreibung genutzt, um eine realitätsnahe 3D-Rekonstruktion der Landschaft für Simulations- und Visualisierungszwecke zu erzeugen. Dabei versucht man, aus zweidimensionalen Luftbildern die dreidimensionale Form der Szenenobjekte zurückzugewinnen. Durch die Bedeutungszuweisung während der Interpretationsphase wird es möglich, bei der Rekonstruktion der Szene zusätzliches Vorwissen über die

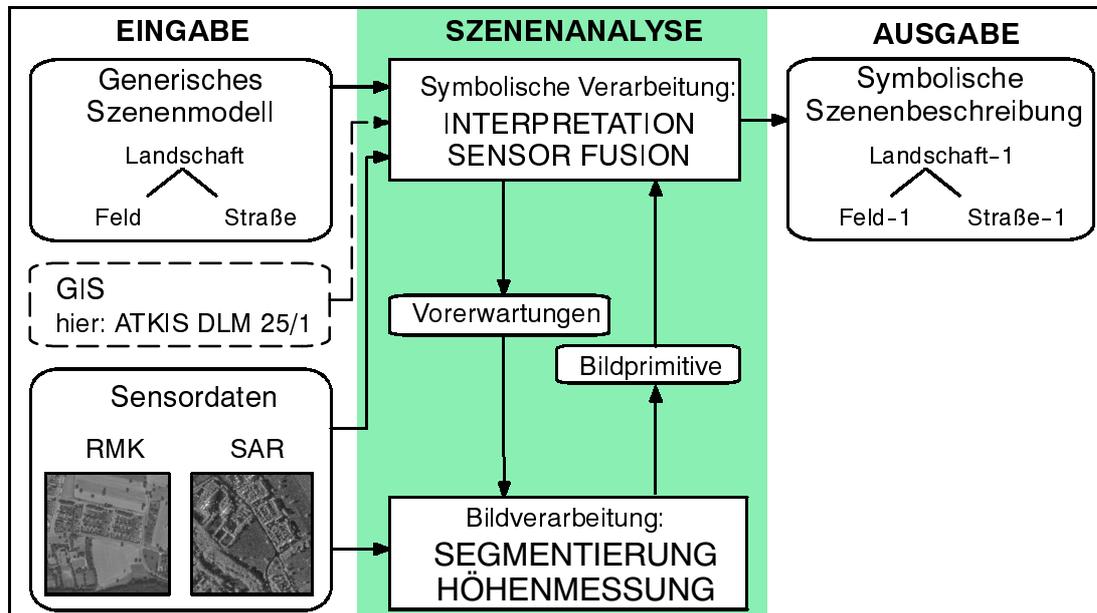


Bild 2.2: Architektur des wissensbasierten Bildinterpretationssystems AIDA

Objektgeometrie einzubringen. Allein aus der Kenntnis der Objektbedeutung können z. B. Straßen eben modelliert oder Höhengsprünge an Waldkanten eingefügt werden, auch wenn diese Informationen nicht in den Senkrechtaufnahmen enthalten sind [154]. Unzureichende Datenqualität und Mehrdeutigkeiten bei der Gewinnung von Höheninformationen lassen sich so in gewissem Rahmen kompensieren. Da jedoch der Rekonstruktionsteil für die vorliegende Arbeit nicht relevant ist, wird an dieser Stelle auf eine nähere Darstellung verzichtet.

Die Eingangsdaten für das Interpretationssystem bestehen aus den multisensoriellen Luftbildern und dem Vorwissen über die Szene. Bei den Bilddaten wird vorausgesetzt, daß die Abbildungsparameter der Sensoren bekannt und die Bilder in einem gemeinsamen (Geo-)Koordinatensystem registriert sind. Das Vorwissen setzt sich zum einen aus dem generischen Szenenmodell mit allgemeingültigem Wissen über die Szenenobjekte und zum anderen - falls vorhanden - aus szenenspezifischen Informationen eines Geoinformationssystems (GIS) zusammen. Derartige Systeme enthalten maschinell verarbeitbare Daten, die u.a. die geographische Lage ausgewählter Objektarten, wie z. B. Straßen und Flüsse, beschreiben. Die GIS-Daten repräsentieren bereits eine symbolische Szenenbeschreibung, die im Vergleich zum Luftbild jedoch unvollständig und veraltet sein kann. Trotzdem liefert das GIS zuverlässige Hinweise für eine erwartungsgesteuerte Bildanalyse. Das vorgestellte Interpretationssystem nutzt Daten des Digitalen Landschaftsmodells DLM 25 des deutschen Amtlich Topogra-

---

phisch-Kartographischen Informationssystem (ATKIS) [1][49]. Die GIS-Daten vereinfachen den Interpretationsprozeß und erhöhen die Vollständigkeit und Korrektheit der symbolischen Szenenbeschreibung. Da geeignete GIS-Daten jedoch nicht flächendeckend vorhanden sind, kann eine Interpretation auch allein auf Basis des generischen Szenenmodells erfolgen.

Im Bildverarbeitungsmodul erfolgt die Segmentierung der Sensordaten und die anschließende Extraktion von Bildprimitiven wie Regionen und Konturen. Gegebenenfalls werden auch lokale Merkmale der Bildprimitive wie z. B. Flächeninhalt, Hauptachse oder Konturlänge berechnet. Das Interpretationsmodul ist für die symbolische Verarbeitung zuständig. Es fusioniert korrespondierende Informationen aus den verschiedenen Sensoren, gruppiert die Bildprimitive zu komplexeren Strukturen und weist ihnen eine symbolische Bedeutung zu. Neben dieser datengetriebenen (*bottom-up*) Strategie kann die Analyse auch modellgetrieben (*top-down*) erfolgen. Dabei leitet das System aus dem generischen Szenenmodell Hypothesen über die Existenz, Form, Farbe und Lage der gesuchten Objekt ab. Diese Vorerwartungen werden anschließend in den Sensordaten verifiziert. Die intelligente Systemsteuerung verfügt über verschiedene Analysestrategien, mit denen gezielt nach fehlenden Objektteilen gesucht und flexibel auf Konflikte reagiert werden kann. Das System liefert schließlich eine symbolische Beschreibung der beobachteten Szene, die den Bildinhalt mit dem vorhandenen Wortschatz des generischen Szenenmodells am besten wiedergibt. Objektklassen, die nicht im Vorwissen enthalten sind, können prinzipbedingt nicht erkannt werden.

Das generische Szenenmodell ist - wie für wissensbasierte Systeme typisch - explizit in einer Wissensbasis repräsentiert. Durch Austausch dieser Wissensbasis lassen sich auf einfache Weise unterschiedliche Anwendungsgebiete realisieren. Das System wurde erfolgreich für die Detektion von Straßen [47][153] und die Erkennung komplexer Objekte wie ein Klärwerk [154] aus multisensoriellen Luftbildern, sowie im Gebiet der Nahbereichsphotogrammetrie für die Rekonstruktion von Gebäuden aus Stereoaufnahmen [46] eingesetzt.

Von besonderem Interesse bei jedem wissensbasierten System ist die Frage, welche Wissensinhalte auf welche Weise in der Wissensbasis repräsentiert sind und wie diese automatisch für einen Informationsverarbeitungsprozeß genutzt werden. Antworten liefern die folgenden Abschnitte.

## 2.3 Repräsentation des generischen Szenenmodells

Für die automatische Interpretation multisensorieller Luftbilder benötigt ein Szenenanalyse-System Vorwissen über die möglichen Szenenobjekte und die Abbildungseigenschaften der eingesetzten Sensoren. Neben reinem deklarativen Faktenwissen ist zusätzlich prozedurales Wissen zur Nutzung dieser Fakten erforderlich. In der Literatur sind zahlreiche Formalismen zur expliziten Repräsentation von Wissensinhalten vorgeschlagen worden [163]. Beispielhaft seien hier die Prädikatenlogik, Regeln und semantische Netze genannt.

Die *Prädikatenlogik* stellt eine formale Sprache dar und gestattet die Darstellung von Begriffen und Beziehungen zwischen diesen. Als Wissensrepräsentationsform für ein Bildinterpretationssystem erweist sie sich jedoch zu starr, da sich z. B. Unsicherheiten nur schwer beschreiben lassen.

*Regeln* oder auch *Produktionen* formulieren Wissensinhalte in einer Wenn-Dann-Beziehung. Falls der Bedingungsteil erfüllt ist, wird der Aktionsteil der Regel ausgeführt. Trifft dieses auf mehrere Regeln der Regelbasis zur gleichen Zeit zu, so muß eine übergeordnete Steuerung, die sog. Inferenzmaschine, mit Hilfe einer Konfliktlösungsstrategie eine Regel auswählen. Große Regelbasen verlangsamen diesen Auswahlprozeß und werden zudem schnell unübersichtlich. Zahlreiche aus der Literatur bekannte wissensbasierte Interpretationssysteme wie ACRONYM [17], BPI [92][146], ERNEST [101][124], MESSIE [21], MOSES [113][114], SIGMA [94] und SPAM [96] nutzen Regeln zur Steuerung des Analyseablaufs. In den Systemen SPAM und BPI ist ferner das gesamte Szenenwissen in Form von Regeln bzw. Produktionen formuliert. Die Systeme ERNEST, MOSES, SIGMA und VISIONS [53] beschreiben das szenenspezifische Wissen mit Hilfe *semantischer Netze*, auf die der folgende Abschnitt näher eingeht.

### 2.3.1 Semantische Netze für die Wissensrepräsentation

Zur anschaulichen Modellierung von Objekten und ihren Beziehungen untereinander eignen sich in besonderem Maße Graphen. Eine Form der gerichteten Graphen sind die semantischen Netze. Ihre Knoten repräsentieren Begriffe oder Sachverhalte und ihre Kanten modellieren die Relationen zwischen den Begriffen. Ein einfaches Beispiel für ein semantisches Netz zeigt Bild 2.3. Die explizite Repräsentation in Knoten und Kanten erleichtert die Strukturierung umfangreicher Wissensinhalte. Zudem lassen sich semantische Netze – im Gegensatz zu einer Regelbasis – für den Menschen anschaulich visualisieren.

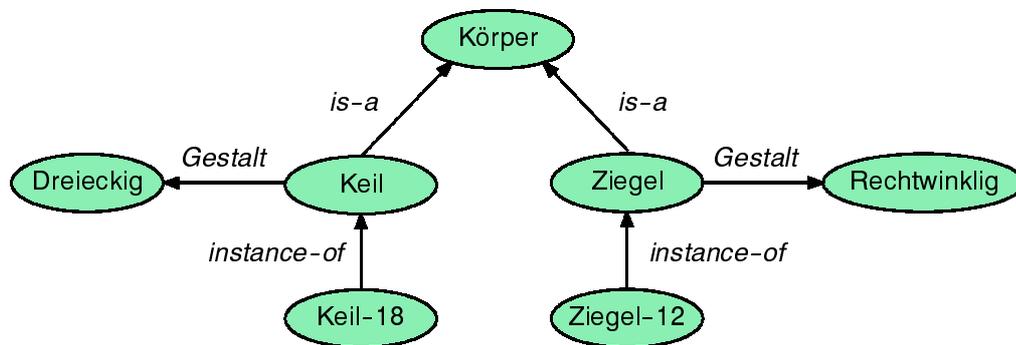


Bild 2.3: *Beispiel für ein semantisches Netz aus [163]: Die Kanten repräsentieren Klassenhierarchie (is-a), Objektform (Gestalt) bzw. die Relation zwischen Objektmodell und -instanz (instance-of).*

Prinzipiell kann den Kanten eines semantischen Netzes eine beliebige Bedeutung zugewiesen werden. Für eine automatisierte Verarbeitung hat es sich jedoch als sinnvoll erwiesen, die möglichen Kantentypen auf einige wenige zu beschränken. Das System SIGMA benutzt Kanten zur Darstellung von Klassenhierarchien (*is-a*), zur Zerlegung komplexer Begriffe in seine Bestandteile (*part-of*) und zur Beschreibung der Abbildung vom Szenenbereich in den Bildbereich (*appearance-of*). In ERNEST kommen ähnliche Kanten zum Einsatz, deren Richtungen im Vergleich zu anderen Systemen vertauscht sind und vom Vaterknoten zum Kindknoten gerichtet sind. Die *Spezialisierungskante* beschreibt die Ableitung eines speziellen Begriffs von einem allgemeineren. Die Dekomposition eines Objektes erfolgt über die *Bestandteilkante*. Die *Konkretisierungskante* verbindet Knoten aus unterschiedlichen Begriffswelten miteinander, wie z. B. Begriffe aus dem Szenenbereich mit solchen aus dem Bildbereich. Sie stellt somit eine Verallgemeinerung der *appearance-of*-Kante aus SIGMA dar.

Das an dieser Stelle betrachtete Interpretationssystem AIDA [89][151][152] verwendet ebenfalls semantische Netze zur Darstellung des deklarativen Wissens. Die prozeduralen Wissensinhalte werden in Regeln und Methoden beschrieben. Zur Definition des generischen Szenenmodells ist vor Beginn der Bildanalyse ein semantisches Netz zu definieren, auf dessen Basis während des Analyseprozesses nach und nach eine symbolische Beschreibung des Bildinhalts erstellt wird. AIDA stellt eine Netzwerksprache zur Verfügung, die die Formulierung des problemspezifischen Wissens erleichtert. Diese Netzwerksprache lehnt sich in großen Teilen an die im System ERNEST eingesetzte Syntax an und erweitert sie in einigen Punkten. Im

folgenden werden die wichtigsten Knoten- und Kantentypen vorgestellt. Für eine ausführliche Darstellung sei auf [151][152] verwiesen.

### 2.3.1.1 Knoten

Die Objekte in der Szene und im Bild werden durch Knoten im semantischen Netz repräsentiert. Es wird zwischen zwei Knotentypen unterschieden: *Konzepte* sind generische Modelle der Objekte und werden vor Beginn der Analyse in der Wissensbasis definiert. Spezifische Realisierungen der Objekte, die in der Szene beobachtet worden sind, werden durch *Instanzen* der korrespondierenden Konzepte im semantischen Netz dokumentiert. Während es in der Wissensbasis z. B. lediglich ein Konzept des Objekts *Straße* existiert, können im Bild mehrere *Straßen*-Instanzen vorhanden sein. Der Prozeß zur Erzeugung einer Instanz wird *Instanzierung* genannt.

Somit besteht die Wissensbasis ausschließlich aus Konzepten, während die symbolische Szenenbeschreibung durch ein Netz von Instanzen gegeben ist. Je nach Interpretationsstand können Instanzen eines Knoten  $n$  den Status *Hypothese*  $I_H(n)$ , *partielle Instanz*  $I_P(n)$ , *komplette Instanz*  $I_K(n)$  oder *fehlende Instanz*  $I_F(n)$  einnehmen. Hypothesen sind Vermutungen, die aus der Wissensbasis oder aus der bisherigen Szenenbeschreibung abgeleitet, aber noch nicht durch Daten bestätigt wurden. Konnte eine Hypothese teilweise in den Daten verifiziert werden, wechselt der Status zur partiellen Instanz. Wurden ein Objekt vollständig in seiner Struktur detektiert, so handelt es sich um eine komplette Instanz. Mit fehlenden Instanzen werden solche Hypothesen dokumentiert, die in den Daten nicht verifiziert werden konnten.

Jedes Konzept verfügt über eine Reihe von eigenen Methoden. Instanzen besitzen keine eigenen Methoden, sondern sie benutzen die ihrer Konzepte. Für Objekte, die direkt aus den Daten gewonnen werden können, läßt sich eine *Segmentierungsmethode* angeben, die entsprechende Bildprimitive aus den Sensordaten extrahiert. Eine *Bewertungsmethode* dient zur Beurteilung, inwieweit die aus der Wissensbasis abgeleiteten Vorerwartungen mit den Messungen im Bild übereinstimmen. Über ein *Knotengewicht* läßt sich einstellen, in welchem Maße der Knoten in die Bewertung der Szenenbeschreibung eingeht. Relevante Objekteigenschaften werden in einer Menge von *Attributen* beschrieben.

### 2.3.1.2 Attribute

Die spezifischen Eigenschaften der Objekte werden in Attributen repräsentiert. Beispielsweise läßt sich eine Straße durch ihre Länge und Breite charakterisieren. Attribute besitzen einen *Attributwert*  $E$  (für: Evidenz), der aus den Sensordaten oder den

---

Attributen untergeordneter Knoten bestimmt wird, und einen *Wertebereich H* (für: Hypothese), der dem erwarteten Attributwert entspricht. Die Wertebereiche der Attribute stellen Expertenwissen dar und sind vor Beginn der Analyse in der Wissensbasis zu definieren. Gegebenenfalls können diese Erwartungsbereiche während der Analyse weiter eingeschränkt werden. Sowohl Attributwert als auch Wertebereich sind in der Regel ungenaue Größen und werden deshalb durch Intervalle repräsentiert.

Zur Bestimmung des Attributwerts und des Wertebereichs existieren entsprechende *Berechnungsmethoden*, die während der Analyse aktiviert werden. Mit ihnen können sowohl Messungen in den Sensordaten als auch funktionale Zusammenhänge zwischen den Attributen formuliert werden (z. B. die Straßenbreite ergibt sich aus der Summe der einzelnen Fahrbahnbreiten). Die Argumente der Methoden (hier: Fahrbahnbreite) müssen zunächst allgemein beschrieben und im Moment der Berechnung mit den aktuellen Werten gefüllt werden. Eine Pfadgrammatik [151] erleichtert dabei den Zugriff auf benachbarte Knoten und Attribute.

Ähnlich wie die Knoten des semantischen Netzes besitzen auch die Attribute eine *Bewertungsmethode*. Ihre Aufgabe ist die Beurteilung, inwieweit der gemessene Attributwert mit dem Erwartungsbereich übereinstimmt. Das Ergebnis der Bewertung wird im Attribut dokumentiert. Aus allen Attributbewertungen eines Knotens wird mittels der Bewertungsmethode des Knotens (s. Kapitel 2.3.1.1) eine Knotenbewertung bestimmt. Dabei geht jedes Attribut mit einem vom Benutzer definierbaren *Attributgewicht* ein. Der Anwender kann bei Bedarf die Bewertungsmethoden selbst angeben. Anderenfalls werden die für das gewählte Bewertungsverfahren (s. Kapitel 2.4.3 und Kapitel 4) vordefinierten Standardmethoden genutzt.

### 2.3.1.3 Kanten

Die Knoten werden durch wenige vordefinierte Kantentypen zueinander in Beziehung gesetzt. Die Netzwerkkanten werden auch als *Relationen* bezeichnet. Jede Instanz ist über die *instance-of*-Kante mit dem zugehörigen Konzept verbunden, um den Bezug zwischen Wissensbasis und symbolischer Szenenbeschreibung herzustellen. Darüber hinaus existieren hierarchische und topologische Kanten im semantischen Netz.

#### Hierarchische Kanten

Oftmals läßt sich die Erkennung eines komplexen Musters auf die Detektion mehrerer Teilobjekte abbilden, die aufgrund ihrer einfacheren Struktur leichter aus den Sensordaten zu extrahieren sind. Diese Dekomposition in Teilobjekte wird mit Hilfe der *part-of*-Kante realisiert. Manche Bestandteile eines Objektes sind zur Erkennung des Ganzen zwingend notwendig, während andere nicht unbedingt vorhanden sein müssen.

Derartige obligatorische bzw. optionale Bestandteile lassen sich in der *part-of*-Kante effizient über die Angabe der minimalen und maximalen Kantenzahl definieren. Das Minimum gibt die Zahl der obligatorischen, die Differenz zwischen Maximum und Minimum die Zahl der optionalen Teile an. Neben dem Fall, daß ein Vaterknoten mehrere Kindknoten besitzt, ist es auch möglich, daß ein Kindknoten Bestandteil mehrerer Vaterknoten ist und somit mehrfach gebunden wird. Beispielsweise ist eine Mittelleitplanke einer Fernstraße Teil beider Richtungsfahrbahnen. Derartige Konstellationen können über die Angabe der minimalen und maximalen Bindungszahl definiert werden.

In der Regel werden Objekte in den Sensordaten aufgrund der Form- und Farbmerkmale ihres Abbildes erkannt. Somit muß das generische Szenenmodell Begriffe mit symbolischer Bedeutung (z. B. Straße) mit konkreten geometrischen und photometrischen Begriffen (z. B. Linie) in Beziehung setzen. Die Verbindung dieser unterschiedlichen Begriffswelten - *konzeptionelle Ebenen* genannt - erfolgt ausschließlich über die *concrete-of*-Kante, abgekürzt *con-of*. Bestandteilkanten sind hingegen nur innerhalb derselben konzeptionellen Ebene zugelassen. Dadurch wird die Wissensbasis in separierbare Schichten organisiert. Die im Bild segmentierbaren Primitive werden im semantischen Netz ebenfalls explizit als Knoten repräsentiert. Die *data-of*-Kante verbindet sie mit den übrigen Knoten des generischen Szenenmodells. Sie stellt eine Spezialform der *con-of*-Kante dar.

In manchen Fällen ist es sinnvoll, nach Objektteilen oder -konkretisierungen erst dann zu suchen, wenn andere bereits erkannt wurden und einen bestimmten Kontext etabliert haben. Solche Subknoten werden über kontextabhängige Kanten (*cdpart-of* und *cdcon-of*) angebunden. Eine weitere Reihenfolgesteuerung läßt sich über die Angabe von Prioritäten in den *part-of*- und *con-of*-Kanten realisieren. Besitzt ein Objekt mehrere Bestandteile, sucht das System zuerst nach denen mit der höchsten Priorität.

Klassenhierarchien lassen sich mittels der *is-a*-Kante effizient repräsentieren, mit der ein spezielles Konzept mit einem allgemeineren verbunden wird. Dabei erbt das speziellere Objekt automatisch alle Bestandteile, Konkretisierungen, Methoden und Attribute von seinem Vaterknoten. Die ererbten Objekteigenschaften lassen sich bei Bedarf lokal neu definieren.

### **Topologische Kanten**

Eine Erweiterung gegenüber anderen auf semantischen Netzen basierenden Interpretationssystemen sind die topologischen Kanten. Mit ihnen werden Nachbarschaftsbeziehungen zwischen Objekten derselben Hierarchiestufe explizit beschrieben. Um nicht

---

alle denkbaren topologischen Relationen vordefinieren zu müssen, wird die allgemeine Klasse der *attributierten Relation*, abgekürzt *attr-rel*, eingeführt. Diese besitzt im Gegensatz zu allen anderen Kantentypen Attribute, die dazu genutzt werden können, die erwarteten Attributwerte der verbundenen Objekte weiter einzugrenzen. Eine attributierte Relation *nördlich-von* kann z. B. eingesetzt werden, um die geographische Position eines Objekts bezüglich eines bereits erkannten einzuschränken.

### 2.3.2 Inhalt und Strukturierung der Wissensbasis

Die vorgestellte Netzwerksyntax wird genutzt, um ein generisches Landschaftsmodell für die wissensbasierte Luftbildinterpretation zu definieren. Diese Wissensbasis enthält nur solche Objekte und Beziehungen, die für die automatische Auswertung der Bilder relevant sind. Es wird Vorwissen über die Struktur, Geometrie und Materialeigenschaften der erwarteten Szenenobjekte sowie deren Erscheinungsform in unterschiedlichen Sensoren benötigt. Bild 2.4 zeigt ein Beispiel für ein semantisches Netz zur Interpretation multisensorieller Luftbilder aus [151]. Die Wissensbasis gliedert sich in einen *3D-Szenenbereich* und einen *2D-Bildbereich*. Der 3D-Szenenbereich teilt sich weiter in eine *semantische* und eine *physikalische Ebene*. Sofern ein GIS genutzt werden soll, wird er um eine *GIS-Ebene* erweitert. Der 2D-Bildbereich besteht aus einer *Sensorebene* mit sensorspezifischen Konzepten und der *Bildverarbeitungsebene*, die die segmentierten Bildprimitive repräsentiert.

Die Knoten der einzelnen konzeptionellen Ebenen werden gemäß Definition ausschließlich durch *con-of*-Kanten verbunden. Die Datenknoten der Bildverarbeitungsebene sind durch die *data-of*-Kante angebunden. Innerhalb der Ebenen existieren nur *part-of*- und *is-a*-Kanten, die die Bestandteils- bzw. Klassenhierarchien beschreiben.

Die *semantische Ebene* enthält alle Begriffe mit symbolischer Bedeutung. Neben den eigentlichen Landschaftsobjekten ist in einem Teilbaum des Netzes das Wissen über verschiedene Sensoren und deren Abbildungseigenschaften im photometrischen und geometrischen Sinne repräsentiert. Eine Reihenmeßkamera wird z. B. beschrieben durch die Konzepte *Visuelle Sensitivität* bzw. *Zentralprojektion*. Das erste beinhaltet Informationen darüber, wie einzelne Materialien im visuellen Spektrum erscheinen, woraus sich der Farb- bzw. Grauwert im Bild ableiten läßt. Das zweite Konzept modelliert das geometrische Abbildungsverhalten der Reihenmeßkamera durch eine Lochkamera, die die 3D-Szene zentralperspektivisch auf die Bildebene projiziert. Sind die Parameter der Zentralprojektion [134][76] wie Position, Orientierung und Brennweite der Kamera bekannt, so lassen sich 3D-Raumpunkte in 2D-Bildpunkte umrechnen. Die inverse Abbildung von 2D-Bildkoordinaten in 3D-Raumkoordinaten

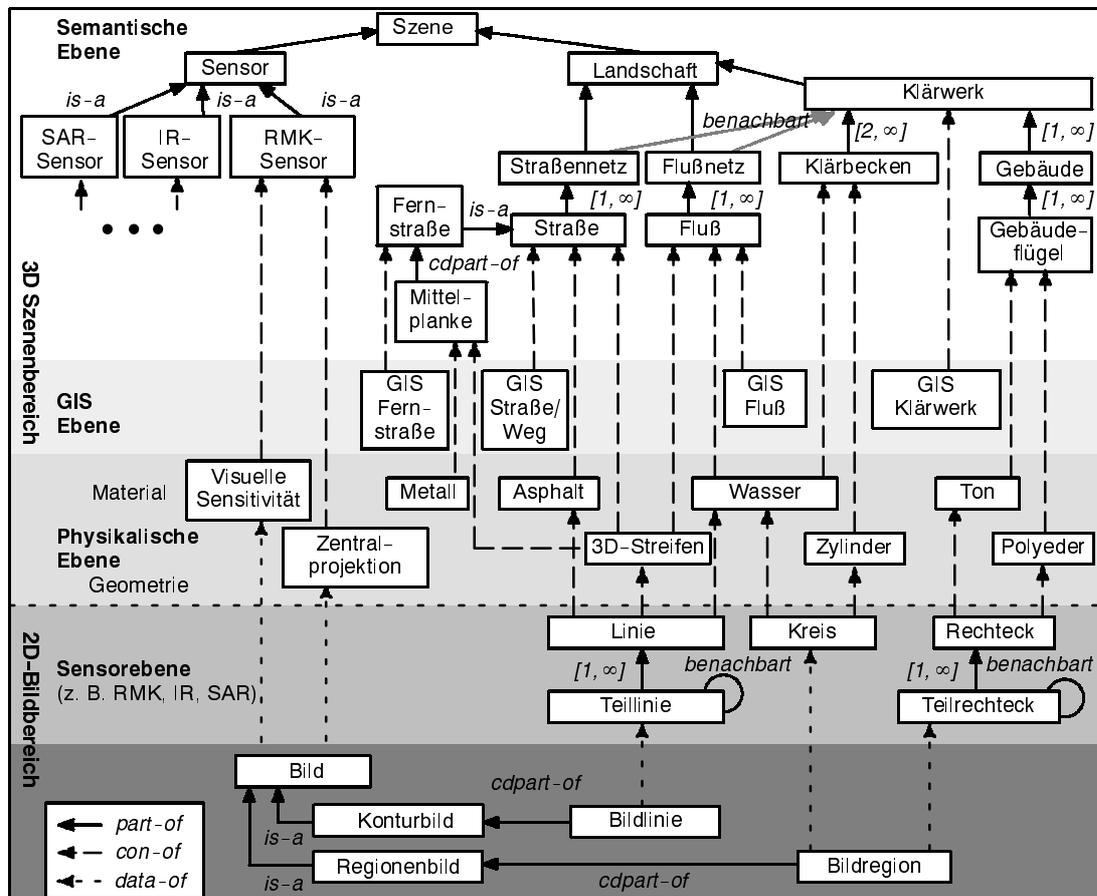


Bild 2.4: Beispiel für ein generisches Szenenmodell zur Detektion eines Klärwerks [151]

ist auf Basis eines einzelnen Bildes nicht eindeutig. Die Lösung liegt auf der Sichtlinie vom Brennpunkt der Kamera durch das Abbild des Raumpunktes in der Bildebene. Erst unter Hinzunahme eines zweiten Bildes ist der Raumpunkt eindeutig durch den Schnittpunkt der Sichtlinien korrespondierender Bildpunkte gegeben. Ein eindeutige Lösung läßt sich auch dann bestimmen, wenn das Höhenrelief der Landschaft bekannt ist oder als Ebene angenommen wird. Dann ergibt sich der Raumpunkt aus dem Schnitt der Sichtlinie mit dem Höhenrelief.

Das semantische Netz in Bild 2.4 beinhaltet die Landschaftsobjekte *Straßen-* und *Flußnetz*. Diese bestehen aus mindestens einer Straße bzw. einem Fluß, was durch die entsprechenden Intervalle an den *part-of*-Kanten angedeutet wird. Eine spezielle Form der Straße ist die *Fernstraße*, die eine *Mittelleitplanke* als zusätzlichen Bestandteil besitzt und deren Erwartungswerte für die Straßenbreite größer als beim Konzept *Straße* sind. Ferner ist das komplexe Objekt *Klärwerk* in der Wissensbasis modelliert. Es

---

besteht aus mindestens zwei *Klärbecken* und einem oder mehreren *Gebäuden*, die wiederum aus *Gebäudeflügeln* zusammengesetzt sind. Darüber hinaus läßt sich ein Klärwerk in Luftbildern daran erkennen, daß es an einem Fluß liegt, in den das geklärte Wasser abgeleitet wird, und eine Straße als Zufahrtsweg besitzt. Diese Nachbarschaftsbeziehungen werden im semantischen Netz explizit durch topologische Relationen namens *benachbart* modelliert.

Die Daten eines Geoinformationssystems stellen zusätzliches szenenspezifisches Wissen dar. Da die GIS-Objekte durch die Einordnung in den Objektartenkatalog des digitalen Landschaftsmodells bereits eine symbolische Bedeutung besitzen, repräsentiert das GIS eine Teilinterpretation der Szene, die allerdings fehlerhaft, unvollständig oder veraltet sein kann. Dennoch werden die GIS-Daten in weiten Teilen mit dem Bildinhalt übereinstimmen und sie sind deshalb eine zuverlässige Quelle zur Ableitung von Hypothesen über die Existenz und Lage der Szenenobjekte. Die GIS-Daten werden in einer *GIS-Ebene* direkt unterhalb der semantischen Ebene modelliert und stellen eine Konkretisierung der Landschaftsobjekte dar. Sind für das Beobachtungsgebiet keine GIS-Daten vorhanden, entfällt die GIS-Ebene ersatzlos.

Um eine automatische Zuordnung der symbolischen Begriffe aus der semantischen Ebene zu den segmentierten Bildprimitiven in der Bildverarbeitungsebene zu ermöglichen, werden die geometrische Form und das Material der Objekte sowie deren Erscheinungsform im Bild in den tieferliegenden Netzwerkschichten modelliert. Als Materialien werden hier *Metall*, *Asphalt*, *Wasser* und *Ton* (für die Dachziegel) definiert. Für die Beschreibung der 3D-Objektgeometrie existieren die Konzepte *3D-Streifen*, *Zylinder* und *Polyeder*. Diese Begriffe des 3D-Szenenbereichs werden in der *Sensorebene* weiter konkretisiert. Ein *3D-Streifen* wird im Bild durch eine *2D-Linie* repräsentiert, die allerdings aufgrund unvollständiger Segmentierung zerstückelt sein kann und deshalb aus mehreren benachbarten *Teillinien* zusammengesetzt wird. Ein Zylinder wird auf einen *Kreis* abgebildet, während ein *Polyeder* durch ein *Rechteck* bestehend aus *Teilrechtecken* im Bild beschrieben wird. Zusammen mit dem Wissen um Sensortyp und dessen Abbildungseigenschaften lassen sich die erwarteten geometrischen und photometrischen Eigenschaften eines Objekts in konkrete bildbezogene Größen transformieren. Längenangaben können von Metern in Bildpunkte, 3D-Koordinaten in Bildkoordinaten und Materialeigenschaften in Farb- oder Grauwerte umgerechnet werden, um so ein passendes Datum aus den segmentierten Bildregionen und -linien auszuwählen.

Die beschriebene Wissensbasis ist geeignet, in multisensoriellen Luftbildern Straßen- und Flußnetze zu detektieren. An Stellen, wo Straßen und Flüsse dicht

beieinander liegen oder wo GIS-Daten einen entsprechenden Hinweis geben, wird zusätzlich nach dem komplexen Objekt Klärwerk gesucht. Je nach vorhandenen Sensordaten werden die korrespondierenden Sensor-Ebenen zum semantischen Netz hinzugefügt. Liegen z. B. keine IR-Daten vor, wird das zugehörige Wissen auch nicht für die Interpretation herangezogen.

Das semantische Netz mit dem generischen Szenenmodell ist vor Beginn der Bildinterpretation als Vorwissen zu definieren. Neben der Strukturierung in unterschiedliche Hierarchien und Netzwerkebenen müssen die relevanten Objekteigenschaften in Attributen und Attributberechnungsmethoden formuliert werden. Diese Wissensbasis wird vom vorgestellten System genutzt, um für die vorliegenden Sensordaten eine automatische Bildinterpretation durchzuführen. Aus dem Konzeptnetz wird ein instanziiertes semantisches Netz generiert, das als symbolische Szenenbeschreibung dient. Die Vorgehensweise während des Analyseprozesses wird im folgenden Abschnitt beschrieben.

## **2.4 Steuerung der Bildinterpretation**

Für eine automatische Bildinterpretation reicht es nicht aus, das Vorwissen über die Szene strukturiert und einer Syntax folgend zu formulieren. Darüber hinaus ist ein Mechanismus erforderlich, der diese Wissensbasis auswertet und automatisch Schlußfolgerungen für die zielgerichtete Steuerung des Bildanalyseprozesses ableitet. Da die Wissensbasis jedoch austauschbar sein soll, sollte diese Steuerungskomponente anwendungsunabhängig ausgelegt werden, so daß sie für jede beliebige, aber korrekt formulierte Wissensbasis einsetzbar ist. Das beschriebene Bildinterpretationssystem AIDA besitzt eine regelbasierte Inferenzmaschine, die die Reihenfolge der Analyse steuert. Sie umfaßt einen Regelsatz, der ausschließlich die Semantik der in Kapitel 2.3 vorgestellten Netzwerksprache nutzt und somit unabhängig von den konkreten Wissensinhalten ist.

Durch die Ausführung einer Inferenzregel wird die aktuelle symbolische Szenenbeschreibung modifiziert. Es werden z. B. Hypothesen generiert oder neue Instanzen gebildet. Jede Szenenbeschreibung - charakterisiert durch die Menge der aktuell gültigen Konzepte und Instanzen - wird in einem Knoten eines Suchbaums dokumentiert. Führt die Aktivierung einer Inferenzregeln zu mehreren alternativen Szenenbeschreibungen, werden sie als konkurrierende Lösungen betrachtet und jede für sich in einem neuen Suchbaumknoten gespeichert. Der Suchbaum spaltet sich auf. Zur Steigerung der Effizienz dieses Suchprozesses werden die konkurrierenden Szenenbeschreibung im Hinblick auf ihre Vollständigkeit und Übereinstimmung mit den Erwartungen bewertet. Die vielversprechendere Alternative wird stets zuerst weiterver-

folgt. Die restlichen Lösungen werden nicht verworfen, da sie eventuell zu einem späteren Zeitpunkt noch untersucht werden müssen.

Im folgenden wird die Inferenzmaschine mit einigen Regeln und Strategien beschrieben. Ferner wird das Problem des Suchprozesses und der Bewertung konkurrierender Szenenbeschreibungen beleuchtet.

### 2.4.1 Inferenzmaschine

Die Inferenzmaschine im Bildinterpretationssystem AIDA ist als regelbasiertes System ausgelegt. Die Interpretationsstrategie ist durch einen Satz von Inferenzregeln definiert. Durch Austausch der Regelbasis kann die verfolgte Strategie verändert werden und ermöglicht so ein äußerst flexibles Systemverhalten. Es existieren Strategien, die eine daten- bzw. eine modellgetriebene Interpretation favorisieren und solche, die zusätzlich die topologischen Relationen im semantischen Netz berücksichtigen. Neue Strategien können leicht durch das Entfernen bzw. Hinzufügen von Regeln realisiert werden. Die Regeln einer Strategie besitzen unterschiedliche Prioritäten, die die Reihenfolge der Regelarbeitung bestimmen. Somit kann auch durch Änderung der Prioritäten die Interpretationsstrategie variiert werden.

Der Inferenzmechanismus testet in der Reihenfolge fallender Prioritäten die Bedingungsteile der vorhandenen Regeln. Die erste anwendbare Regel wird gefeuert. Die jeweiligen Bedingungsteile prüfen, ob für die Knoten der aktuellen Szenenbeschreibung bestimmte Interpretationszustände vorliegen. Die Aktionsteile hingegen erzeugen neue Knoten oder modifizieren die Interpretationszustände der bereits bestehenden Knoten. Durch das wiederholte Ausführen von Inferenzregeln wird nach und nach die symbolische Szenenbeschreibung aufgebaut und damit die Bildinterpretation vorangetrieben.

Eine charakteristische Vorgehensweise bei der Bildinterpretation ist die *Top Down/Bottom Up*-Strategie. Dabei werden zunächst modellgetrieben (*top-down*) Hypothesen über die Existenz und die Eigenschaften von Szenenobjekten aus der Wissensbasis abgeleitet. Anschließend wird in den Daten geprüft, ob und wie gut die Hypothesen mit dem abgebildeten Szeneninhalte übereinstimmen. Dazu werden segmentierte Bildprimitive über die *data-of*-Kante angebunden. Die Hypothese kann datengetrieben (*bottom-up*) verifiziert werden, indem die erwarteten Attributwerte mit den konkret in den Daten gemessenen verglichen werden. Eine derartige Interpretationsstrategie kann durch wenige Regeln zur Instanzierung und Hypothesengenerierung modelliert werden. Weitere Regeln dienen zur Spezialisierung bereits instanzierter Knoten und zur Erzeugung von mehrfach angebundenen Knoten.

**Die Instanzierungsregeln** erzeugen partielle, komplette oder fehlende Instanzen. Existieren z. B. für eine Hypothese  $I_H(n)$  im Suchbaumknoten  $\mathcal{N}$  alle in der Wissensbasis definierten obligatorischen, kontextunabhängigen Bestandteile, Konkretisierungen und die über *data-of*- sowie *attributierten* Relationen angebundene Knoten als komplette Instanz, erzeugt die Regel  $R_P$  eine partielle Instanz  $I_P(n)$ :

**R<sub>P</sub>: Partielle Instanzierung**

WENN für eine Hypothese  $I_H(n)$  alle kompletten Instanzen  $I_K(m)$  existieren mit:

$$m \xrightarrow{r_{obl}} n \quad r \in \{part-of, con-of, data-of, attr-rel\} ; n, m \in \mathcal{N}$$

DANN erzeuge den Interpretationszustand  $I_P(n)$  wie folgt:

falls  $I_H(n)$  bereits im Suchbaumknoten  $\mathcal{N}$  modifiziert wurde,

- ändere Zustand zu  $I_P(n)$
- berechne die Attribute

sonst

- erzeuge und verzeigere eine leere Instanz  $I_P(n)$
- kopiere und berechne die Attribute

markiere  $I_P(n)$  als modifizierten Knoten.

Zusätzlich existieren die folgenden Instanzierungsregeln:

- Die Regel  $R_K$  erzeugt aus einer partiellen Instanz  $I_P(n)$  eine komplette Instanz  $I_K(n)$ , wenn alle obligatorischen Komponenten (über *part-of*, *con-of*, *data-of* oder *attr-rel* angebundene Kindknoten) vollständig instanziiert sind.
- Sollen für einen Knoten  $I_H(n)$  Daten über eine *data-of*-Kante angebundene werden, aktiviert die Regel  $R_{data-of}$  eine lokal definierte Segmentierungsmethode und verbindet die entstandenen Bildprimitive mit  $I_H(n)$ .
- Wenn eine obligatorische Komponente nicht instanziiert werden kann, erzeugt die Regel  $R_F$  eine fehlende Instanz  $I_F(n)$ .

**Die Regeln zur Generierung von Hypothesen** initiieren die Suche nach fehlenden Komponenten oder schränken die Eigenschaften bestehender Hypothesen weiter ein:

- Die Regel  $R_{H:inv(cd)(opt)}$  erzeugt modellgetrieben eine neue Hypothese  $I_H(m)$ , wenn eine obligatorische (optionale) kontext(un)abhängige Komponente  $I(m)$  zur Instanzierung des übergeordneten Knotens  $I_H(n)$  fehlt.
- Wenn ein obligatorischer kontextunabhängiger übergeordneter Knoten  $I(m)$  fehlt oder noch nicht durch den aktuellen (untergeordneten) Knoten  $I(n)$  eingeschränkt wurde, erzeugt die Regel  $R_H$  eine datengetriebene Hypothese  $I_H(m)$ .

- Die Regel  $R_{AR:(inv)(cd)(opt)}$  erzeugt ausgehend von einer kompletten Instanz  $I_K(n)$  eine Hypothese  $I_H(m)$  über obligatorische (optionale) kontext(un)abhängige attributierte (also insbesondere topologische) Relationen (in inverser Richtung).

**Die Spezialisierungsregel** wertet die *is-a*-Kante aus, um speziellere Hypothesen für bereits instanziierte Konzepte zu generieren:

- Wenn eine komplette Instanz  $I_K(n)$  existiert, die noch nicht spezialisiert wurde, ersetzt die Regel  $R_{Spez}$  die Generalisierung  $I_K(n)$  durch jeweils eine Hypothese  $I_H(n')$  für eine der  $i$  möglichen Spezialisierungen von  $n$ . Zusammen mit der Generalisierung  $I_K(n)$  ergeben sich  $i+1$  konkurrierende Szenenbeschreibungen, die getrennt voneinander weiterverfolgt werden.

**Die Regel zur Mehrfachbindung** ermöglicht, daß ein Knoten eine Komponente von mehreren übergeordneten Knoten wird und somit mehrfach gebunden wird:

- Fehlt für eine Instanz  $I_K(n)$  eine Bindung zu einem übergeordneten Knoten, existiert desweiteren eine Instanz  $I_K(n')$  desselben Konzepts mit dieser Bindung und ist  $I_K(n')$  über einen benutzerdefinierten Pfad von  $I_K(n)$  aus erreichbar, dann verschmilzt die Regel  $R_{Binde}$  die beiden Knoten zu  $I_K(n'')$  mit Mehrfachbindung.

Während der Interpretation kann der Bedingungsteil einer Inferenzregel für mehrere Knoten des aktuellen Suchbaumknotens erfüllt sein. In diesem Fall entscheidet die Lage des Knotens im semantischen Netz darüber, für welchen Knoten die Regel gefeuert wird. Zur Bestimmung der Rangfolge wird die maximale und minimale Pfadlänge zum Signal über *is-a*-, *con-of*- und *part-of*-Kanten herangezogen [125]. Der Knoten mit dem höchsten Rang wird zur weiteren Analyse ausgewählt und für ihn die Regel gefeuert.

Mit Hilfe der vorgestellten Regeln ist eine vollständige Szenenanalyse realisierbar, wie am Beispiel in Bild 2.5 deutlich wird: Ausgehend von der einfachen Wissensbasis in Bild 2.5a soll eine *Fernstraße* instanziiert werden. Die Analyse beginnt unter Vorgabe einer initialen Hypothese  $I_H(\text{Straße})$  mit der modellgetriebenen Generierung einer Hypothese für die Konkretisierung *Fahrbahn* (Aktivierung der Regel  $R_{H:inv}(\text{Straße-1})$  in Schritt 1, s. Bild 2.5). Eine *Fahrbahn* kann in den Daten durch eine *Streifen* bestätigt (Schritt 2) und daraufhin partiell (Schritt 3) und schließlich komplett (Schritt 4) instanziiert werden. Das hinzugewonnene Wissen wird genutzt, um die Hypothese  $I_H(\text{Straße})$  datengetrieben weiter einzuschränken (Schritt 5). Anschließend erfolgt die modellgetriebenen Suche nach der zweiten, noch fehlenden *Fahrbahn* (Schritt 6), die analog zu der ersten instanziiert wird (Schritte 7-9). Alle obligatorischen Bestandteile der *Straße* wurden gefunden, so daß sie partiell und komplett instanziiert werden kann (Schritte 10 und 11). Das Konzept *Straße* kann laut Wissensbasis zu einer *Fernstraße*

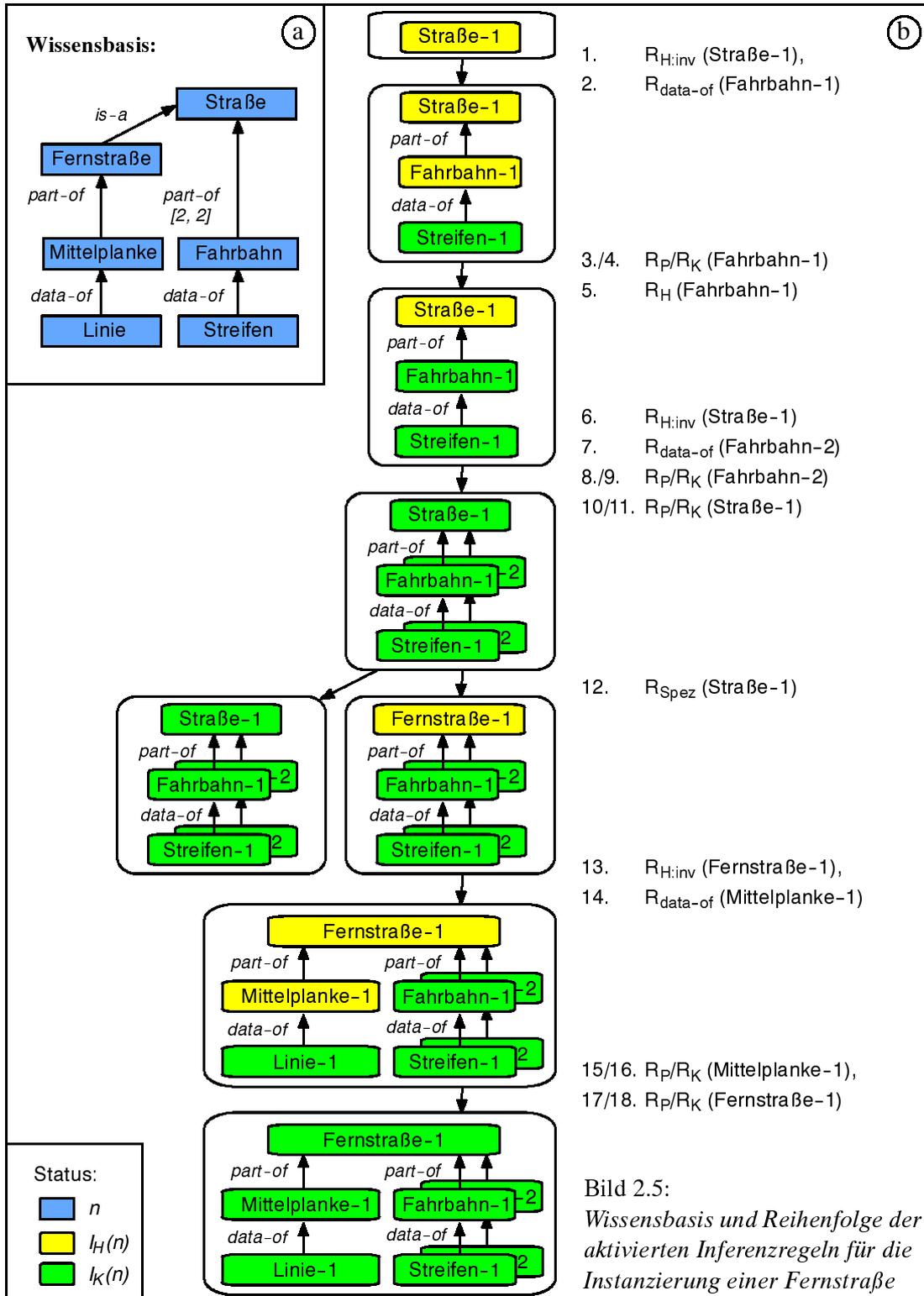


Bild 2.5: Wissensbasis und Reihenfolge der aktivierten Inferenzregeln für die Instanzierung einer Fernstraße

spezialisiert werden. In Schritt 12 wird zum einen die bisherige Interpretation beibehalten und zum anderen  $I_K(\text{Straße})$  durch  $I_H(\text{Fernstraße})$  ersetzt. Der Suchbaum spaltet sich in zwei Äste auf. Die erste Alternative ist vollständig und wird nicht weiter untersucht, da keine Regel mehr feuern kann. Zur Bestätigung der spezielleren Szenenbeschreibung wird die Suche nach dem fehlenden Bestandteil *Mittelplanke* initiiert (Schritt 13). Das System sucht in den Daten gezielt nach einer *Linie*, die zwischen den bereits detektierten Fahrbahnen liegt. Nach der Bindung eines geeigneten Datums (Schritt 14) kann sowohl die *Mittelplanke* als auch das Konzept *Fernstraße* komplett instanziiert werden (Schritte 15–18). Das Analyseziel ist erreicht.

## 2.4.2 Bildinterpretation als Suchprozeß

Das Ziel der Bildinterpretation ist eine vollständige und konsistente symbolische Beschreibung der Szene in Form eines instanziierten semantischen Netzes. Formal wird dazu ein Konzept der Wissensbasis als sog. *Zielkonzept* definiert. Wurde dieses Konzept vollständig instanziiert und somit in den Daten verifiziert, gilt das Ziel als erreicht und die Analyse endet. Aufgrund der unterschiedlichen Eingangsdaten kann a priori nicht festgelegt werden, wie man zum Ziel gelangt. Stattdessen existieren problemunabhängige Inferenzregeln, deren wiederholte Anwendung sukzessive eine Szenenbeschreibung aufbauen. Allerdings treten häufig Situationen auf, die mehrere Möglichkeiten zur Fortsetzung der Analyse erlauben. Welcher der konkurrierenden Wege letztlich der optimale ist und zum Ziel führt, kann oftmals erst im weiteren Verlauf der Auswertung festgestellt werden.

Um die Möglichkeit zu haben, in eine alternative, zunächst zurückgestellte Szenenbeschreibung zurückzukehren und dort die Analyse fortzusetzen, werden die Zwischenergebnisse in einem Suchbaum dokumentiert. Jede konkurrierende Szenenbeschreibung wird mit ihren aktuellen Konzepten und Instanzen in einem Suchbaumknoten abgespeichert. Erzeugt eine Inferenzregel mehrere Ergebnisse, so werden diese als konkurrierende Interpretationen aufgefaßt. Der Suchbaum spaltet sich in entsprechend viele Folgeknoten auf. Im Beispiel aus Bild 2.5 hatte die Spezialisierungsregel  $R_{Spez}$  in Schritt 12 zwei mögliche Interpretationen geliefert und zur Aufspaltung des Suchbaums geführt. Zur Effizienzsteigerung werden die semantischen Netze bei einer Aufspaltung des Suchbaums nicht vollständig kopiert. Stattdessen repräsentieren die Suchbaumknoten jeweils unterschiedliche Teilmengen eines Gesamtnetzes. In Bild 2.6 ist diese Teilmengenbildung für die konkurrierenden Suchbaumknoten aus Bild 2.5 illustriert.

Neben der Spezialisierungsregel kann auch die Regel zur Datenbindung  $R_{data-of}$  konkurrierende Ergebnisse liefern, wenn mehrere zu den Erwartungen passende Bildprimitive existieren. Ferner kann sich der Suchbaum aufspalten, wenn durch die

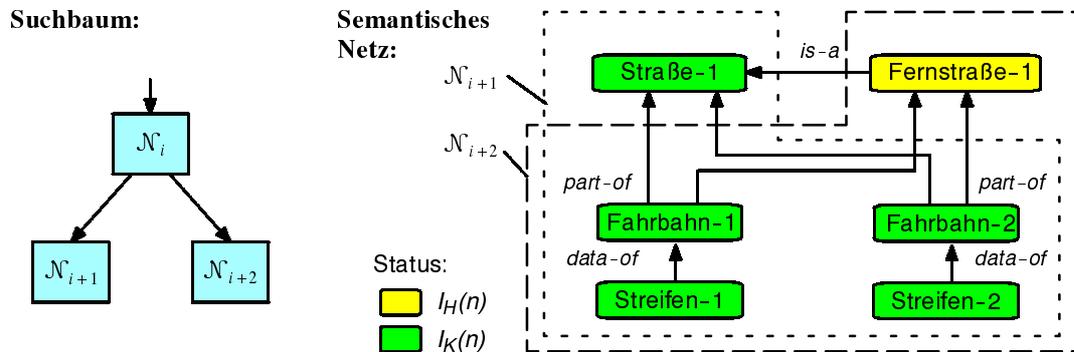


Bild 2.6: Repräsentation von Suchbaumknoten durch Teilmengenbildung

Regeln  $R_{H:(inv)(cd)(opt)}$  sich gegenseitig ausschließende Hypothesen erzeugt oder durch die Regel zur Mehrfachbindung  $R_{Binde}$  mehrere komplementäre Knoten gefunden werden.

Die Steuerungskomponente des Interpretationssystems hat die Aufgabe, einen Pfad durch den Suchbaum zum Ziel zu finden. Um nicht alle kombinatorisch möglichen Pfade untersuchen zu müssen, wird der A\*-Algorithmus [102] als Graphsuchverfahren eingesetzt. Dieser Algorithmus findet stets den optimalen Pfad zum Ziel, vorausgesetzt er existiert. Als Kostenfunktion werden die angefallenen Kosten zum aktuellen Suchbaumknoten  $g(\mathcal{N})$  und die geschätzten Restkosten zum Ziel  $h^*(\mathcal{N})$  herangezogen. Bedingung für das Verfahren ist neben der Monotonie der Kostenfunktion, daß die Restkosten  $h^*(\mathcal{N})$  optimistisch gegenüber den tatsächlichen Kosten  $h(\mathcal{N})$  abgeschätzt werden:

$$f(\mathcal{N}) = g(\mathcal{N}) + h^*(\mathcal{N}) \quad \text{mit : } 0 \leq h^*(\mathcal{N}) \leq h(\mathcal{N}) \quad (2.1)$$

Für alle alternativen Interpretationszustände - repräsentiert durch einen Suchbaumknoten  $\mathcal{N}_i$  - werden die Kosten  $f(\mathcal{N}_i)$  bestimmt. Die Alternative mit den geringsten Kosten ist in der aktuellen Situation die vielversprechendste und wird deshalb für die weitere Analyse ausgewählt. Stellt sich im Verlauf der Auswertung heraus, daß die Interpretation nicht korrekt ist, werden die Kosten über diejenigen einer zunächst zurückgestellten Alternative steigen. Das System wählt dann diese Alternative aus und setzt dort die Analyse fort. Auf diese Weise wird stets der optimal erscheinende Ast des Suchbaum weiterentwickelt, bis schließlich das Ziel erreicht ist.

Der A\*-Algorithmus tendiert in der Praxis oftmals zu einer Breitensuche, wenn sich die konkurrierenden Suchbaumknoten nur gering in ihren Kosten unterscheiden. Anstelle konsequent einen Pfad in die Tiefe zu verfolgen, springt das Verfahren zwischen den ähnlich guten Teilbäumen hin und her und entwickelt sie parallel. Zur

Beschleunigung des Suchprozesses wird deshalb in [152][110] ein modifiziertes Verfahren vorgeschlagen. Der sog.  $\varepsilon$ -A\*-Algorithmus bevorzugt Suchbaumknoten, die näher am Ziel liegen und favorisiert dadurch eine Tiefensuche. Dazu werden die Restkosten der anderen, weniger nah am Ziel liegenden Suchbaumknoten um den Faktor  $\varepsilon$  verschlechtert:

$$f_{\varepsilon}(\mathcal{N}) = g(\mathcal{N}) + (1 + \varepsilon) h^*(\mathcal{N}) \quad (2.2)$$

Das Verfahren verfehlt unter Umständen den optimalen Pfad zum Ziel und liefert gegebenenfalls eine suboptimale Lösung. Die Kosten dieser Lösung weichen maximal um den Faktor  $1+\varepsilon$  von denen der optimalen Lösung ab [108].

Entscheidend für die Konvergenz der Verfahren ist die Ermittlung der Kostenterme  $g(\mathcal{N})$  und  $h^*(\mathcal{N})$ , die während der Suche minimiert werden sollen. Im Falle des Interpretationssystem lassen sich schwer die Kosten eines Suchbaumknotens bestimmen. Es fällt leichter die Qualität oder Güte der Interpretation zu bewerten, indem die Meßergebnisse mit den erwarteten Werten verglichen werden. Anstelle der Minimierung von Kosten muß dementsprechend die Güte der Interpretation maximiert werden, um die optimale Lösung zu finden. Der folgende Abschnitt behandelt das bislang verwendete Bewertungssystem im Bildinterpretationssystem AIDA.

### 2.4.3 Possibilistisches Bewertungssystem

Das Bewertungssystem für die Szenenanalyse soll ein Gütemaß für konkurrierende Interpretationszustände liefern, so daß eine Bearbeitungsrangfolge für die einzelnen Alternativen festgelegt werden kann. Dabei ist darauf zu achten, daß die Bewertungsfunktion monoton ist, damit die Interpretationszustände trotz unterschiedlicher Analysetiefen vergleichbar bleiben. Die Steuerung wählt den bestbewerteten Interpretationszustand für die weitere Analyse aus.

Zur Definition einer Rangfolge wäre ein relatives Qualitätsmaß ausreichend, das lediglich eine Aussage darüber trifft, ob die Interpretation repräsentiert durch den Suchbaumknoten  $\mathcal{N}_i$  besser oder schlechter ist als diejenige des Suchbaumknotens  $\mathcal{N}_j$ . Es ist jedoch wünschenswert, daß die Bewertung auch ein absolutes Vertrauensmaß liefert, das angibt, wie gut die Interpretation mit den Erwartungen übereinstimmt. In der Regel wird das Vertrauensmaß über den normierten Wertebereich  $[0, 1]$  definiert. Der Minimalwert steht für fehlendes und der Maximalwert für vollstes Vertrauen.

Die Bewertung muß sowohl unsichere als auch ungenaue Informationen verarbeiten können, da die Daten oft fehlerbehaftet und unvollständig und Klassifikationsergebnisse selten konfliktfrei sind. Ein Aussage ist *unsicher*, wenn sie nicht eindeutig als wahr

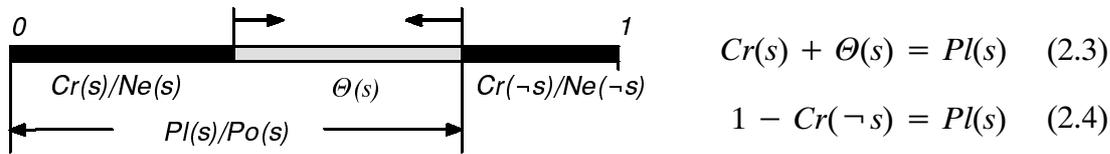


Bild 2.7: Zusammenhang zwischen Unwissenheit  $\Theta(s)$  sowie Glaubwürdigkeit  $Cr(s)$  und Plausibilität  $Pl(s)$  bzw. Notwendigkeit  $Ne(s)$  und Möglichkeit  $Po(s)$  gemäß der Possibility-Theorie.

oder falsch klassifiziert werden kann. Dagegen nennt man eine Aussage *ungenau*, wenn sie keinen exakten Wert besitzt, sondern nur mit einer gewissen Toleranz angegeben werden kann. In der Literatur wurden zahlreiche Ansätze vorgestellt, um Unsicherheiten repräsentieren und verarbeiten zu können [79]. Hier sind zum Beispiel die Evidenztheorie [136], Bayes-Netze [100][107] und die Possibility-Theorie [34] zu nennen. Zur Modellierung von Ungenauigkeiten wurden u. a. Fuzzy-Sets [166] und linguistische Variablen [167] vorgeschlagen.

Tönjes [152] entwickelt in seiner Arbeit ein Bewertungssystem für die Bildinterpretation, das die Faktoren Unsicherheit und Ungenauigkeit integriert. Die Verarbeitung der Unsicherheit ist aus der Possibility-Theorie abgeleitet, während ungenaue Aussagen durch Fuzzy-Sets modelliert werden:

Zur Beurteilung der Sicherheit einer Aussage wird ein Vertrauensmaß definiert, das die Menge aller Aussagen auf das Einheitsintervall  $[0, 1]$  abbildet. Eine Hypothese, die noch nicht in den Daten verifiziert wurde, muß notwendigerweise weder richtig noch falsch sein. Es handelt sich um Unwissenheit. Klassische Vertrauensmaße sind nicht in der Lage, Unwissenheit zu repräsentieren. Die Dempster-Shafer-Theorie [135] hingegen teilt zur Modellierung der Unwissenheit das Einheitsintervall in drei disjunkte Bereiche: die Glaubwürdigkeit der Aussage  $Cr(s)$  (engl.: Credibility), die Unwissenheit  $\Theta(s)$  und die Unglaubwürdigkeit  $Cr(\neg s)$  (die Glaubwürdigkeit der entgegengesetzten Aussage) (Bild 2.7).

Zu Beginn der Analyse liegt weder Glaubwürdigkeit noch Unglaubwürdigkeit vor, die Unwissenheit nimmt das gesamte Intervall  $[0, 1]$  ein. Der fortschreitende Analyseprozeß verringert durch jede hinzukommende Evidenz, d.h. durch jeden Meßwert, die Unwissenheit und erhöht entweder die Glaubwürdigkeit oder Unglaubwürdigkeit der Interpretation. Die Summe aus Glaubwürdigkeit und Unwissenheit ist der Wert, auf den die Glaubwürdigkeit bis zum Ende der Analyse maximal steigen kann, und wird Plausibilität  $Pl(s)$  genannt. Sie stellt eine optimistische Abschätzung der Glaubwürdigkeit dar. Für den Sonderfall konsonanter fokaler Aussagen [34] verein-

facht sich die Glaubwürdigkeit zur Notwendigkeit  $Ne(s)$  (engl.: Necessity) und die Plausibilität zur Möglichkeit  $Po(s)$  (engl.: Possibility).

Eine sichere Aussage kann ungenau sein, z. B. durch Toleranzen bei der Messung. Die Modellierung von ungenauen Aussagen erfolgt im vorgestellten Ansatz mit Hilfe von Fuzzy-Sets im Sinne von Zadeh [166]. Dabei wird die Zugehörigkeit einer Meßgröße zu einem Intervall nicht binär entschieden, sondern durch einen fließenden Übergangsbereich zwischen Zugehörigkeit und Nichtzugehörigkeit relativiert. Die Zugehörigkeitsfunktion  $\mu(x)$  wird durch eine unscharfe Menge in Form eines Trapezes modelliert (Bild 2.8). Analog zur Verknüpfung scharfer Mengen lassen sich auch für Fuzzy-Sets die Verknüpfungen *Vereinigung*, *Durchschnitt* und *Komplement* definieren, auf deren Basis die Attributbewertung im semantischen Netz realisiert ist:

Sowohl Attributwert  $E$  (für Evidenz) als auch der erwartete Attributwert  $H$  (für Hypothese) werden als unscharfe Mengen interpretiert. Aufgabe der Bewertung ist es, die Übereinstimmung von Meßbereich und Erwartungsbereich zu beurteilen. Dazu werden für eine Meßgröße  $x$  (z. B. die Breite einer Straße) die Möglichkeit  $Po(H|E)$  und die Notwendigkeit  $Ne(H|E)$ , daß die Aussage “ $x$  ist  $H$ ” (z. B. “Die Straße ist schmal”) wahr ist unter der Bedingung “ $x$  ist  $E$ ” (z. B. “Die gemessene Straßenbreite ist 7m”) gemäß (2.5) und (2.6) bestimmt. Die Größen  $po_H(x)$  und  $po_E(x)$  stellen die Möglichkeitsverteilungen der Hypothese bzw. der Evidenz dar.

$$Po(H|E) = \sup_{x \in X} \min(po_H(x), po_E(x)) \quad (2.5)$$

$$Ne(H|E) = \inf_{x \in X} \max(po_H(x), 1 - po_E(x)) \quad (2.6)$$

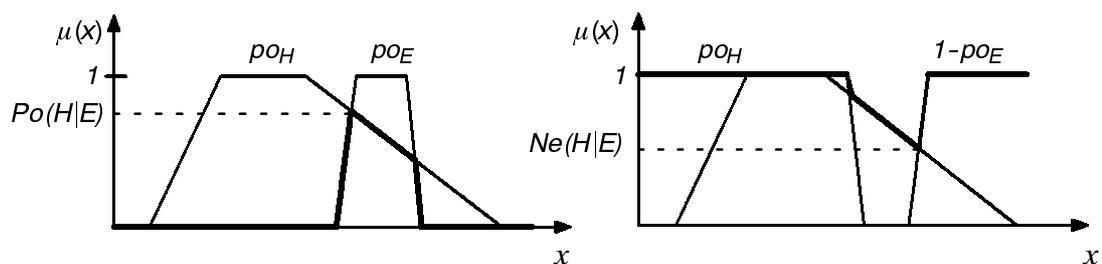


Bild 2.8: Bestimmung der Möglichkeit  $Po$  und Notwendigkeit  $Ne$  einer Hypothese  $H$  bei gegebener Evidenz  $E$ . Die Hypothese entspricht dem erwarteten Attributwert, die Evidenz dem gemessenen Attributwert, die jeweils durch trapezförmige Fuzzy-Sets modelliert werden.

Bei der Bewertung einer Szeneninterpretation müssen nicht nur einzelne Attribute bewertet werden. Vielmehr müssen alle Attributbewertungen einer Instanz zu einer

Gesamtbewertung zusammengefaßt werden. Schließlich sind alle Instanzbewertungen zu einer Bewertung für die ganze Szeneninterpretation, also den Suchbaumknoten, zu aggregieren. Bei der Integration verschiedener Einzelbewertungen unterscheidet Tönjes in die Kombination komplementärer und die Kombination redundanter Aussagen:

Komplementäre Aussagen sind voneinander unabhängig und ergänzen sich gegenseitig, wie z. B. mehrere orthogonale Attribute einer Instanz oder die einzelnen Bestandteile eines komplexen Objekts. Zur Kombination der Einzelbewertungen werden die Notwendigkeiten und Möglichkeiten der Teilaussagen durch eine gewichtete geometrische Mittelung [152] zu einem Gesamtwert für Notwendigkeit bzw. Möglichkeit der Gesamtaussage zusammengefaßt. Dabei wird jeder Knoten und jedes Attribut des semantischen Netzes entsprechend seines Knoten- bzw. Attributgewichts berücksichtigt. Das Bildinterpretationssystem nutzt diese Rechenvorschrift, um die Bewertungen mehrerer Attribute sowie obligatorischer bzw. optionaler Objektteile zu kombinieren.

Redundante Aussagen sind nicht orthogonal zueinander und beschreiben (teilweise) dasselbe Ereignis. Ein Beispiel ist die Konkretisierung eines Objekts in mehreren (multisensoriellen) Datensätzen. Die Sicherheit für die Existenz eines Objekts kann erhöht werden, wenn man es nicht nur in einem Bild detektiert, sondern es auch in anderen erkennen kann. Andererseits sollte die Bewertung des Objekts sich nicht verschlechtern, wenn man es nicht in einem zweiten Bild detektieren kann. Vielleicht ist der zweite Sensor nicht geeignet, das Objekt abzubilden, oder das Objekt wird aufgrund der Aufnahmeperspektive durch andere verdeckt. Aus diesem Grund werden die Bewertungen redundanter Aussagen mit Hilfe der Regel von Dempster [79] kombiniert. Sie summiert alle Anteile der Einzelbewertungen, die die Gesamtaussage unterstützen, und normiert die Summe auf alle Anteile, die der Gesamtaussage nicht widersprechen [152]. Die Berechnungsvorschrift wird vom Bildinterpretationssystem zur Bewertung von Objekten angewendet, die mehrere Konkretisierungen besitzen.

Wie bereits angedeutet stellt die Möglichkeit eine optimistische Abschätzung der Notwendigkeit dar. Die Notwendigkeit  $Ne(\mathcal{N})$  einer durch den Suchbaumknoten  $\mathcal{N}$  repräsentierten Interpretation kann maximal auf den Wert der Möglichkeit  $Po(\mathcal{N})$  steigen. Die Restgüte ist durch die Unwissenheit  $\Theta(\mathcal{N}) = Po(\mathcal{N}) - Ne(\mathcal{N})$  begrenzt. Die Gesamtgüte der Interpretation wird entsprechend der Kostenfunktion des A\*-Algorithmus aus Gl. (2.1) ermittelt:

$$f(\mathcal{N}) = Ne(\mathcal{N}) + \Theta(\mathcal{N}) = Po(\mathcal{N}) \quad (2.7)$$

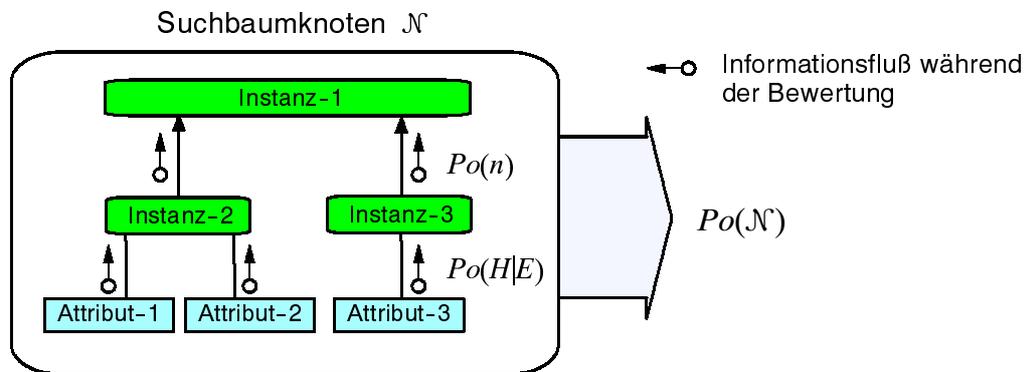


Bild 2.9: Datengetriebene Aggregation der Einzelbewertungen von Instanzen und Attributen zu einer Gesamtbewertung der Szenenbeschreibung

Somit dient die Möglichkeit  $Po(\mathcal{N})$  als Entscheidungskriterium für die Auswahl der besten Alternative. Soll der  $\varepsilon$ -A\*-Algorithmus genutzt werden, so sind die Bewertungen der restlichen, konkurrierenden Suchbaumknoten  $\mathcal{N}'$  nach Gl. (2.8) zu mindern.

$$f(\mathcal{N}') = Ne(\mathcal{N}') + (1 - \varepsilon) \Theta(\mathcal{N}') = \varepsilon Ne(\mathcal{N}') + (1 - \varepsilon) Po(\mathcal{N}') \quad (2.8)$$

Die Besonderheit des vorgestellten Bewertungssystems liegt in der Möglichkeit, die Unsicherheit der Interpretation zu repräsentieren. Diese kann elegant als Restgüte für den A\*-Algorithmus genutzt werden. Ferner wird bei der Kombination von mehreren Aussagen unterschieden, ob die Aussagen komplementär oder redundant sind. Im semantischen Netz wird diese Unterteilung direkt auf die Aggregation von Attributbewertungen bzw. die Kantentypen *part-of* bzw. *con-of* abgebildet. Das Verfahren faßt die in den unterschiedlichen Attributen durch Meßwerte eingebrachte Evidenz zu einer Gesamtbewertung des Suchbaumknotens zusammen. Dabei werden die Bewertungen von unten nach oben, also datengetrieben kombiniert (Bild 2.9).

Unterscheiden sich jedoch mehrere konkurrierende Interpretationen lediglich durch eine neue Hypothese, z. B. durch unterschiedliche Spezialisierungen eines Begriffs, so kommt keine neue Evidenz aus den Daten hinzu, da es sich um eine rein modellgetriebene Modifikation der Szenenbeschreibung handelt. Alle Alternativen werden somit gleich gut bewertet. Die Analyse wählt dann zufällig einen der Suchbaumknoten für die weitere Analyse aus, obwohl eventuell Vorwissen über die Wahrscheinlichkeiten der unterschiedlichen Hypothesen existiert. In der vorliegenden Arbeit wird ein neuartiges Bewertungssystem vorgestellt, das a priori bekannte Auftretenswahrscheinlichkeiten von Szenenobjekten mit in den Bewertungsprozeß einbezieht und durch die Selektion der jeweils wahrscheinlicheren Alternative ein deterministisches Systemverhalten bewirkt (s. Kapitel 4).

### 3. Repräsentation und Nutzung von temporalen Wissensinhalten

Das in Kapitel 2 vorgestellte Bildinterpretationssystem ist in der Lage, automatisch eine symbolische Beschreibung der in den Eingangsdaten dargestellten Szene zu erzeugen. Aufgrund der Trennung von allgemeingültigem und sensorspezifischem Wissen, ist das System nicht auf einen bestimmten Bildtyp festgelegt. Vielmehr können während der Szeneninterpretation die Bilder mehrerer Sensoren gleichzeitig ausgewertet werden. Die Aufnahmezeitpunkte der einzelnen Luftbilder bleiben jedoch unberücksichtigt. Das System geht stets davon aus, daß in allen Bildern eine identische Szene abgebildet wurde, auch wenn die Aufnahmen unterschiedlichen Datums sind. Veränderungen, die zwischen den Aufnahmezeitpunkten aufgetreten sind, können nicht erklärt werden. Ferner ist das System nicht in der Lage, Nutzen aus eventuell vorhandenem temporalen Wissen zu ziehen.

Ziel der vorliegenden Arbeit ist es deshalb, das Interpretationssystem dahingehend zu erweitern, daß auch multitemporale Luftbilder ausgewertet werden können. Die Analyse soll unter Berücksichtigung der jeweiligen Aufnahmezeitpunkte erfolgen. Wurde eine Szenenbeschreibung für einen Zeitpunkt  $t_k$  erstellt, so sollen die Erkenntnisse bei der Untersuchung des folgenden Zeitpunkts  $t_{k+1}$  berücksichtigt werden. Die zuletzt erstellte Interpretation kann, ähnlich wie GIS-Daten, als Vorwissen für die Analyse des aktuellen Bildes genutzt werden, da sie relativ zuverlässige Prognosen für die Lage der gesuchten Objekte liefert. Bild 3.1 zeigt die prinzipielle Vorgehensweise.

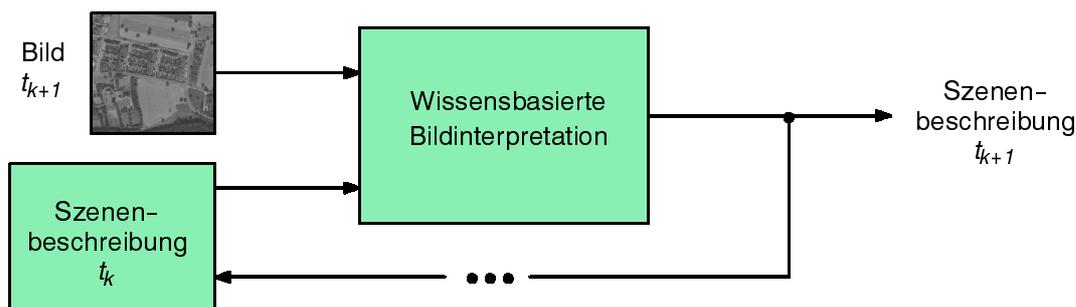


Bild 3.1: Nutzung der vorherigen Szenenbeschreibung für die Interpretation der aktuellen Sensordaten.

Die Kenntnis, daß zwischen den beiden Zeitpunkten  $t_k$  und  $t_{k+1}$  eine gewisse Zeit verstrichen ist, kann auf unterschiedliche Art und Weise verarbeitet werden. Der einfachste Weg ist die Annahme, daß die Objekte sich über die Zeit nicht verändert

---

haben. Folglich können während des Interpretationsvorgangs die Objekte der vorhergehenden Szenenbeschreibung in Bedeutung und Eigenschaften unverändert als neue modellgetriebene Hypothese übernommen werden. Diese Hypothesen werden anschließend in den aktuellen Daten überprüft. Kann man eine Hypothese nicht bestätigen, sind die aktuellen Daten entweder nicht dazu geeignet, das Objekt erneut zu detektieren, oder die Landschaft hat sich an der betrachteten Stelle verändert.

Das Problem der Datenunzulänglichkeit läßt sich prinzipiell nicht ganz eliminieren. Nur der Einsatz von qualitativ besseren Daten und die Nutzung mehrerer Sensoren kann die Analyseergebnisse verbessern. Die Veränderung der Landschaft hingegen läßt sich in gewissem Maße vorhersagen. So kann man allein aufgrund des Aufnahmezeitpunkts Vermutungen über die Farbe von Laubbäumen oder die Existenz von Schnee anstellen. Sollen aber strukturelle Veränderungen der Landschaft vorhergesagt werden, so benötigt man Wissen darüber, welche Änderungen in welchen Zeiträumen möglich oder zumindest wahrscheinlich sind. Eine Stelle, an der zuletzt ein Gebäude detektiert wurde, kann z. B. nicht eine Woche später bewaldet sein. Der Mensch verfügt über Wissen, welche Landschaftsveränderungen plausibel sind und welche nicht. Um die automatische Szeneninterpretation sicherer zu gestalten, soll derartiges temporales Wissen im System repräsentiert und nutzbar gemacht werden. Die folgenden Abschnitte behandeln zunächst notwendige Anforderungen an die Wissensrepräsentation und -nutzung. Anschließend werden Ansätze aus der Literatur zur Repräsentation und Verarbeitung von zeitlichen Zusammenhängen beschrieben. Schließlich erfolgt die ausführliche Diskussion der erforderlichen Erweiterungen am Bildinterpretationssystem AIDA.

### **3.1 Systemanforderungen für die Auswertung multitemporaler Luftbilder**

Die Verarbeitung zeitlicher Zusammenhänge in einem Bildinterpretationssystem erfordert die Repräsentation und die automatische Nutzung von temporalem Wissen. Wie auch bei der multisensoriellen Luftbildinterpretation gilt, daß nur solche Wissensinhalte repräsentiert werden sollen, die für die Analyse relevant sind und generell von einem automatischen Inferenzmechanismus ausgewertet werden können. Das temporale Wissen muß dazu geeignet sein, Veränderungen in der Landschaft zu einem gewissen Grad erklären und vorhersagen zu können. Im folgenden wird erörtert, welche zeitlichen Wissensinhalte in einem Bildinterpretationssystem repräsentiert werden sollten und auf welche Weise daraus bei der Analyse Nutzen gezogen werden kann.

### 3.1.1 Wissensrepräsentation

Bei der Repräsentation von temporalem Wissen muß zunächst die Frage beantwortet werden, welche Wissensinhalte existieren und welcher Teil davon für eine automatische Luftbildauswertung wichtig und nutzbar ist und somit im System repräsentiert werden soll. Da die Syntax der Wissensrepräsentation anwendungsunabhängig bleiben muß, ist ein allgemeingültiges Konzept zur Darstellung zeitlicher Zusammenhänge zu entwickeln. Anhand eines Beispiels werden die wesentlichen Aspekte temporaler Wissensinhalte erörtert:

Eine aktuelle Aufgabe der Europäischen Union besteht in der Überprüfung, ob die von den Landwirten gemeldeten (und von der EU geförderten) Brachlandflächen tatsächlich brachliegen oder nicht doch landwirtschaftlich genutzt werden [103][128][149]. Dieses wird in der Regel durch die interaktive Auswertung von Luft- oder Satellitenbildern realisiert. Allerdings läßt sich in einem einzigen Bild eine landwirtschaftliche Nutzfläche nicht zuverlässig von Brachland unterscheiden, da auch Brachland mit landwirtschaftlich nicht nutzbaren Pflanzen bewachsen sein kann. Die Auswertung einer Reihe von multitemporalen Bildern ermöglicht hingegen eine bessere Überwachung, da gegebenenfalls der gesamte Zyklus *Pflügen, Säen, Reifen* und *Ernten* beobachtet und das Gebiet somit als Brachland ausgeschlossen werden kann. Will man diese Analyse automatisieren, so ist es sinnvoll, den genannten Zyklus auch bei der automatischen Bildauswertung zu nutzen. Fraglich ist jedoch, ob alle Ereignisse des Zyklus überhaupt im Bild erkennbar sind. Eine aufgehende Saat und reifes Getreide lassen sich an ihrer charakteristischen Farbe, ein gepflügter Acker an den Ackerfurchen erkennen. Ein Unterschied zwischen einem Feld vor und nach dem Säen ist jedoch in Luftbildern in der Regel nicht zu sehen, so daß dieses Ereignis für die Bildauswertung nicht nutzbar ist und somit auch nicht in der Wissensbasis enthalten sein sollte.

Oftmals wird man nicht das Ereignis selbst, z. B. das *Pflügen*, beobachten können, sondern lediglich das Resultat des Ereignisses oder den Zustand, in dem sich das Objekt anschließend befindet. Die Zustände einer landwirtschaftlichen Nutzfläche könnten z. B. durch die Begriffe *gepflügtes Feld*, *bewachsenes Feld*, *geerntetes Feld* und *brachliegendes Feld* beschrieben werden. Visuell erkennbar sind diese Zustände wie bereits angedeutet an parallelen Ackerfurchen, an der charakteristischen Farbe bzw. an der Existenz von Stroh oder Strohbällen auf dem Feld.

Nach der Festlegung, welche Ereignisse oder Zustände in der Wissensbasis repräsentiert werden sollen und wie diese in den Bildern detektiert werden können, sind die zeitlichen Zusammenhänge zwischen ihnen zu definieren. Im genannten Beispiel wird das Feld zunächst gepflügt, bestellt und schließlich geerntet, bevor der Zyklus von

Neuem beginnt. Die Zustände haben demnach eine definierte Reihenfolge. Prinzipiell können in einer dynamischen Welt mehrere Ereignisse gleichzeitig eintreten und länger andauernde Zustände können sich überlappen. In der vorliegenden Anwendung der Luftbildinterpretation, bei der sich die Zustände auf dieselbe Region, z. B. ein bestimmtes Flurstück, beziehen, werden sie jedoch als zeitlich disjunkt angesehen. Eine Überlappung ist somit ausgeschlossen. Die Ereignisse oder Zustände folgen aufeinander, gegebenenfalls mit einer Übergangszeit, in der keines der definierten Ereignisse vorliegt.

Neben der Art der Zustände und ihrer Reihenfolge lassen sich eventuell Informationen über den *Anfangszeitpunkt des Zustands*, die *Dauer des Zustands* und die *Dauer des Übergangs* von einem Zustand zum Folgezustand machen. Ein Feld wird z. B. ab Herbst für einige Monate im Zustand *abgeerntet* bleiben, bis im Frühjahr gepflügt und gesät wird. Derartige zeitliche Angaben sollten ebenfalls in der Wissensbasis repräsentiert werden. Die Bestimmung eines Folgezustands ist nicht immer eindeutig, da oftmals mehrere Folgezustände möglich sind. Die Folgezustände sind aber in der Regel nicht gleich wahrscheinlich. So ist es zwar denkbar, daß aus einem abgeernteten Feld eine Baustelle und später ein Haus entsteht, aber wahrscheinlicher ist es, daß das Feld weiter landwirtschaftlich genutzt, also gepflügt und bestellt wird. Sofern es dem Experten möglich ist, *Wahrscheinlichkeiten für Zustände und Zustandsübergänge* anzugeben, so sollten sie auch in der Wissensbasis abgelegt und für die spätere Analyse genutzt werden.

Zusammenfassend können folgende Informationen als relevante temporale Wissensinhalte identifiziert werden:

- Menge  $\mathcal{S}$  zeitlich disjunkter Ereignisse bzw. Zustände  $S_i \in \mathcal{S}$
- Charakteristische, im Bild detektierbare Merkmale der Zustände  $\mathcal{F}_i$
- Mögliche Zustandsübergänge (Reihenfolge der Zustände)  $T_{i,j}$
- Anfangszeitpunkt  $t_0^{S_i}$  und Dauer  $d^{S_i}$  der Zustände
- Dauer der Zustandsübergänge  $d^{T_{i,j}}$
- Wahrscheinlichkeit der Zustände  $P(S_i)$
- Wahrscheinlichkeit der Zustandsübergänge  $P(T_{i,j}) = P(S_j|S_i)$

Eine grafische Darstellung der genannten Größen zeigt Bild 3.2. Ein Zustand  $S_i$  wird als zeitliches Intervall mit einem Anfangszeitpunkt  $t_0^{S_i}$  und einer Dauer  $d^{S_i}$  interpretiert.

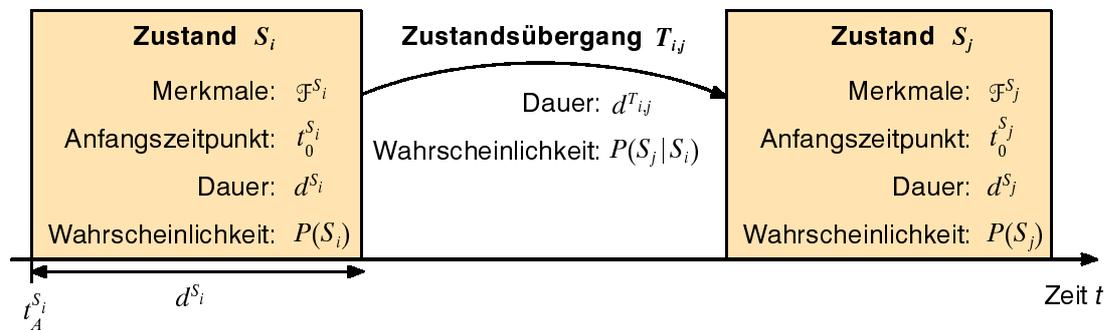


Bild 3.2: Modellierung von temporalem Wissen durch Zustände und Zustandsübergänge

Die Zustände können nahtlos aufeinander folgen, sich jedoch nicht überlappen. Der Übergang zwischen zwei Zuständen kann aber auch eine gewisse Zeit in Anspruch nehmen. So wird z. B. der Übergang eines Wohngebiets in einen Wald mehrere Jahre dauern.

Die Netzwerksprache des beschriebenen Bildinterpretationssystems ist derart zu erweitern, daß alle genannten temporalen Wissensinhalte gespeichert werden können. Dabei muß berücksichtigt werden, daß Anfangszeitpunkt, Dauer und Wahrscheinlichkeit von Zuständen und Zustandsübergängen in der Regel nur unsicher oder gar nicht angegeben werden können.

### 3.1.2 Wissensnutzung

Das temporale Wissen soll während der Bildinterpretation dazu genutzt werden, die Effizienz des Analysevorgangs und die Qualität der resultierenden Szenenbeschreibung zu steigern. Dieses wird erreicht, indem die Interpretationsergebnisse für das vorhergehende Bild mit in den Analyseprozeß des aktuellen Bildes einbezogen werden. Wurde für eine Region im letzten Bild ein bestimmter Zustand detektiert, so kann für das aktuelle Bild mit Hilfe der Information, wieviel Zeit zwischen den beiden Aufnahmen verstrichen ist, und durch Ausnutzung des temporalen Wissens ein oder mehrere Folgezustände für diese Region prädiert werden. Eine Lösung kann dabei auch sein, daß der Zustand unverändert geblieben ist. Somit deckt das Verfahren auch den einfachen Ansatz aus Kapitel 3, die Annahme der sich nicht ändernden Szene, ab. Kann jedoch aufgrund des Zusatzwissens gezielt ein Folgezustand vorhergesagt oder ein anderer Zustand sicher ausgeschlossen werden, so wird die Analyse effizienter, da die entsprechenden Alternativen nicht verfolgt werden müssen. Es ist jedoch darauf zu achten, daß die Erzeugung der Folgezustände optimistisch erfolgen muß. Um die wahre Lösung nicht übersehen zu können, müssen alle denkbaren Folgezustände – aber nicht die unmöglichen – untersucht werden.

---

Das temporale Wissen dient also dazu, modellgetriebene Hypothesen über Folgezustände der Szenenobjekte zu generieren, die im weiteren Verlauf der Analyse in den Daten verifiziert werden. Je exakter die Prädiktion, desto schneller kann die richtige Szenenbeschreibung gefunden werden. Hierzu trägt auch die intelligente Auswertung der gegebenenfalls vorhandenen Zustands- und Zustandsübergangswahrscheinlichkeiten bei. Sie sollten genutzt werden, die möglichen Alternativen in der Reihenfolge absteigender Wahrscheinlichkeiten zu untersuchen. Dadurch kann im Mittel die Analysezeit reduziert werden.

Die wesentlichen Vorteile, die man sich von der Nutzung des temporalen Wissens bei der Auswertung multitemporaler Bilder verspricht, sind im folgenden zusammengefaßt:

- Prädiktion von Folgezuständen für die Szenenobjekte ausgehend von einem zuvor detektiertem Zustand und der verstrichenen Zeit:
  - Generierung von Hypothesen für alle denkbaren Folgezustände
  - Ausschluß der unmöglichen Folgezustände
- Untersuchung der Alternativen in der Reihenfolge abnehmender Auftretenswahrscheinlichkeiten.

Nachdem die grundsätzlichen Aspekte zur Repräsentation und Nutzung von temporalem Wissen für die Bildinterpretation erörtert wurden, bleibt zu klären, wie das in Kapitel 2 beschriebene Bildinterpretationssystem AIDA modifiziert werden muß, um multitemporale Bilder auswerten zu können. Neben der Erweiterung der Netzwerksprache sind auch Eingriffe in die Systemsteuerung und die Entwicklung neuartiger Interpretationsstrategien notwendig.

Es muß eine geeignete Repräsentationsform für Zustände und deren Übergänge gefunden werden, die sich möglichst nahtlos in das semantische Netz des Interpretationssystems integrieren läßt. Der folgende Abschnitt gibt zunächst einen Überblick über verschiedene Ansätze aus der Literatur, die dazu geeignet sind, zeitliche Zusammenhänge darzustellen und aus ihnen Schlußfolgerungen zu ziehen. Mehrere dieser Ansätze aufgreifend werden in Kapitel 3.3 die realisierten Erweiterungen am Bildinterpretationssystem vorgestellt.

## **3.2 Ansätze zur Verarbeitung zeitlicher Zusammenhänge**

In den unterschiedlichsten wissenschaftlichen Disziplinen wird seit mehreren Jahrzehnten an der effizienten Repräsentation und Verarbeitung von zeitbezogenen Informatio-

nen geforscht. Vor allem der Bereich der Künstlichen Intelligenz versucht den natürlichen Umgang des Menschen mit der Zeit zu kopieren. Aber auch in der Betriebswirtschaftslehre und der Informationstechnik wird an zeitbezogenen Forschungsthemen gearbeitet. Erst die Existenz von Zeit erlaubt es, Veränderungen und dynamische Vorgänge zu beschreiben. Hat man das Ziel, derartige Prozesse im Rechner zu modellieren, so ist die Verarbeitung der Dimension Zeit zwingend erforderlich. Dennoch wird in vielen Programmsystemen die Repräsentation der Zeit vernachlässigt. Nicht umsonst behauptet McDermott in [95]:

*“Dealing with time correctly would change everything in an AI program.”*

Der Schwerpunkt aller Forschungsarbeiten liegt in der Repräsentation von temporalen Randbedingungen (engl.: *temporal constraints*) und dem zeitlichen Schlußfolgern (engl.: *temporal reasoning*). Ein Anwendungsgebiet ist z. B. die medizinische Diagnose, bei der für eine optimale Therapie berücksichtigt werden soll, wann und in welcher Reihenfolge die Krankheitssymptome aufgetreten sind. Desweiteren ist das Gebiet der Planung zu nennen. Aus dem Wissen über die Dauer einzelner Vorgänge und Arbeitsschritte soll automatisch ein Zeitplan erstellt werden. Im Bereich der industriellen Prozeßüberwachung wird angestrebt, die Geschichte der Prozeßvariablen und der vergangenen Aktionen bei der Prozeßsteuerung zu berücksichtigen. Schließlich wird seit kurzer Zeit an temporalen GIS-Systemen geforscht, deren Geo-Daten nicht nur eine Momentaufnahme der Welt darstellen, sondern die auch zeitliche Veränderungen dokumentieren und zeitlich-räumliche Anfragen gestatten [61][118].

In allen genannten Anwendungsgebieten werden die temporalen Informationen letztlich für eine oder mehrere der folgenden Aufgaben genutzt [157]:

- *Konsistenzprüfung*: Das temporale Wissen wird, vor allem nach dem Einfügen neuer Information, auf Konsistenz geprüft. Inkonsistente Teile werden lokalisiert.
- *Beantwortung zeitlicher Fragen*: Mit Hilfe des temporalen Wissens können Fragen bezüglich der Zeit beantwortet werden. Diese reichen von der Lokalisierung eines Faktums in der Zeit bis zur Bestimmung, wann eine Menge von Behauptungen gleichzeitig wahr wird.
- *Erklärung*: Man versucht, den aktuellen Zustand der Welt über die Erklärung der Vergangenheit zu rechtfertigen.
- *Prädiktion*: Aus dem temporalen Wissen wird eine Prognose über den zukünftigen Zustand der Welt erstellt.
- *Lernen*: Aus Beschreibungen der Welt zu unterschiedlichen Zeitpunkten sollen automatisch Regeln über die Veränderung abgeleitet werden.

Für die vorliegende Anwendung der automatischen Auswertung von multitemporalen Luftbildern ist vor allem der Bereich der Prädiktion relevant. Ausgehend von den aktuellen Informationen aus der letzten symbolischen Szenenbeschreibung sollen mit Hilfe des temporalen Wissens Prognosen über die Zustände der Objekte zum nächsten Aufnahmezeitpunkt erstellt werden, um die Bildinterpretation effizienter steuern zu können. Hierzu ist ein geeignetes Verfahren zur Repräsentation und Nutzung der zeitlichen Zusammenhänge zu wählen.

Im folgenden werden unterschiedliche Methoden zur Modellierung von zeitlichen Zusammenhängen vorgestellt. Nach einer kurzen Abhandlung über Zustandsübergangsdigramme und Markov-Ketten wird ein Überblick über die Ansätze des zeitlichen Schlußfolgerns im Bereich der Künstlichen Intelligenz gegeben. Darüber hinaus werden die Verfahren der Netzplantechnik, die vor allem in der Betriebswirtschaftslehre eingesetzt werden, vorgestellt.

### 3.2.1 Zustandsübergangsdigramme und Markov-Ketten

Zustandsübergangsdigramme werden vor allem in der Systemtheorie genutzt, um die logischen Zusammenhänge von Prozessen darzustellen. Ein Beispiel ist die Beschreibung endlicher Automaten (engl.: *Finite State Machine*).

#### 3.2.1.1 Zustandsübergangsdigramme zur Beschreibung endlicher Automaten

Endliche Automaten [2][60] beschreiben Systeme mit dedizierten Zuständen, die aufgrund äußerer Signale, z. B. durch Eingabe von Symbolen, ihren Zustand verändern. Formal sind sie definiert durch:

- eine endliche Menge  $\mathcal{S}$  von Zuständen  $S_i \in \mathcal{S}$ ,
- ein Eingabealphabet  $\Sigma$ , das auch das leere Zeichen  $\emptyset$  enthalten kann,
- einer Übergangsrelation  $next : (\mathcal{S} \times \Sigma) \rightarrow \mathcal{S}$ , die die Zustände ineinander überführt,
- einen Startzustand  $S_0 \in \mathcal{S}$  und
- einer Menge von Endzuständen  $S_\infty \subseteq \mathcal{S}$

Ist  $(S_i, x, S_{i+1}) \in next$ , so nennt man  $S_{i+1}$  Folgezustand von  $S_i$  bei Eingabe von  $x$  ( $x \in \Sigma$ ). Es ist möglich, daß  $S_i = S_{i+1}$  gilt und somit ein Zustand sein eigener Nachfolger ist.

Die erlaubten Zustandsübergänge eines Automaten lassen sich entweder durch eine Zustandsübergangstabelle beschreiben oder aber graphisch, und damit anschaulich, in

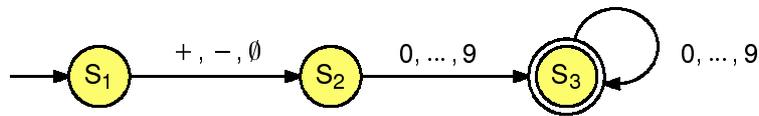


Bild 3.3: Zustandsübergangsdiagramm zur Beschreibung eines endlichen Automaten

einem *Zustandsübergangsdiagramm* darstellen. Das Diagramm ist ein Graph, in dem die Zustände durch Knoten und die Übergänge durch Kanten repräsentiert werden, die mit den Eingangssymbolen beschriftet sind. Start- und Endzustände werden besonders markiert. Bild 3.3 zeigt ein Beispiel für ein Zustandsübergangsdiagramm eines einfachen Automaten mit drei Zuständen, der die Eingabe positiver und negativer ganzer Zahlen mit der Mindestlänge 1 erlaubt.

Die Zustandsübergänge sind nicht an bestimmte Zeiten oder Zeitspannen gebunden, sondern sie treten aufgrund von Ereignissen, hier die Eingabe von Symbolen, auf. Es werden demnach keine expliziten zeitlichen Aspekte modelliert, wenn auch die Ereignisse, die die Übergänge auslösen, in einem festen zeitlichen Takt auftreten oder durch ein Zeitgeber-Signal beschrieben sein können.

### 3.2.1.2 Markov-Ketten

In der Signaltheorie versucht man, natürliche Signalquellen durch Modelle zu beschreiben, um Meßergebnisse analysieren, simulieren und vorhersagen zu können. Neben den deterministischen Signalmodellen haben vor allem die stochastischen Modelle an Bedeutung gewonnen. Stochastische Signalmodelle beschreiben ein Signal als einen parametrisierten Zufallsprozeß. Es wird ein Modell des Zufallsprozesses ausgewählt und die Erwartungswerte der Modellparameter werden anhand der oftmals stichprobenartigen Signalmessungen geschätzt. Vertreter stochastischer Modelle sind z. B. Gaußsche-Prozesse, Poisson-Prozesse und Markov-Prozesse.

Diskrete Markov-Prozesse werden *Markov-Ketten* [2][10][60] genannt. Sie lassen sich als ein diskreter stochastischer Automat beschreiben und haben somit eine große Ähnlichkeit mit den beschriebenen Zustandsübergangsdiagrammen. Markov-Ketten modellieren die Signalquelle als einen Automaten, der ausgehend vom aktuellen Zustand mit einer gewissen Wahrscheinlichkeit in einen anderen Zustand übergeht. Die Markov-Bedingung besagt dabei, daß der Automat ein Gedächtnis der Länge 1 hat, also die Übergangswahrscheinlichkeit nur vom aktuellen Zustand und nicht von vorherigen Zuständen abhängt. Formal ausgedrückt gilt demnach:

$$P(S_j | S_i, S_k, \dots) = P(S_j | S_i) \quad (3.1)$$

Markov-Ketten lassen sich ähnlich wie Zustandsübergangsdiagramme in Form eines Graphen visualisieren. Die Kanten des Graphen werden hier jedoch mit den Übergangswahrscheinlichkeiten beschriftet. Aufgrund der wahrscheinlichkeitstheoretischen Gesetze addieren sich alle Übergangswahrscheinlichkeiten eines Zustands zu 1 (Gl. (3.2)). Aus ihnen lassen sich gemäß Gl. (3.3) Zustandswahrscheinlichkeiten berechnen, die sich wiederum über alle Zustände zu 1 aufsummieren (Gl (3.4)).

$$\sum_{j=1}^n P(S_j|S_i) = 1 \quad \forall i = 1 \dots n \quad (3.2)$$

$$P(S_i) = \sum_{j=1}^n P(S_i|S_j) \cdot P(S_j) \quad \forall i = 1 \dots n \quad (3.3)$$

$$\sum_{i=1}^n P(S_i) = 1 \quad (3.4)$$

Markov-Ketten enthalten keine zeitlichen, dafür aber stochastische Abhängigkeiten von Systemzuständen. Sie werden u. a. im Bereich der Quellencodierung eingesetzt.

### 3.2.2 Zeitliches Schlußfolgern in der Künstlichen Intelligenz

In dem Forschungsgebiet der Künstlichen Intelligenz (KI) begann die Analyse und Modellierung der Zeit mit einigen anwendungsorientierten Arbeiten wie z. B. das System *Chronos* [18] im Bereich der Spracherkennung oder *Time Specialist* [71] bei der medizinischen Entscheidungsunterstützung. Erst in den frühen achtziger Jahren wurden allgemeine Theorien zum Thema Zeit entwickelt [4][95][158]. In den eher theoretischen Veröffentlichungen werden logische Formalismen beschrieben, Algorithmen zum zeitlichen Schlußfolgern und ihre Komplexität vorgestellt oder Repräsentationsformen für zeitliche Änderungen vorgeschlagen. Die temporalen Theorien werden vor allem in Systemen zur Spracherkennung [45][165], in Planungssystemen [6][119][155] und medizinischen Expertensystemen [72][109][122][123] eingesetzt.

Ein wesentliches Unterscheidungsmerkmal der entwickelten Theorien ist die Modellierung von Zeit. Einige Ansätze nutzen den Zeitpunkt, andere das Zeitintervall als temporales Primitiv. Die Zeit wird in einigen Arbeiten als diskrete Menge von Zeitelementen betrachtet. Andere Autoren modellieren sie als kontinuierliche Größe, was die Komplexität des zeitlichen Schlußfolgerns in der Regel erhöht.

McDermott [95] definiert die Zeit als unendliche, kontinuierliche Menge von Zeitpunkten. Ein Zeitintervall wird durch ein Paar von Zeitpunkten repräsentiert. James

F. Allen hingegen sieht das Zeitintervall als einziges temporales Element. Zeitpunkte können in seinen Augen als sehr kurze Zeitintervalle interpretiert werden. In [3] entwickelt er eine Intervall-Logik basierend auf dreizehn Relationen, die zwischen zwei Zeitabschnitten existieren können.

Dennoch ist der Begriff des Augenblicks in unserem Zeitverständnis allgegenwärtig. Formulierungen wie “die Temperatur des Patienten um 9:00 Uhr” lassen sich mit rein zeitintervallbezogenen Größen nicht beschreiben. Auch der abrupte Übergang von einem Zustand in einen Folgezustand (“wenn das Licht ausgeschaltet wird, wird ein beleuchteter Raum schlagartig dunkel”) kann mit Zeitintervallen nicht modelliert werden. Vilain [158] erweitert die Intervall-Logik von Allen um Zeitpunkte und Punkt-zu-Punkt- bzw. Punkt-zu-Intervall-Relationen. Allen selbst führt in [5] die sog. Momente ein. Seiner Theorie folgend sind Momente einerseits Zeitpunkte, die lang genug sind, um Ereignisse beinhalten zu können. Andererseits repräsentieren sie sehr kurze, nicht weiter zerlegbare Zeitintervalle mit Anfangs- und Endpunkt und einer Dauer ungleich Null. Die Erweiterung der Theorie um Momente macht das gesamte Modell jedoch weniger intuitiv und ändert nichts daran, daß augenblickliche Zustandsänderungen nicht beschrieben werden können.

Theorien, wie die von Allen, beschreiben rein qualitative Beziehungen zwischen den zeitlichen Elementen Zeitpunkt oder -intervall. Oftmals liegen aber konkrete, quantitative Zeitangaben vor. Im einfachsten Fall handelt es sich dabei um absolute Zeitreferenzen wie ein Datum oder eine Uhrzeit. Jedes Element erhält dadurch einen Zeitstempel. Die Dauer eines Intervalls oder der Abstand zweier Zeitpunkte kann leicht durch Subtraktion der entsprechenden Zeitstempel bestimmt werden.

Oftmals stehen keine präzisen numerischen Informationen zur Verfügung. Vielmehr liegen lediglich Informationen über den zeitlichen Abstand zwischen den Ereignissen vor. In solchen Fällen haben sich gerichtete Graphen als geeignete Repräsentationsformen erwiesen, in denen die Knoten die Ereignisse und die Kanten die zeitlichen Distanzen zwischen den Ereignissen modellieren. Einen derartigen Ansatz stellen Dean und McDermott in [27] vor. Zeitliche Zusammenhänge werden in sog. *Time Maps* graphisch dargestellt, wobei unpräzisen Zeitangaben durch minimale und maximale Distanzangaben in den Kanten des Graphen Rechnung getragen wird. Vollkommene Unwissenheit über die quantitative Distanz zweier Ereignisse wird durch die Angabe  $[0, \infty]$  modelliert, wodurch die Relation zu einer rein qualitativen Reihenfolgebeziehung entartet. Einen ähnlichen, mathematisch fundierten Ansatz beschreiben Dechter et al. in [29] mit dem sog. *Temporal Constraint Network*. Im Gegensatz zu [27] werden

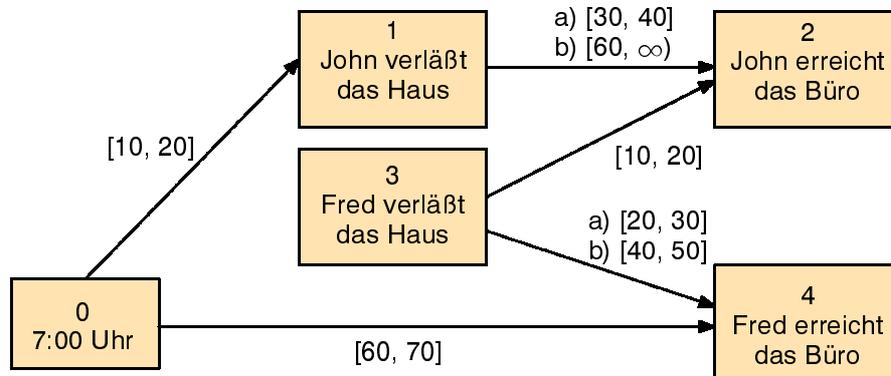


Bild 3.4: “Temporal Constraint Network” für das Beispiel 4.1

mehrere Distanzangaben an den Kanten des Graphen zugelassen. Betrachtet sei folgendes Beispiel aus [29]:

**Beispiel 4.1:** *John fährt entweder mit dem Auto (30-40 Minuten) oder dem Bus (mindestens 60 Minuten) zur Arbeit. Fred benutzt entweder das Auto (20-30 Minuten) oder eine Fahrgemeinschaft (40-50 Minuten). Heute verläßt John das Haus zwischen 7:10 und 7:20 Uhr und Fred trifft im Büro zwischen 8:00 und 8:10 Uhr ein. 10-20 Minuten nachdem Fred sein Haus verläßt, erreicht John das Büro. Gesucht ist eine Lösung dafür, wann die beiden Personen das Haus verlassen und wann sie im Büro angekommen sind. Welches Transportmittel haben sie dabei gewählt ?*

Die Fakten des obigen Beispiels lassen sich in einem *Temporal Constraint Network* formulieren (Bild 3.4). Je nach gewähltem Transportmittel gelten unterschiedliche Distanzen zwischen den jeweiligen Knoten des Graphen. Die Autoren schlagen vor, derartige Graphen in einfachere zu zerlegen, die jeweils nur eine Zeitangabe pro Kante besitzen. Im vorliegenden Beispiel bedeutet dies, man betrachtet die vier möglichen Kombinationen der Transportmittel als separate Graphen. Anschließend werden mit Hilfe des Verfahrens von Floyd-Warshall [105] die kürzesten Pfade durch den Graphen berechnet. Dabei werden die frühesten Anfangszeitpunkte und spätesten Endzeitpunkte der einzelnen Knoten bestimmt. Angenommen John benutzt das Auto und Fred die Fahrgemeinschaft, dann ergeben sich folgende kürzeste Pfade durch den Graphen:

Start-/Endknoten	0	1	2	3	4
0	0	[10, 20]	[40, 50]	[20, 30]	[60, 70]

Die Zahlen repräsentieren die gültigen Zeiträume der einzelnen Ereignisse in Minuten nach 7:00 Uhr. Eine Lösung wäre somit, daß John das Haus um 7:10 Uhr verläßt und um 7:50 Uhr das Büro mit dem Auto erreicht, während Fred um 7:30 Uhr

mit der Fahrgemeinschaft aufbricht und um 8:10 Uhr im Büro eintrifft. Eine entsprechende Rechnung für den Fall, daß John den Bus benutzt, führt zu Widersprüchen. Diese Kombination der Transportmittel ist demnach nicht möglich.

Man erkennt, daß mit Hilfe der graphenbasierten Techniken qualitative und quantitative zeitliche Zusammenhänge repräsentiert werden können, auf deren Basis auch ein zeitliches Schlußfolgern möglich ist. Quantitative Zeitangaben sind in der Regel mit einer gewissen Ungenauigkeit und Unsicherheit verbunden. Einige Autoren versuchen dieser Tatsache mit probabilistischen [28] und possibilistischen Ansätzen [35][36][37] Rechnung zu tragen.

Abschließend läßt sich festhalten, daß eine Vielzahl von Ansätzen zur Modellierung und Nutzung von zeitlichen Zusammenhängen existieren, von denen keiner als optimal bezeichnet werden kann. Es herrscht zwar allgemein Einigkeit darüber, daß das Zeitintervall, das die Dauer von Ereignissen beschreibt, die fundamentale temporale Größe ist. Dennoch ist die Intervall-Logik in vielen Fällen zu komplex, so daß eine Modellierung der Zeit auf Basis von Zeitpunkten angemessen erscheint. Deshalb muß für jede Anwendung neu entschieden werden, welcher Ansatz sich hinsichtlich seiner Präzision und seines Aufwands am besten eignet.

### 3.2.3 Netzplantechnik

In vielen wirtschaftlichen Bereichen müssen Projekte oder andere Vorhaben geplant, gesteuert und überwacht werden. Nachdem sich traditionelle Planungstechniken als unzulänglich erwiesen hatten, setzte ab 1957 zunächst in den USA und später auch in Europa die Entwicklung eines neuen Ansatzes ein. Basierend auf den Erkenntnissen der Graphentheorie und der aus der Elektrotechnik bekannten Netzwerkanalyse entstanden die Verfahren der sog. *Netzplantechnik* [59][133]. Die neuen Planungstechniken waren von Anfang an sehr erfolgreich. So erzielte die US-Navy bei der Vorbereitung und Überwachung des Polaris-Raketen-Projekts durch Anwendung der neuen Verfahren einen Zeitgewinn von zwei Jahren.

Die beiden wichtigsten und bekanntesten Verfahren wurden unabhängig voneinander in den USA entwickelt. Dabei handelt es sich um CPM (*Critical Path Method*) und PERT (*Program Evaluation and Review Technique*). Etwa zur selben Zeit entstand in Frankreich das Verfahren MPM (*Metra-Potential-Method*). Allen Verfahren ist gemein, daß die Strukturelemente eines Projekts in einem gerichteten Graphen, dem sog. Netzplan, repräsentiert werden, der für die Ablauf-, Zeit-, Kapazitäts- und Kostenanalyse genutzt werden kann.

Die Strukturelemente eines Projekts sind *Vorgänge* und *Ereignisse*. Vorgänge stellen Zeit und Betriebsmittel in Anspruch nehmende Geschehnisse mit definiertem Anfang und Ende dar. Ereignisse sind wohldefinierte Zustände im Projektablauf. Vorgänge und Ereignisse des Projekts werden auf Knoten und Kanten des Netzplans abgebildet, wobei die Folgebeziehungen im Graphen die Reihenfolge der Vorgänge widerspiegeln. Dabei ist es jedoch nicht wie bei Zustandsübergangsdiagrammen zulässig, daß ein Vorgang sein eigener Nachfolger ist.

Je nach Abbildung auf die Graphenelemente unterscheidet man *Vorgangsknoten-*, *Vorgangspfeil-* und *Ereignisknotennetze*. In Vorgangsknotennetzen werden die Vorgänge durch Knoten des Netzplans modelliert, die entsprechend des Zeitverbrauchs des Vorgangs einen Anfangs- und Endzeitpunkt haben. Die Kanten setzen die Vorgänge miteinander in eine Reihenfolgebeziehung. Ein Repräsentant dieses Ansatzes ist das eingangs erwähnte MPM. Bei Vorgangspfeilnetzen hingegen, wie sie bei der CPM-Technik genutzt werden, stellen die Netzwerkkanten den Zeitverbrauch dar, während die Knoten einen Zeitpunkt oder ein Ereignis widerspiegeln. Die Ereignisknotennetze ähneln den Vorgangspfeilnetzen. Anstelle der Vorgänge bilden sie jedoch Ereignisse auf Knoten ab. Der bekannteste Vertreter dieser Gattung ist das PERT-Verfahren. Ereignisknotennetze werden an dieser Stelle nicht näher behandelt. Die Unterschiede zwischen Vorgangspfeil- und Vorgangsknotennetzen sollen anhand eines Beispiels aus [133] veranschaulicht werden.

**Beispiel 4.2:** *Drei Freunde planen eine gemeinsame Reise mit dem Auto ins Ausland. Zur Planung ihrer Reisevorbereitungen erstellen sie ein Liste der notwendigen Vorgänge und der geschätzten Vorgangsdauern (s. Tab. 3.1). Die Freunde einigen sich auf folgende Aufteilung: Der Erste beschafft zunächst die Visa und reserviert anschließend das Hotel. Der Zweite kümmert sich um die Ausrüstung, während der Dritte das Auto überprüft. Der Erste und der Zweite sollen gemeinsam nach Erledigung ihrer Aufgaben das Proviant besorgen.*

Tabelle 3.1 *Geplante Vorgänge und geschätzte Vorgangsdauern aus Beispiel 4.2*

<b>Vorgang:</b>	<b>Dauer in Tagen:</b>
Visabeschaffung	5
Hotelreservierung	15
Überprüfung/Ergänzung der Ausrüstung	25
Überprüfung des Autos	20
Proviantbeschaffung	5

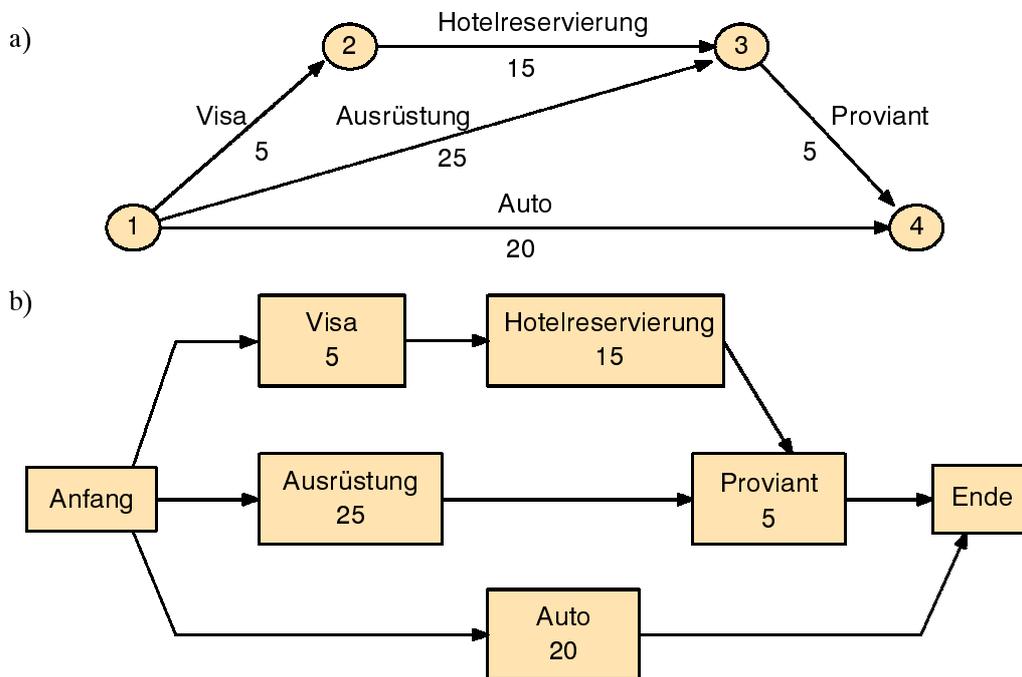


Bild 3.5: a) Vorgangspfeil- und b) Vorgangsknotenetz zu Beispiel 4.2

Die genannten Fakten lassen sich sowohl als Vorgangspfeil- (Bild 3.5a) als auch als Vorgangsknotenetz anschaulich darstellen (Bild 3.5b). Im Vergleich zu den sonst üblichen Zeit-Balken-Diagrammen haben Netzpläne den Vorteil, daß die technischen, wirtschaftlichen und logischen Abhängigkeiten zwischen den Vorgängen berücksichtigt werden. Darüber hinaus ist es vorteilhaft, daß nur relative Zeitangaben enthalten sind, so daß eine Terminverschiebung, wie z. B. die Verzögerung des Beginns der Reisevorbereitungen, keinen Einfluß auf den Netzplan hat.

Zwischen zwei Vorgängen mit definiertem Anfangs- und Endzeitpunkt sind vier Anordnungsbeziehungen möglich. Bislang wurden ausschließlich die sog. *Normalfolge* berücksichtigt, die aussagt, daß ein Vorgang *B* frühestens *a* Zeiteinheiten nach dem Ende von Vorgang *A* beginnen kann (in der Praxis gilt zumeist  $a=0$ ). Darüber hinaus existieren folgende Beziehungen:

- *Anfangsfolge*: *B* kann frühestens *a* Zeiteinheiten nach dem Beginn von *A* starten
- *Endfolge*: *B* muß spätestens *a* Zeiteinheiten nach dem Ende von *A* enden.
- *Sprungfolge*: *B* muß spätestens *a* Zeiteinheiten nach dem Anfang von *A* enden.

Durch die Zeitangabe *a* werden Abstandsbeziehungen zwischen den Anfangs- bzw. Endzeitpunkten der Vorgänge definiert, wodurch sich z. B. vorgeschriebene Wartezei-

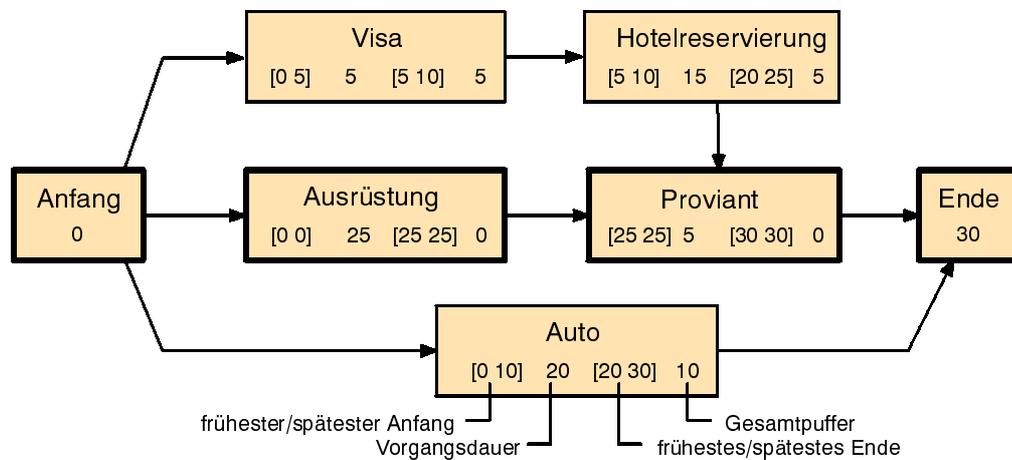


Bild 3.6: Vorgangsknotennetz mit Anfangs- und Endzeitpunkten sowie kritischem Pfad (fett gedruckt)

ten zwischen zwei aufeinanderfolgenden Arbeitsschritten wie Lackieren und Verpacken beschreiben lassen. Kann der zeitliche Abstand nicht exakt angegeben werden, so sind obere und untere Grenzen für den Abstand zu definieren. Im Vorgangsknotennetz werden die unterschiedlichen Anordnungsbeziehungen mit ihren Abstandsangaben durch entsprechende Annotation der Kanten markiert (s. DIN 69900). Vorgangspfeilnetze sind für die Darstellung derartiger Relationen nicht geeignet.

Ziel der Netzplantechnik ist nicht allein die Repräsentation der relativen Zeitabhängigkeiten, sondern vielmehr die Ableitung absoluter Zeitpunkte oder Intervalle, in denen Vorgänge beginnen oder enden müssen. Aus den Vorgangsdauern und den Anordnungsbeziehungen lassen sich mit Hilfe der sog. *Vorwärts-* und *Rückwärtsrechnung* [133] die *frühesten* und *spätesten* Anfangs- bzw. *Endzeitpunkte* der Vorgänge berechnen. Aus dem frühesten Anfang des Endknotens ergibt sich die *minimale Projektdauer*. Die Differenz zwischen frühestem und spätestem Anfangszeitpunkt eines Vorgangs wird *Gesamtpuffer* genannt. Sie beschreibt den zeitlichen Spielraum, der für die Durchführung des Vorgangs besteht. Eine Verzögerung des Vorgangs, die kleiner als die Gesamtpuffer bleibt, hat keine Verlängerung der Gesamtprojektdauer zur Folge, vorausgesetzt die folgenden Vorgänge werden ihrerseits nicht verzögert oder verlängert. Die Vorgänge, deren Gesamtpuffer null ist, bilden den *kritischen Pfad* im Netzplan. Ihre Verzögerung oder Verlängerung wirkt sich direkt auf die Projektdauer aus. In Bild 3.6 wurde das Vorgangsknotennetz aus Bild 3.5b um die Anfangs-, End- und Pufferzeiten erweitert. Die minimale Projektdauer beträgt 30 Tage. Den kritischen Pfad bilden hier die Vorgänge der Ausrüstungsüberprüfung und der Proviantbeschaffung.

Ein gravierendes Problem bei der Netzplantechnik ist, daß sich die Vorgangsdauern oftmals nur vage angeben oder abschätzen lassen. Es existieren zahlreiche Ansätze zur Verarbeitung unsicherer Zeitangaben. Dabei werden sowohl stochastische Modelle wie gleich-, dreiecks- oder normalverteilte Zufallsvariablen [132][150] als auch Fuzzy-Mengen [111][115] eingesetzt.

### 3.3 Erweiterung des Bildinterpretationssystems

Aus den vorgestellten Verfahren zur Repräsentation und Verarbeitung zeitlicher Zusammenhänge wird im folgenden ein Ansatz abgeleitet, der die Auswertung multitemporaler Luftbilder mit dem in Kapitel 2 beschriebenen Bildinterpretationssystem AIDA ermöglicht. Er umfaßt sowohl die Erweiterung der Wissensrepräsentation als auch die Modifikation der Wissensnutzung.

Das temporale Wissen soll analog zum strukturellen und topologischen Wissen innerhalb des semantischen Netzes gespeichert werden können. Aus diesem Grund erscheinen graphenbasierte Ansätze zum zeitlichen Schlußfolgern - wie die erwähnten Zustandsübergangsdigramme, Markov-Ketten, *Temporal Constraint Networks* und Netzpläne - grundsätzlich dazu geeignet, in das semantische Netz integriert zu werden.

Grundlegende Voraussetzung zur Verarbeitung zeitpunktbezogener Informationen ist der Einsatz von Zeitstempeln. Im Gegensatz zum monotemporalen Fall, wo stillschweigend vorausgesetzt wird, daß alle auszuwertenden Bilder die Szene zum selben Zeitpunkt beschreiben, besitzen bei der multitemporalen Auswertung alle Luftbilder einen Zeitstempel, der durch den jeweiligen Aufnahmezeitpunkt definiert ist. Die Aufnahmezeitpunkte und deren zeitlichen Abstände sollen explizit mit in den Interpretationsprozeß einbezogen werden.

Allein das Wissen um den Aufnahmezeitpunkt kann bei der Interpretation eines einzelnen Luftbildes nützlich sein. Wird z. B. ein Bild analysiert, das im Herbst aufgenommen wurde, so erwartet man Laubbäume mit gelben und roten Blättern. In einer Aufnahme aus dem Frühjahr oder Sommer wird das Laub grün sein. Dieses Wissen läßt sich bei der modellgetriebenen Hypothesenbildung ausnutzen, indem z. B. gezielt die erwartete Farbe für Laubbäume aus dem Zeitstempel des zu analysierenden Bildes abgeleitet wird.

Liegt hingegen eine Reihe von multitemporalen Luftbildern vor, so lassen sich neben dem absoluten Aufnahmezeitpunkt auch die Zeitdifferenzen zwischen aufeinanderfolgenden Bildern ausnutzen. Je nach Länge der verstrichenen Zeit sind unterschiedliche Veränderungen in der Landschaft möglich. Welche Änderungen in welchen Zeiten plausibel sind, muß als temporales Wissen vor Beginn der Analyse in der Wissensbasis

---

abgelegt werden. Die Netzwerksprache des Bildinterpretationssystems ist so zu erweitern, daß dieses temporale Wissen in Form eines semantischen Netzes beschrieben werden kann. Darüber hinaus muß das System in der Lage sein, das temporale Wissen automatisch auf die zu untersuchenden Bilder anzuwenden, um mögliche Folgezustände für die einzelnen Objekte zu präzisieren. Beide Aspekte werden in den folgenden Abschnitten behandelt.

### 3.3.1 Repräsentation von Zustandsübergangsdigrammen

Die in Kapitel 3.1.1 genannten Anforderungen an die Beschreibung temporaler Wissensinhalte werden nun aufgegriffen und in ein konkretes Repräsentationsschema für das genannte Bildinterpretationssystem umgesetzt:

Ziel ist es, Informationen über Objektzustände und deren Zustandsübergänge in zeitlicher und probabilistischer Hinsicht zu erfassen. Zustandsübergangsdigramme enthalten keine Wahrscheinlichkeiten, Markov-Ketten wiederum keine Zeitangaben. In Netzplänen und *Temporal Constraint Networks* sind keine auf sich selbst zurückführenden Zustandsübergänge erlaubt und der Zeitverbrauch ist entweder auf die Knoten oder die Kanten beschränkt. Somit ist keines der vorgestellten Konzepte für die hier vorliegende Aufgabe direkt nutzbar. Die Repräsentation der zeitlichen Zusammenhänge erfolgt deshalb mit Hilfe von Zustandsübergangsdigrammen, die derart um Wahrscheinlichkeiten erweitert werden, daß die Bedingungen für Markov-Ketten eingehalten sind. Für die Nutzung des temporalen Wissens während der Bildinterpretation (s. Kapitel 3.3.2) werden die Prinzipien der Vorwärtsrechnung aus der Netzplantechnik eingesetzt, wobei sowohl Zustands- als auch Zustandsübergangsdauern berücksichtigt werden.

Das erweiterte Zustandsübergangsdigramm wird direkt in das semantische Netz integriert. Die Zustände werden durch Konzepte im semantischen Netz repräsentiert, die Zustandsübergänge durch entsprechende Kanten. Um sie von topologischen und hierarchischen Relationen unterscheiden zu können, wird eine neue Kante eingeführt: die *temporale Relation*. Ähnlich wie die attributierte Relation verläuft diese Kante horizontal im semantischen Netz und verbindet zwei Knoten derselben konzeptionellen Ebene. Da zeitliche Zusammenhänge nur sinnvoll zwischen Objekten mit symbolischer Bedeutung angegeben werden können und z. B. nicht zwischen rein geometrischen Objekten, dürfen temporale Relationen nur innerhalb der semantischen Ebene des Konzeptnetzes definiert werden (vgl. Bild 2.4). Die Einführung von Zustandsübergängen innerhalb der physikalischen oder der Sensorebene ist nicht zulässig.

Die Knoten des semantischen Netzes müssen im Vergleich zur Beschreibung in Kapitel 2.3.1.1 um einige Einträge erweitert werden, um die notwendigen zeitlichen Angaben repräsentieren zu können. Jeder Knoten erhält einen *Zeitstempel*, mit dem während der Analyse der Zeitpunkt der Instanzierung dokumentiert wird. Dadurch kann die Instanz eindeutig einem der multitemporalen Luftbilder zugeordnet werden. Ein Zeitstempel wird im Datumsformat *JJJJMMTT* (Jahr, Monat, Tag) angegeben.

Für den Fall, daß ein Konzept einen Zustand des Übergangsdiagramms repräsentiert, sind einige zusätzliche Zeitangaben erforderlich (Bild 3.7). Der Ungenauigkeit dieser Angaben wird durch Intervalle mit Minimal- und Maximalwerten Rechnung getragen. Für jeden Zustand kann eine *Zustandsdauer*  $d^{S_i}$  definiert werden. Unkenntnis bezüglich der Zustandsdauer wird durch das Intervall  $[0 \infty]$  dargestellt. Im Feld *Anfangszeitpunkt* läßt sich ein absoluter Zeitraum  $t_0^{S_i}$  für den zulässigen Beginn eines Zustands ablegen, was besonders bei saisonalen oder wiederkehrenden Ereignissen sinnvoll ist. Hier kennzeichnet das Intervall  $[0 \infty]$ , daß der Zustand zu jeder Zeit auftreten kann. Schließlich läßt sich eine *Zustandswahrscheinlichkeit*  $P(S_i)$  definieren, sofern über die Auftretenswahrscheinlichkeit des Zustands Expertenwissen existiert. Anderenfalls werden alle Zustände als gleichwahrscheinlich angenommen.

Durch die *temporale Relation* werden die möglichen Übergänge zwischen jeweils zwei Zuständen beschrieben. Neben der Information über den Ausgangs- und den Nachfolgezustand ist die Angabe einer *Zustandsübergangsdauer*  $d^{T_{ij}}$  möglich. So erfordert z. B. der Übergang eines Wohngebiets in einen Wald einige Zeit und ist nicht von einem Tag auf den anderen möglich. Vordefiniert ist das Intervall  $[1 1]$ , was in Tagesfrist aufeinanderfolgende Zustände beschreibt. Schließlich kann für eine temporale Relation eine bedingte *Zustandsübergangswahrscheinlichkeit*  $P(S_j | S_i)$  definiert werden. Analog zur Zustandswahrscheinlichkeit gilt, daß diese Angabe nur aus Expertenwissen bzw. -schätzungen oder statistischen Messungen stammen kann. Ansonsten sind alle Zustandsübergänge desselben Startzustands mit derselben Wahrscheinlichkeit zu versehen. Um eine konsistente Markov-Kette zu erhalten, ist darauf zu achten, daß die Bedingungen gemäß Gl. (3.2) – (3.4) eingehalten sind, sich also alle Zustandswahrscheinlichkeiten und alle Zustandsübergangswahrscheinlichkeiten zu 1 addieren.

Für eine exakte Modellierung der statistischen Gegebenheiten müßten streng genommen zeitliche Wahrscheinlichkeitsverteilungen genutzt werden. Mit solchen Verteilungsfunktionen könnte man beschreiben, daß die Wahrscheinlichkeit eines Zustands mit seiner zunehmenden Dauer abnimmt und schließlich bei einer Überschrei-

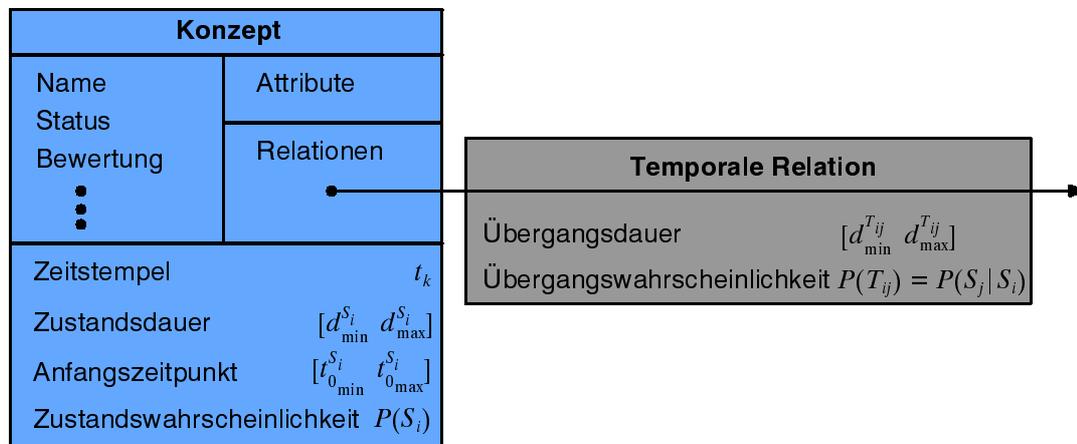


Bild 3.7: Repräsentation eines Zustandsübergangsdiagramms im semantischen Netz

tung der maximal erlaubten Zustandsdauer gegen Null geht. Wie im weiteren Verlauf der Ausführungen noch deutlich werden wird, kommt es jedoch an dieser Stelle nicht auf eine absolute Exaktheit der vorgegebenen Wahrscheinlichkeiten an. Vielmehr sollen sie während der Analyse lediglich dazu genutzt werden, um wahrscheinlichere Hypothesen bevorzugt untersuchen zu können. Somit sind nur die Größenverhältnisse der Wahrscheinlichkeiten untereinander relevant. Der Vorteil einer mathematisch korrekten Wahrscheinlichkeitsverteilung würde den Aufwand für deren Verarbeitung nicht rechtfertigen, zumal derartige Verteilungen noch schlechter zu messen oder zu schätzen wären. Aus diesem Grund werden lediglich skalare Wahrscheinlichkeiten benutzt.

### 3.3.1.1 Beispiele zur Repräsentation von temporalem Wissen

An zwei Beispielen soll im folgenden illustriert werden, wie sich mit dem beschriebenen Repräsentationsschema zeitliche Zusammenhänge modellieren lassen:

Zunächst wird das Beispiel aus Kapitel 3.1.1 aufgegriffen. Eine landwirtschaftliche Nutzfläche wurde dort durch die Zustände *gepflügetes Feld*, *bewachsenes Feld*, *geerntetes Feld* und *brachliegendes Feld* charakterisiert. Diese können in einem semantischen Netz mit entsprechenden Zustandsknoten dargestellt werden (Bild 3.8). Die jeweiligen Anfangszeitpunkte geben eine absolute zeitliche Referenz für die verschiedenen Phasen. So ist z. B. der Beginn des Zustands *geerntetes Feld* auf die Monate August und September mit einer Dauer von ein bis zwei Monaten beschränkt. Jeder Zustand wird durch seine spezifischen Merkmale wie z. B. parallele Ackerfurchen beschrieben. Die Zustandsübergänge modellieren den jährlichen Zyklus von Pflügen, Säen, Reifen, Ernten. Hier werden die temporalen Relationen genutzt. Die Zustandsän-

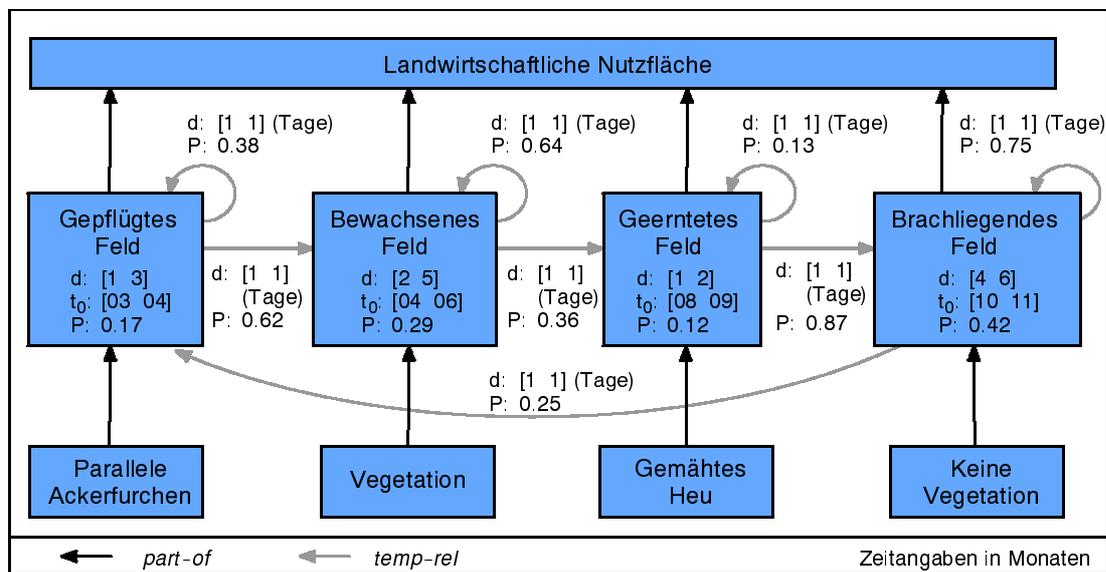


Bild 3.8: Semantisches Netz mit integriertem Zustandsübergangsdiagramm für die Erkennung von landwirtschaftlichen Nutzflächen

derungen können spontan, z. B. durch das Pflügen des Feldes, eintreten. Deshalb wird eine eintägige Zustandsübergangsdauer durch das Intervall [1 1] definiert. Die angegebenen Zustandswahrscheinlichkeiten entsprechen den mittleren Häufigkeiten mit der ein Zustand über das Jahr gesehen auftritt. Sie summieren sich zu 1. Gelingt es, in einer Reihe von Luftbildern alle vier Zustände für ein Flurstück zu detektieren, kann man die Existenz einer landwirtschaftlichen Nutzfläche bestätigen und somit ausschließen, daß es sich bei dieser Region um Brachland handelt.

Neben saisonalen Veränderungen, die also zu bestimmten Jahreszeiten eintreten, gibt es andere Fälle, bei denen die Zustandsänderungen zu beliebigen Zeitpunkten im Jahr auftreten können. Dafür haben aber die einzelnen Zustände eine feste Dauer. So läßt sich z. B. ein Messegelände durch eine typische Abfolge von Ereignissen mit relativ festen Zeiten charakterisieren: Nach einer Phase der Inaktivität werden im Vorfeld einer Messe die Messestände aufgebaut, was in der Regel ein bis zwei Wochen dauert. Anschließend findet für wenige Tage die eigentliche Messe statt. Nachdem die Stände wieder abgebaut sind, beginnt wieder die Zeit der Inaktivität. Die einzelnen Zustände eines Messegeländes lassen sich anhand charakteristischer Merkmale auch in Luftbildern erkennen. Während der Auf- und Abbauphasen stehen viele LKWs zwischen den Messehallen, die das Material für die Messestände transportieren. Die Parkplätze des Messegeländes sind jedoch noch relativ leer. Während der Messe hingegen sind die LKWs vom Gelände verschwunden, aber die Parkplätze sind durch die Fahrzeuge der Besucher belegt. In

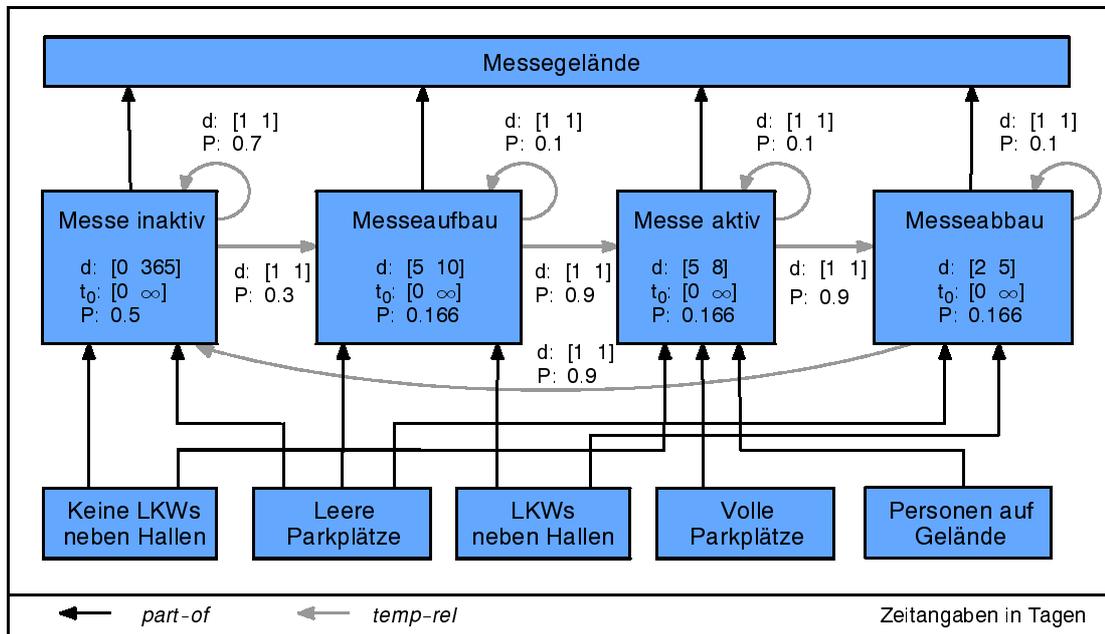


Bild 3.9: Semantisches Netz mit integriertem Zustandsübergangsdiagramm für die Erkennung eines Messegeländes

hochaufgelösten Luftbildern können sogar die Besucher selbst auf dem Messegelände ausgemacht werden. Findet keine Messe statt, so sind das Messegelände und die Parkplätze weitestgehend leer.

In Bild 3.9 sind diese Zusammenhänge in Form eines semantischen Netzes dargestellt. Der Zustand *Messe inaktiv* ist ein stabiler Zustand und kann unter der Annahme, daß mindestens einmal jährlich eine Messe stattfindet, bis zu einem Jahr andauern. Diese Tatsache wird durch die vorgegebene Zustandsdauer  $[0 \ 365]$  Tage repräsentiert. Die anderen Zustände haben eine Dauer von zwei bis zehn Tagen. Da die Zustände binnen eines Tages wechseln können, werden auch hier die Zustandsübergangsdauern auf einen Tag gesetzt. Die Anfangszeitpunkte der Zustände sind nicht eingeschränkt, was durch das Intervall  $[0 \ \infty]$  ausgedrückt wird. Über das Jahr gesehen ist der Zustand *Messe inaktiv* sicherlich der häufigste. Deshalb erhält er mit 0.5 die höchste Zustandswahrscheinlichkeit. Die restlichen Zustände werden als gleichwahrscheinlich angesehen. Auch durch die Übergangswahrscheinlichkeiten wird ausgedrückt, daß der Inaktivitätszustand stabil und die restlichen nur von kurzer Dauer und somit instabil sind. Erst wenn in einer Folge von multitemporalen Luftbildern alle Zustände in der richtigen Reihenfolge detektiert werden können, kann die Hypothese,

daß es sich um ein Messegelände handelt, bestätigt werden. Somit sind alle vier Zustände notwendige Bestandteile des Konzepts *Messegelände*.

Man erkennt, daß sich komplexe zeitliche Zusammenhänge im semantischen Netz repräsentieren lassen. Die folgenden Abschnitte beschreiben, wie das zusätzliche Wissen für die Bildinterpretation genutzt werden kann.

### 3.3.2 Modellgetriebene Prädiktion von Folgezuständen

Das Ziel der Einführung von temporalen Wissensinhalten ist es, zeitliche Veränderungen in den Bildern erklären oder aufgrund der zeitlichen Veränderungen ein komplexes Objekt, wie z. B. ein Messegelände, erkennen zu können. Das Zustandsübergangsdiagramm beschreibt dabei, welche Veränderungen im Laufe welcher Zeit und mit welcher Wahrscheinlichkeit möglich sind. Hat man im ersten Bild der Zeitreihe für ein Objekt einen bestimmten Zustand detektiert, so kann auf Basis dieser Erkenntnis für das folgende Bild eine Prädiktion über den Zustand dieses Objekts abgeleitet werden. Da eine solche Hypothese ausschließlich auf Basis des Vorwissens generiert wird, nennt man sie modellgetrieben, im Gegensatz zu einer datengetriebenen Vorgehensweise.

Eine modellgetriebene Prädiktion von Folgezuständen erhöht die Effizienz der Bildanalyse, da unmögliche Folgezustände ausgeschlossen werden können und so der Suchraum verkleinert wird. Wie bereits in Kapitel 3.1.2 erläutert, muß darauf geachtet werden, daß die Prädiktion optimistisch erfolgt, um bei der Analyse nicht fälschlicherweise einen möglichen Folgezustand zu übersehen.

Ausgehend vom aktuellen Zeitpunkt  $t_k$  und einem für ein bestimmtes Objekt detektierten Zustand  $S_i$  sollen für den Zeitpunkt des nächsten Bildes  $t_{k+1}$  die möglichen Folgezustände  $\{S_j\}$  dieses Objekts bestimmt werden. Dazu werden, ähnlich wie in der Netzplantechnik, der früheste Anfangszeitpunkt  $t_{A_{\min}}^{S_j}$  und der späteste Endzeitpunkt  $t_{E_{\max}}^{S_j}$  aller Zustände des Zustandsübergangsdiagramms berechnet. Ein Zustand ist dann ein möglicher Nachfolger, wenn gilt:

$$\begin{aligned} t_{A_{\min}}^{S_j} &\leq t_{k+1} \leq t_{E_{\max}}^{S_j} \\ \wedge \quad t_{0_{\min}}^{S_j} &\leq t_{k+1} \leq t_{0_{\max}}^{S_j} + d_{\max}^{S_j} \end{aligned} \tag{3.5}$$

Der betrachtete Zeitpunkt  $t_{k+1}$  muß einerseits innerhalb des aus den Zustands- und Zustandsübergangsdauern berechneten Zeitintervalls liegen. Andererseits müssen auch die vordefinierten Zeiten eines möglichen Auftretens des Zustands eingehalten werden.

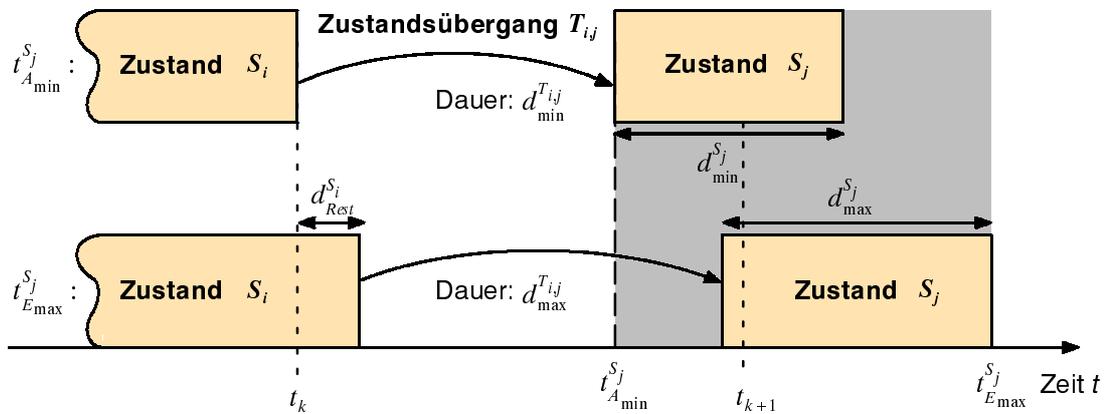


Bild 3.10: Berechnung des frühesten Anfangszeitpunkts und spätesten Endzeitpunkts eines möglichen Folgezustands

Diese sind a priori durch das Intervall der Anfangszeitpunkte  $[t_{0_{min}}^{S_j}, t_{0_{max}}^{S_j}]$  und die maximale Dauer eines jeden Zustands  $d_{max}^{S_i}$  festgelegt. Diese Vorgehensweise entspricht der geforderten optimistischen Abschätzung.

Die Berechnung des frühesten Anfangszeitpunkts  $t_{A_{min}}^{S_j}$  und des spätesten Endzeitpunkts  $t_{E_{max}}^{S_j}$  eines Zustands (Bild 3.10) erfolgt gemäß Gl. (3.6) bzw. Gl. (3.7). Der früheste Anfangszeitpunkt wird durch die minimale Zustandsübergangsdauer von  $S_i$  zu  $S_j$  bestimmt, wenn optimistischerweise vorausgesetzt wird, daß der Zustandsübergang sofort beginnt. Die Berechnung des spätesten Endzeitpunkts legt die maximale Übergangsdauer und die maximale Dauer von  $S_j$  zugrunde. Darüber hinaus wird die maximale restliche Verweildauer  $d_{Rest}^{S_i}$  im Ausgangszustand  $S_i$  berücksichtigt. Diese ergibt sich aus dessen maximaler Zustandsdauer abzüglich der bisher beobachteten Verweildauer in  $S_i$ . Wurde z. B. ein Zustand, der maximal 14 Tage anhalten kann, bereits seit 5 Tagen beobachtet, so beträgt die restliche Verweildauer noch maximal 9 Tage.

$$t_{A_{min}}^{S_j} = t_k + d_{min}^{T_{ij}} \quad (3.6)$$

$$t_{E_{max}}^{S_j} = t_k + d_{Rest}^{S_i} + d_{max}^{T_{ij}} + d_{max}^{S_j} \quad (3.7)$$

Besonders berücksichtigt werden muß die Möglichkeit, daß ein Zustand im Zustandsübergangsdiagramm aufgrund eines großen Zeitsprungs von  $t_k$  nach  $t_{k+1}$  übersprungen werden kann. In einem solchen Fall addieren sich die einzelnen Zeitangaben entlang des Pfads vom Ausgangs- zum Endknoten.

Hat man z. B. den Zustand *Messe inaktiv* in einem Bild erkannt, so ergeben sich folgende Zustände als mögliche Nachfolger in einem 9 Tage später aufgenommenen Bild (vgl. Wissensbasis aus Bild 3.9):

- *Messe inaktiv*: Der Zustand kann unverändert sein, da die Zustandsdauer maximal 365 Tage beträgt.
- *Messeaufbau*: Da eine Aufbauphase bis zu zehn Tage dauern kann und keine absoluten Zeitvorgaben für das erlaubte Auftreten des Zustands existieren, ist auch dieser Zustand denkbar.
- *Messe aktiv*: Falls die Aufbauphase weniger als acht Tage gedauert hat, kann inzwischen eine Messe begonnen haben. Der Zustand *Messeaufbau* wäre dann übersprungen worden.

Der Zustand *Messeabbau* läßt sich hingegen auf Basis der genannten Wissensbasis nicht innerhalb von neun Tagen erreichen, da selbst bei einer fünftägigen Messe und einer minimal kurzen Aufbauzeit mindestens zwölf Tage bis zum Abbau verstreichen würden. In diesem Beispiel müssen somit nur drei von vier möglichen Zuständen auf ihre Existenz untersucht werden. Der vierte Zustand kann von vornherein durch Ausnutzung des temporalen Wissens ausgeschlossen werden, wodurch der Interpretationsprozeß effizienter wird.

An dem Beispiel erkennt man außerdem, daß durch die Berücksichtigung der vorherigen Erfahrungen auch visuell nicht unterscheidbare Zustände differenziert werden können. In einem monotemporalen Luftbild können die Zustände *Messeaufbau* und *Messeabbau* nicht voneinander unterschieden werden, beide prägen sich gleichermaßen durch die neben den Hallen stehenden LKWs im Bild aus. Dabei kann man natürlich nicht erkennen, ob das Material aus- oder eingeladen wird. Erst durch die Einordnung in einen zeitlichen Kontext läßt sich in dem obigen Gedankenexperiment feststellen, daß es sich um eine Aufbau- und nicht um eine Abbauphase handeln könnte. Die Einbringung des temporalen Wissens und die entsprechende Nutzung bei der multitemporalen Bildanalyse ermöglichen somit eine verbesserte Szeneninterpretation mit spezialisierteren Objektbedeutungen.

### **3.3.3 Strategie für die Interpretation multitemporaler Luftbilder**

Das temporale Wissen läßt sich durch die beschriebene modellgetriebene Prädiktion von Objektzuständen für eine zeitliche Schlußfolgerung bei der Bildinterpretation einsetzen. Diese temporale Inferenz muß jedoch in die bisherige Auswertestrategie

eingebettet werden, ohne sie zu beeinträchtigen. Bislang besteht die Strategie in einer modellgetriebenen Generierung von Hypothesen, die anschließend in den Sensordaten, also datengetrieben, verifiziert oder auch falsifiziert werden. Desweiteren werden mit hoher Priorität Hypothesen entlang der topologischen Kanten im semantischen Netz erzeugt, um möglichst schnell die Nachbarschaftsbeziehungen auswerten zu können.

Die Annahme, daß alle Bilddaten die beobachtete Szene zum selben Zeitpunkt darstellen, gilt nicht mehr. Für die Unterscheidung der verschiedenen Aufnahmezeitpunkte wurden für die Knoten des semantischen Netzes Zeitstempel eingeführt. Zur Laufzeit der Analyse werden diese mit dem Zeitpunkt ihrer Instanzierung bzw. ihrer letzten Aktualisierung gefüllt. Dadurch wird eine zeitliche Zuordnung der Knoten möglich, was Voraussetzung für eine temporale Inferenz ist.

Die Ausnutzung des temporalen Wissens erfolgt mit Hilfe einer zusätzlichen Inferenzregel  $R_{Temp}$ . Sie wird aktiviert, sofern zum aktuellen Interpretationsstand noch ein weiteres auszuwertendes Bild vorliegt, und somit ein weiterer Zeitstempel existiert, und gleichzeitig für eine komplette Instanz Folgezustände über eine temporale Relation prädiert werden können. Für jeden der möglichen Folgezustände wird eine Hypothese generiert und der Suchbaum in entsprechend viele Blattknoten aufgespalten, da die einzelnen Alternativen konkurrierend zu betrachten sind. Die Analyse schaltet einen Zeitschritt weiter und setzt mit der Interpretation des Folgebildes fort:

**$R_{Temp}$ : Temporale Inferenz**

WENN für den Zeitpunkt  $t_{k+1}$  Bilddaten vorliegen und für eine Instanz  $I_K(n)$  mit dem Zeitstempel  $t_k$  noch keine Instanzen  $I(m)$  existieren mit:

$$n \xrightarrow{r} m \quad r = temp-rel ; n, m \in \mathcal{N}$$

$$\wedge \quad t[I(m)] = t_{k+1}$$

DANN prädiere mögliche Folgezustände  $\{m\}$  von  $I_K(n)$  für  $t_{k+1}$ , erzeuge für jeden Kandidaten aus  $\{m\}$  eine Hypothese wie folgt:

- erzeuge und verzeigere eine leere Instanz  $I_H(m)$
- setze den Zeitstempel auf  $t_{k+1}$
- kopiere und berechne die Attribute
- markiere  $I_H(m)$  als modifizierten Knoten

erzeuge und verzeigere für jeden Kandidaten aus  $\{m\}$  einen neuen Suchbaumknoten mit Zeitstempel  $t_{k+1}$  als Nachfolger von  $\mathcal{N}$ .

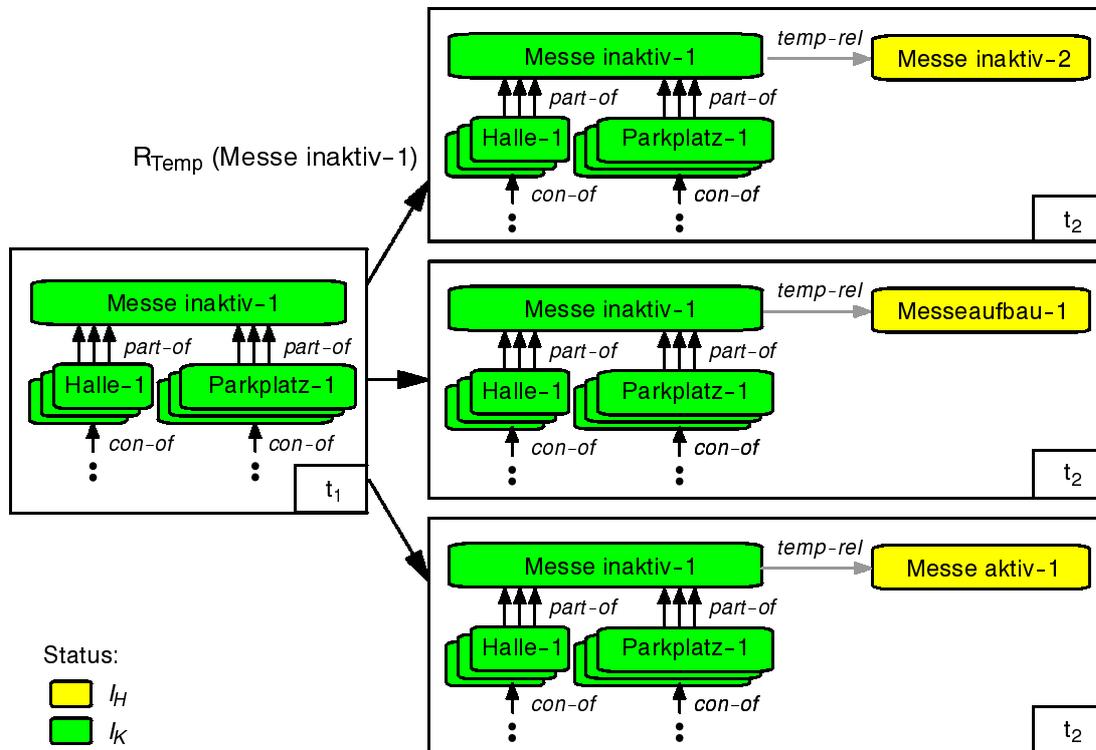


Bild 3.11: Prädiktion von Folgezuständen mit Hilfe der Regel für temporale Inferenz  $R_{Temp}$

Bild 3.11 illustriert die Anwendung der Regel  $R_{Temp}$  für das Beispiel des Messegeländes aus Bild 3.9. Wie zuvor erläutert, sind mögliche Folgezustände für ein inaktives Messegelände nach neun Tagen die Zustände *Messe inaktiv*, *Messeaufbau* und *Messe aktiv* (s. Kapitel 3.3.2). Die Regel zur temporalen Inferenz erzeugt die entsprechenden Hypothesen und spaltet den Suchbaum in drei neue Knoten auf. Die Systemsteuerung muß nun auf Basis einer Bewertung der einzelnen Alternativen entscheiden, mit welchem Suchbaumknoten fortgefahren werden soll. Kapitel 4 behandelt das Thema der Bewertung detailliert.

Die temporale Inferenz soll erst dann aktiviert werden, wenn das aktuelle Bild vollständig interpretiert worden ist, da sie das Weiterschalten zum nächsten Bild in der Zeitreihe bewirkt. Die Regel  $R_{Temp}$  erhält deshalb eine niedrige Priorität, so daß ihr Bedingungsteil erst dann geprüft wird, wenn keine der in Kapitel 2.4.1 beschriebenen Inferenzregeln aktiviert werden kann. Darüber hinaus wird nicht einfach die erste Regel, deren Bedingungsteil erfüllt ist, gefeuert, sondern alle aktivierbaren Regelkandidaten werden zunächst in einer Konfliktmenge gesammelt. Anschließend wird diejenige Regel aktiviert, die die Instanz mit dem frühesten Zeitstempel bearbeitet.

---

Kann z. B. die Regel  $R_{Temp}$  sowohl für eine Instanz vom Zeitpunkt  $t_k$  als auch für eine mit dem Zeitstempel  $t_{k+1}$  feuern, so wird die Regel für die Instanz aus  $t_k$  ausgeführt. Dadurch wird gewährleistet, daß nicht voreilig in der Zeitreihe weitergesprungen, sondern zuvor die Interpretation für den aktuellen Zeitpunkt vervollständigt wird.

### 3.3.3.1 Multitemporale Inferenzregeln

Die Instanzierung von Objekten erfolgt in den Regeln aus Kapitel 2.4.1 ohne Berücksichtigung der Zeitstempel von Objektteilen und -konkretisierungen. Demnach wird ein Objekt auch dann instanziiert, wenn seine Objektteile zu unterschiedlichen Zeiten beobachtet wurden. Bei übergeordneten Konzepten, wie dem Messegelände aus Bild 3.9, ist dieses auch erwünscht. Man kann die Existenz eines Messegeländes erst dann bestätigen, wenn zu unterschiedlichen Zeitpunkten die einzelnen Zustände eines Messegeländes detektiert worden sind. Hier wird das komplexe Objekt an seinem zeitlichen Verhalten erkannt.

Anders verhält es sich bei der Detektion von Landnutzungsänderungen. Hier soll derselben (geographischen) Region, z. B. einem bestimmten Flurstück, zu jedem einzelnen Beobachtungszeitpunkt eine Bedeutung wie landwirtschaftliche Fläche, Siedlung oder Wald zugeordnet werden. Das bedeutet, daß ein im semantischen Netz repräsentiertes Konzept *Flurstück* in jedem Zeitschritt seine Konkretisierungen oder Bestandteile für die Instanzierung benötigt. Alle Kindknoten müssen somit denselben Zeitstempel tragen, bevor der übergeordnete Knoten instanziiert werden darf. Um ein derartiges Verhalten zu erreichen, werden die bisherigen Instanzierungsregeln durch solche ersetzt, die den Zeitstempel berücksichtigen. Beispielsweise wird die Regel zur multitemporalen partiellen Instanzierung  $R_{PMT}$  (s. Kasten) erst dann aktiviert, wenn alle obligatorischen Kindknoten vorhanden sind und denselben Zeitstempel tragen wie der Vaterknoten. Während der Instanzierung wird der aktuelle Zeitstempel in die neue Instanz eingetragen. Analog werden die übrigen Regeln zur Instanzierung, zur Datenbindung und zur modellgetriebenen Hypothesengenerierung durch ihre jeweilige multitemporale Version ersetzt.

Wann immer ein Bild vollständig interpretiert wurde und die Regel zur temporalen Inferenz für ein Weiterschalten zum nächsten Bild gesorgt hat, wird überprüft, ob die vorhandenen kompletten Instanzen der symbolischen Ebene auf Basis des neuen Zeitstempels noch immer den richtigen Status besitzen. Gegebenenfalls wird des Status von *komplette Instanz* auf *Hypothese* zurückgesetzt, um zu ermöglichen, daß neue Kindknoten mit dem nunmehr aktuellen Zeitstempel gebunden werden können. Dadurch wird beispielsweise ermöglicht, daß eine Instanz vom Typ *Flurstück* in jedem

**R<sub>PMT</sub>: Multitemporale partielle Instanzierung**

WENN für eine Hypothese  $I_H(n)$  mit dem Zeitstempel  $t_k$  alle kompletten Instanzen  $I_K(m)$  existieren mit:

$$m \xrightarrow{r_{obl}} n \quad r \in \{part-of, con-of, data-of, attr-rel\} ; n, m \in \mathcal{N}$$

$$\wedge \quad t[I_K(m)] = t_k$$

DANN erzeuge den Interpretationszustand  $I_P(n)$  wie folgt:

falls  $I_H(n)$  bereits im Suchbaumknoten  $\mathcal{N}$  modifiziert wurde,

- ändere Zustand zu  $I_P(n)$
- berechne die Attribute

sonst

- erzeuge und verzeigere eine leere Instanz  $I_P(n)$
- setze den Zeitstempel auf  $t_k$
- kopiere und berechne die Attribute

markiere  $I_P(n)$  als modifizierten Knoten.

Bild des multitemporalen Datensatzes der jeweils aktuellen Objektklasse zugeordnet wird.

Die beschriebenen Inferenzregeln werden zu einer neuen Strategie zusammengefaßt, die nunmehr in der Lage ist, multitemporale Bilddatensätze zu verwalten, Folgezustände auf Basis des Zustandsübergangsdiagramms zu präzisieren und diese im jeweils aktuellen Bild zu überprüfen. Die Bilder werden in ihrer chronologischen Reihenfolge jeweils vollständig interpretiert, bevor zum nächsten Bild übergegangen wird.

### 3.3.4 Zusammenfassung der temporalen Erweiterungen

Folgende Maßnahmen ermöglichen die angestrebte multitemporale Bildanalyse:

- Jede Instanz des semantischen Netzes erhält einen Zeitstempel, damit sie einem der multitemporalen Luftbilder zugeordnet werden kann.
- Vorwissen über temporale Zusammenhänge wird durch erweiterte Zustandsübergangsdiagramme beschrieben. Die Zustände werden durch Konzeptknoten, die Zustandsübergänge durch die neu eingeführte temporale Relation direkt im semantischen Netz repräsentiert. Es können Zustands- und Zustandsübergangsdauer sowie deren Auftretenswahrscheinlichkeit definiert werden. Darüber hinaus ist es möglich, einen absoluten Anfangszeitpunkt für einen

---

Zustand anzugeben. Der Unsicherheit aller Zeitangaben wird durch die Verwendung von Minimal- und Maximalwerten Rechnung getragen.

- Während der Analyse wird das temporale Wissen dazu genutzt, mögliche Folgezustände für die Szenenobjekte zu präzisieren. Mit Hilfe der definierten absoluten und relativen Zeitangaben werden Hypothesen für alle Zustandsänderungen generiert, die in der zwischen zwei aufeinanderfolgenden Bildern verstrichenen Zeit eingetreten sein können. Anschließend werden alle konkurrierenden Hypothesen in den Daten verifiziert. Durch Ausschluß unmöglicher Zustandsänderungen kann die Analyse effizienter gestaltet werden. Das beschriebene Verhalten wird durch den Einsatz neuer Inferenzregeln zur temporalen Schlußfolgerung erzielt.

Somit sind die in den Kapiteln 3.1.1 und 3.1.2 genannten Anforderungen an ein multitemporales Bildinterpretationssystem weitestgehend erfüllt. Alle repräsentierten zeitlichen Angaben des temporalen Wissens werden während der Analyse ausgewertet. Lediglich die Auftretenswahrscheinlichkeiten der Zustands- und Zustandsübergänge werden bislang nicht genutzt. Diese Wissensinhalte können dazu eingesetzt werden, um naheliegende von abwegigen Prognosen zu unterscheiden und die ersteren bevorzugt zu behandeln. Hier wird ein Bewertungssystem benötigt, das die a-priori-Wahrscheinlichkeiten bei der Beurteilung einer Szenenbeschreibung mit berücksichtigt. Das folgende Kapitel widmet sich diesem Thema.

## 4. Ein probabilistisches Bewertungssystem

Wie bereits in Kapitel 2.4 deutlich wurde, kommt es während der Bildanalyse in verschiedenartigen Situationen zu Mehrdeutigkeiten. Die Segmentierungsergebnisse sind unsicher und lassen unterschiedliche Interpretationen zu. Aber auch bei Schlußfolgerungen, die ausschließlich aus dem Vorwissen gewonnen werden, können Mehrdeutigkeiten auftreten: Oftmals sind mehrere Spezialisierungen für ein detektiertes Objekt denkbar oder bei der temporalen Inferenz lassen sich mehrere Folgezustände für ein Objekt präzisieren.

In derartigen Situationen wird, wie bereits erläutert, jede Alternative in einem separaten Knoten eines Suchbaums abgelegt. Jeder Suchbaumknoten repräsentiert eine in sich schlüssige Interpretation der Szene. Um die Suche nach der optimalen Interpretation zu beschleunigen, werden die einzelnen Alternativen getrennt voneinander bezüglich der Kompatibilität zwischen den Messungen in den Daten und den aus der Wissensbasis und den Zwischenergebnissen abgeleiteten Erwartungen bewertet. In Kapitel 2.4.3 wurde ein Verfahren vorgestellt, das basierend auf der Possibility-Theorie aus den Attributen und Knoten des semantischen Netzes ein Bewertungsmaß für einen Suchbaumknoten bestimmt. Hierbei werden die Einzelbewertungen sukzessive von den Blattknoten bis hoch zu den Wurzelknoten aggregiert, bis schließlich eine Gesamtbewertung für das semantische Netz des Suchbaumknotens ermittelt ist.

Das Verfahren berücksichtigt ausschließlich solche Knoten und Attribute, bei denen die Erwartungen bereits durch Messungen in den Daten verifiziert wurden. Liegen noch keine Messungen vor, wird eine perfekte Bewertung angenommen, um eine optimistische Qualitätsabschätzung (die Konvergenzbedingung des A\*-Algorithmus, s. Kapitel 2.4.2) zu gewährleisten. Eine derartige Vorgehensweise bewirkt einen ungewollten Effekt: Alle Suchbaumknoten, die sich lediglich durch eine neu hinzugekommene Hypothese unterscheiden, erhalten identische Bewertungen. Denn die unverifizierten Hypothesen werden alle gleichermaßen optimistisch abgeschätzt, während der Rest des semantischen Netzes für alle Alternativen gleich ist. Dieses Phänomen tritt im Beispiel aus Bild 3.11 auf: Die drei konkurrierenden Interpretationen unterscheiden sich lediglich durch die prädierten Folgezustände *Messe inaktiv*, *Messeaufbau* bzw. *Messe aktiv*. Durch die Erzeugung der Hypothesen ist keine neue Evidenz in das semantische Netz geflossen. Aus diesem Grund sind nur die restlichen Knoten des semantischen Netzes bewertungsrelevant. Da sie in allen drei Interpretationen identisch sind, werden die Alternativen gleich bewertet. Das System hat demnach keinen Anhaltspunkt, mit welcher Interpretation es fortfahren soll, und wählt deshalb zufällig einen der drei Suchbaumknoten für die weitere Analyse aus.

---

Durch die Zufälligkeit dieser Entscheidung ist es möglich, daß das System zunächst viel Zeit benötigt, um eine sehr unwahrscheinliche Lösung zu widerlegen, bis es schließlich die naheliegende Lösung untersucht. Zur Steigerung der Effizienz ist es deshalb sinnvoll, in Situationen, in denen nicht aufgrund von Hinweisen aus den Daten entschieden werden kann, Erfahrungen und a-priori-Wissen zu nutzen, um die wahrscheinlichere Alternative zuerst zu verfolgen. In all den Fällen, in denen nur aufgrund von Erfahrungen entschieden werden kann, ist diese Heuristik eine plausible Strategie. Voraussetzung dafür ist natürlich, daß entsprechende Erfahrungen z. B. in Form von Auftretenswahrscheinlichkeiten für die einzelnen Ereignisse bekannt sind. Im erwähnten Beispiel stellt die Lösung mit dem Folgezustand *Messe inaktiv* die wahrscheinlichste Alternative dar (vgl. Wissensbasis in Bild 3.9), da es der am häufigsten beobachtete Zustand eines Messengeländes ist. Somit sollte das Bildinterpretationssystem diese Alternative als erste untersuchen.

Um das gewünschte Systemverhalten zu erreichen, muß das Bewertungsverfahren dafür sorgen, daß wahrscheinlichere Lösungen bei ansonsten identischer Evidenz besser bewertet und somit bevorzugt werden. Es muß sowohl die Evidenz aus den Daten berücksichtigen, als auch das in Form von Auftretenswahrscheinlichkeiten definierte Vorwissen mit in den Bewertungsprozeß einbeziehen. Die beiden Arten von Informationen sind gewissenhaft voneinander zu unterscheiden. Vorwissen kann zwar hilfreich sein, den Suchraum einzuschränken und abwegige Hypothesen zurückzustellen. Letztlich entscheiden aber die Beobachtungen in den Daten über das Vorhandensein oder das Fehlen von Objekten.

Erwünscht ist demnach ein Bewertungssystem, das sowohl einen Informationsfluß *bottom-up* von den beobachteten Daten zu ihrer Interpretation als auch *top-down* von einer Hypothese zu den erwarteten Beobachtungen gewährleistet. Bei der modellgetriebenen Vorgehensweise soll insbesondere Vorwissen über Auftretenswahrscheinlichkeiten von Objekten, wie es im Falle des temporalen Wissens definiert wird, mit berücksichtigt werden.

Die sog. *Bayes-Netze* bieten die genannte Eigenschaft. Sie modellieren Ereignisse und ihre Abhängigkeiten in einem Graphen und sind den semantischen Netzen deshalb strukturell sehr ähnlich. Bayes-Netze gestatten ein Schlußfolgern unter Berücksichtigung der Unsicherheit der Daten und des Vorwissens. Die Unsicherheit wird durch Wahrscheinlichkeiten repräsentiert. Deshalb spricht man bei Bayes-Netzen auch von probabilistischen Netzwerken (engl.: *probabilistic (belief) networks*). In Kapitel 4.1 wird die Theorie der Bayes-Netze in einem für das Verständnis der hier vorgestellten Anwendung ausreichenden Maße erläutert.

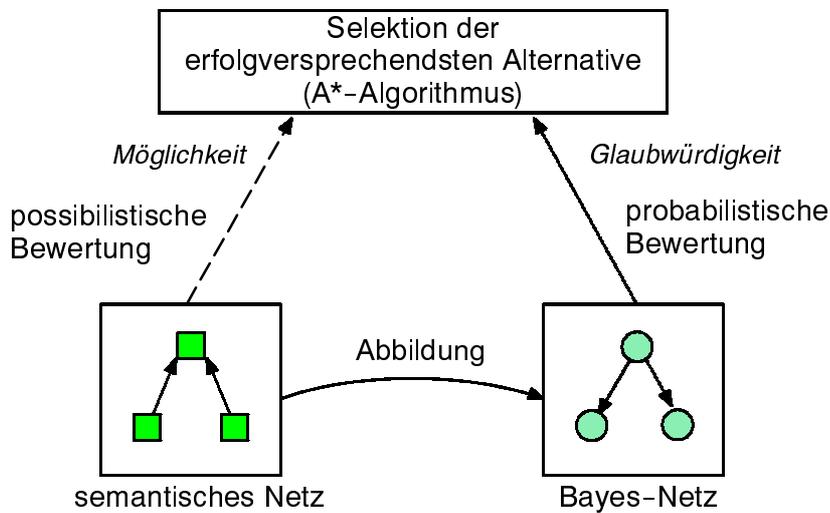


Bild 4.1: *Bewertung des semantischen Netzes durch Abbildung auf ein Bayes-Netz und Ableitung einer Glaubwürdigkeit für die Szenenbeschreibung*

Die zu bewertende Szenenbeschreibung liegt als instanziiertes semantisches Netz vor, das bei dem possibilistischen Bewertungsverfahren aus Kapitel 2.4.3 direkt ausgewertet wurde, um ein Qualitätsmaß für den übergeordneten A\*-Algorithmus zu ermitteln (Bild 4.1). Dabei wird die *Möglichkeit*  $Po(\mathcal{N})$  des Suchbaumknotens als Entscheidungskriterium herangezogen. Für das hier vorgeschlagene probabilistische Bewertungssystem wird das semantische Netz zunächst auf ein Bayes-Netz abgebildet (Bild 4.1). Mit Hilfe eines noch zu erläuternden Verfahrens wird aus den im Bayes-Netz enthaltenen Informationen ein *Glaubwürdigkeitswert* für die im Bayes-Netz repräsentierte Szenenbeschreibung abgeleitet. Dabei werden Auftretenswahrscheinlichkeiten von Knoten und Kanten, also insbesondere die Zustands- und Zustandsübergangswahrscheinlichkeiten des temporalen Wissens, mit berücksichtigt. Dadurch erhalten wahrscheinlichere Szenenbeschreibungen bei ansonsten identischer Evidenz eine bessere Bewertung.

Nach einer Einführung in die Theorie der Bayes-Netze (s. Kapitel 4.1) erörtert Kapitel 4.2, wie das semantische Netz auf ein Bayes-Netz abgebildet und auf welche Weise daraus die Bewertung der Szenenbeschreibung bestimmt wird. Schließlich wird das vorgeschlagene probabilistische Bewertungssystem mit dem possibilistischen Ansatz aus Kapitel 2.4.3 verglichen.

## 4.1 Theorie der Bayes-Netze

In einem Bayes-Netz [66][100][107] werden kausale Abhängigkeiten von Ereignissen durch einen Graphen explizit repräsentiert. Jedes Ereignis wird als diskrete Zufallsvariable mit endlich vielen, sich gegenseitig ausschließenden Werten modelliert und als Knoten des Bayes-Netzes dargestellt. Die Kanten, die die Knoten verbinden, modellieren die kausalen Abhängigkeiten zwischen den Zufallsvariablen. Die Richtung der Kanten entsprechen der natürlichen Vorstellung der kausalen Abhängigkeit und sind von der Ursache zur Wirkung gerichtet. Die Stärke der kausalen Beeinflussung wird durch bedingte Wahrscheinlichkeiten quantifiziert. Somit werden die Abhängigkeiten durch die Struktur des Bayes-Netzes qualitativ und durch die bedingten Wahrscheinlichkeiten quantitativ beschrieben.

Neben der reinen Repräsentation von kausalen Abhängigkeiten können Bayes-Netze auch zur Entscheidungsunterstützung genutzt werden, indem für die Zufallsvariablen Glaubwürdigkeitswerte (engl.: *belief*) auf Basis von Beobachtungen in einem konkreten Anwendungsfall berechnet werden. Die im Bayes-Netz repräsentierten a-priori-Wahrscheinlichkeiten für bestimmte Ereignisse werden durch mehr oder weniger gesicherte Beobachtungen ersetzt. Diese sog. *Evidenz* beeinflusst nicht nur die Glaubwürdigkeit des beobachteten Ereignisses, sondern indirekt auch die Glaubwürdigkeiten der davon kausal abhängigen Ereignisse, d. h. der benachbarten Knoten im Bayes-Netz. Die Evidenz wird nach festen Regeln im Bayes-Netz verteilt (s. Kapitel 4.1.3), bis alle Glaubwürdigkeiten aktualisiert sind. Aus diesen lassen sich schließlich Entscheidungen oder Klassifikationen ableiten.

Ein Anwendungsfeld der Bayes-Netze ist z. B. die medizinische Diagnostik. Hierbei werden die medizinischen Zusammenhänge, welche Symptome und Risikofaktoren in welchem Maß auf bestimmte Krankheiten hindeuten, in Form eines Bayes-Netzes formuliert (Bild 4.2). Werden bei einem Patienten bestimmte Symptome oder Risikofaktoren beobachtet, so verändern sich die jeweiligen Glaubwürdigkeiten, daß er an einer der Krankheiten leidet. Auf diese Weise lassen sich Hinweise für eine medizinische Diagnose und eine anschließende Therapie ableiten. Beispiele für medizinische Diagnosesysteme auf Basis von Bayes-Netzen wurden sowohl im Bereich der Humanmedizin [8][64][87][164] als auch in der Veterinärmedizin [116][156] vorgestellt.

Im Bereich der Bildinterpretation werden Bayes-Netze u. a. bei der 2D-Bildanalyse [65][69][82][112][121][126][127][143] sowie zur Erkennung von 3D-Objekten wie Volumenmodelle [77][99], Fahrzeuge [13], Gebäude [80][81] oder Personen [117] eingesetzt. Dynamische Bayes-Netze, die ihre Struktur während der Analyse verän-

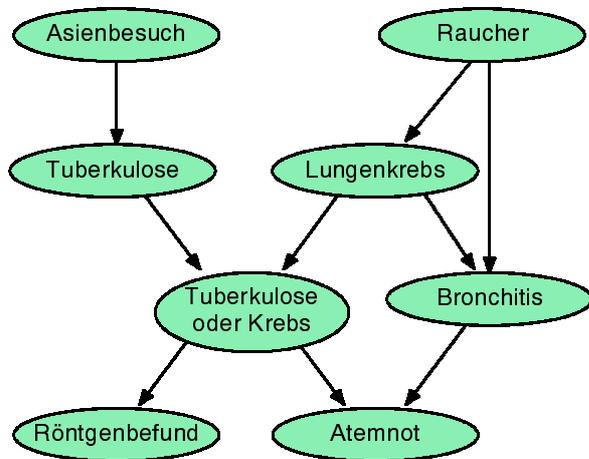


Bild 4.2:

*Modellierung von medizinischem Wissen durch ein Bayes-Netz (aus [87]): Atemnot kann u. a. durch Tuberkulose, Lungenkrebs oder Bronchitis verursacht werden. Ein Besuch in Asien erhöht seinerseits das Tuberkuloserisiko, während Rauchen ein Risikofaktor für Lungenkrebs und Bronchitis darstellt. In einer Thorax-Röntgenaufnahme kann man Tuberkulose und Lungenkrebs nicht differenzieren und eine Bronchitis gar nicht erkennen.*

dem, werden z. B. bei der Überwachung von Verkehrsszenen aus Bildsequenzen genutzt [16][62].

Die Erstellung eines Bayes-Netzes ist keineswegs trivial, da die Abhängigkeiten und Wahrscheinlichkeiten in vielen Fällen schwer anzugeben sind. Sie können sowohl aus Expertenwissen stammen oder aus statistischen oder empirischen Untersuchungen abgeleitet werden. Darüber hinaus existieren Ansätze, Struktur und Wahrscheinlichkeiten von Bayes-Netzen automatisch aus Daten zu erlernen [19][55][56][78]. In der vorliegenden Anwendung ist die Struktur des Bayes-Netzes weitestgehend durch das zu bewertende semantische Netz vorgegeben.

#### 4.1.1 Begriffe und Definitionen

Zunächst werden einige Begriffe eingeführt, die im weiteren Verlauf der Arbeit benötigt werden:

##### 4.1.1.1 Zufallsvariablen und bedingte Unabhängigkeit

Mit  $U$  wird eine endliche Menge von diskreten Zufallsvariablen bezeichnet. Die Namen der Zufallsvariablen werden durch Großbuchstaben (z. B.  $X, Y, Z$ ), deren Belegung mit einem konkreten Variablenwert durch Kleinbuchstaben (z. B.  $x, y, z$ ) gekennzeichnet. Fettgedruckte Buchstaben stehen für Mengen von Variablen (z. B.  $X, Y, Z$ ) bzw. für die Wertebelegung oder die *Konfiguration* dieser Variablenmenge (z. B.  $x, y, z$ ).

Jede Zufallsvariable  $X \in U$  kann disjunkte Werte  $x$  aus einem endlichen Bereich  $D_X$  annehmen. Die Wahrscheinlichkeit, daß  $X$  den Wert  $x$  annimmt, ist:

$$P(X = x) = P(x), \quad x \in D_x$$

**Definition:** Sei  $U$  eine Menge von diskreten Zufallsvariablen.  $P(\cdot)$  sei die Verbundwahrscheinlichkeitsverteilung der Variablen in  $U$ .  $X$ ,  $Y$  und  $Z$  seien drei Untermengen der Variablen in  $U$ .  $X$  und  $Y$  werden als **bedingt unabhängig gegeben  $Z$**  bezeichnet (engl.: *conditionally independent given  $Z$* ), wenn gilt:

$$P(x|y,z) = P(x|z) \quad \text{für } P(y,z) > 0 \quad (4.1)$$

Im Falle bedingter Unabhängigkeit gilt außerdem:

$$P(x,y|z) = P(x|z) P(y|z) \quad (4.2)$$

$$P(x,y,z) = P(x|z) P(y,z) \quad (4.3)$$

#### 4.1.1.2 Definition eines Bayes-Netzes

Ein *Bayes-Netz* ist definiert durch:

- Eine Menge von *Bayes-Knoten* und *Bayes-Kanten*.
- Jeder Knoten repräsentiert eine diskrete Zufallsvariable  $X \in U$  mit ihren  $n$  disjunkten Werten aus  $D_X$ .
- Bayes-Knoten und -Kanten bilden einen gerichteten, azyklischen Graphen (engl.: *directed acyclic graph* oder kurz: DAG). Zyklen über rückgekoppelte Kanten sind nicht erlaubt (s. Bild 4.3a).
- Jedem Knoten  $X$  mit übergeordneten Knoten  $U^1, \dots, U^N$  ist eine Matrix von bedingten Wahrscheinlichkeiten  $\underline{M}_{x|u} = P(x|u) = P(x|u^1, \dots, u^N)$  zugeordnet, die die kausalen Abhängigkeiten zwischen Eltern- und Kindknoten modellieren. Hat der Knoten  $X$  keine Eltern, so entartet die genannte Matrix zum Vektor  $P(x) = P(X=x_1, \dots, X=x_n)$ .

Jede Zufallsvariable nimmt mit einer gewissen Wahrscheinlichkeit einen ihrer Werte  $x_i$  an. Liegt völlige Unwissenheit vor, so sind die Wahrscheinlichkeiten über den Wertebereich gleichverteilt. Beispielsweise läßt sich ein Würfel durch eine Zufallsvariable bzw. einen Bayes-Knoten  $X$  mit sechs möglichen Werten modellieren. Er erhält eine a-priori-Wahrscheinlichkeitsverteilung von:

$$P(x) = P(X='1', X='2', X='3', X='4', X='5', X='6') = \left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right).$$

#### 4.1.1.3 Evidenz

Wird in den Sensordaten ein bestimmter Wert für eine der Zufallsvariablen beobachtet, so verändert sich die Wahrscheinlichkeitsverteilung des zugehörigen Knotens. Die

Wahrscheinlichkeit des beobachteten Wertes steigt, während die der restlichen Werte sinken. Derartige Beobachtungen werden *Evidenz*  $e$  genannt. Wurde z. B. mit absoluter Sicherheit beobachtet, daß mit dem obigen Würfel eine "2" gewürfelt wurde, so ergibt sich eine aktualisierte Wahrscheinlichkeitsverteilung des Knotens  $X$  unter Beobachtung der Evidenz  $e$  von:

$$P(x|e) = (0, 1, 0, 0, 0, 0).$$

Die Evidenz ist unsicher, wenn die Korrektheit der Beobachtung nicht zugesichert werden kann. Läßt sich im obigen Beispiel nicht sicher entscheiden, ob eine "2" oder "3" gewürfelt wurde, da die Augen des Würfels schlecht erkennbar sind, so ergibt sich z. B. eine Wahrscheinlichkeitsverteilung von:

$$P(x|e) = (0, 0.8, 0.2, 0, 0, 0).$$

#### 4.1.1.4 Glaubwürdigkeit

Die Glaubwürdigkeit (engl.: *Belief*)  $BEL(x)$  ist der totale Glaube an die Aussage  $\{X = x\}$  bei gegebener Evidenz  $e$ :

$$BEL(x) \hat{=} P(x|e) \quad (4.4)$$

Die Glaubwürdigkeit entspricht der aktualisierten Wahrscheinlichkeitsverteilung eines Bayes-Knoten unter Beobachtung der aktuell vorhandenen Evidenz. Sie ist zu unterscheiden von der a-priori-Wahrscheinlichkeitsverteilung  $P(x|u)$ .

#### 4.1.1.5 Bayes-Regel

Die Grundlage aller Bayes-Techniken (Bayes-Klassifikatoren, Bayes-Netze etc.) ist die Bayes-Regel:

$$P(H|e) = \frac{P(e|H) P(H)}{P(e)} \quad (4.5)$$

Sie besagt, daß die Glaubwürdigkeit einer Hypothese  $H$  unter Beobachtung der Evidenz  $e$  aus der a-priori-Wahrscheinlichkeit  $P(H)$ , dem Likelihood-Wert  $P(e|H)$  und der Wahrscheinlichkeit für die Beobachtung  $P(e)$  berechnet werden kann. Letztere dient zur Normierung und kann aus Gl. (4.6) abgeleitet werden. Dabei wird zugrunde gelegt, daß sich die Wahrscheinlichkeiten  $P(H|e)$  und  $P(\neg H|e)$  zu 1 addieren.

$$P(e) = P(e|H)P(H) + P(e|\neg H)P(\neg H) \quad (4.6)$$

Die Bayes-Regel ist deshalb von großer Bedeutung, da sich mit ihrer Hilfe die posteriori-Wahrscheinlichkeit  $P(H|e)$ , die in der Regel nur schwer quantifizierbar ist, aus Termen berechnet werden kann, die sich aus Vorwissen ableiten lassen.

**Beispiel:** Am Nachbartisch im Spielcasino wird das Resultat “Zwölf” ausgerufen, ohne daß man weiß, ob es sich um Roulette oder um ein Würfelspiel mit zwei Würfeln handelt. Die Wahrscheinlichkeit  $P(H|e) = P(\text{Würfel}|\text{Zwölf})$  ist ohne Detailkenntnisse des Casinos nur schwer anzugeben. Hingegen kann man durch Anwendung der Bayes-Regel die gesuchte Wahrscheinlichkeit indirekt berechnen:

Aus allgemeinem Vorwissen kann der Wahrscheinlichkeit  $P(e|H) = P(\text{Zwölf}|\text{Würfel})$  der Wert  $1/36$  zugeordnet werden.  $P(H) = P(\text{Würfel})$  läßt sich aus dem Anteil der Würfeltische an der Gesamtzahl der Spieltische schätzen und  $P(e) = P(\text{Zwölf})$  wird gemäß Gl. (4.6) bestimmt.

### 4.1.2 Bayes-Netze als Wissensbasis

Mit Hilfe eines Bayes-Netzes läßt sich die Verbundwahrscheinlichkeitsverteilung  $P(x^1, \dots, x^N)$  einer Menge von Zufallsvariablen  $\{X^1, \dots, X^N\} \in U$  darstellen, indem sie durch Anwendung der Kettenregel auf bedingte Wahrscheinlichkeiten zurückgeführt wird:

$$P(x^1, \dots, x^N) = P(x^N|x^{N-1}, \dots, x^1) P(x^{N-1}|x^{N-2}, \dots, x^1) \dots P(x^3|x^2, x^1) P(x^2|x^1) P(x^1) \quad (4.7)$$

Gilt für einige Zufallsvariablen bedingte Unabhängigkeit, so vereinfachen sich die Terme der bedingten Wahrscheinlichkeiten gemäß Gl. (4.1). Die bedingten Wahrscheinlichkeiten der kausal abhängigen Variablen werden im Bayes-Netz modelliert. Das in Bild 4.3b dargestellte Bayes-Netz aus [107] repräsentiert beispielsweise die folgende Verbundwahrscheinlichkeitsverteilung:

$$P(x^1, x^2, x^3, x^4, x^5, x^6) = P(x^6|x^5) P(x^5|x^2, x^3) P(x^4|x^1, x^2) P(x^3|x^1) P(x^2|x^1) P(x^1) \quad (4.8)$$

Die Abhängigkeiten eines Knoten  $X$  von seinen Elternknoten  $U^1, \dots, U^N$  (Bild 4.3c) werden durch die Matrix  $\underline{M}_{x|u} = \underline{M}_{x|u^1, \dots, u^N}$  gemäß Gl. (4.9) beschrieben. Dabei sei angenommen, daß der Knoten  $X$  eine  $n$ -wertige und - zur Vereinfachung - die Knoten  $U^i$  jeweils  $m$ -wertige diskrete Zufallsvariablen sind. Die Matrix  $\underline{M}_{x|u}$  enthält für jede Wertekombination von  $X$  und  $U^i$  eine Wahrscheinlichkeit, die den Einfluß einer konkreten Wertebelegung der  $U^i$  auf den jeweiligen Wert von  $X$  quantifiziert. Die Matrix besteht somit aus  $n$  Spalten und  $m^N$  Zeilen. Entlang einer Zeile summieren sich die Wahrscheinlichkeiten zu 1.

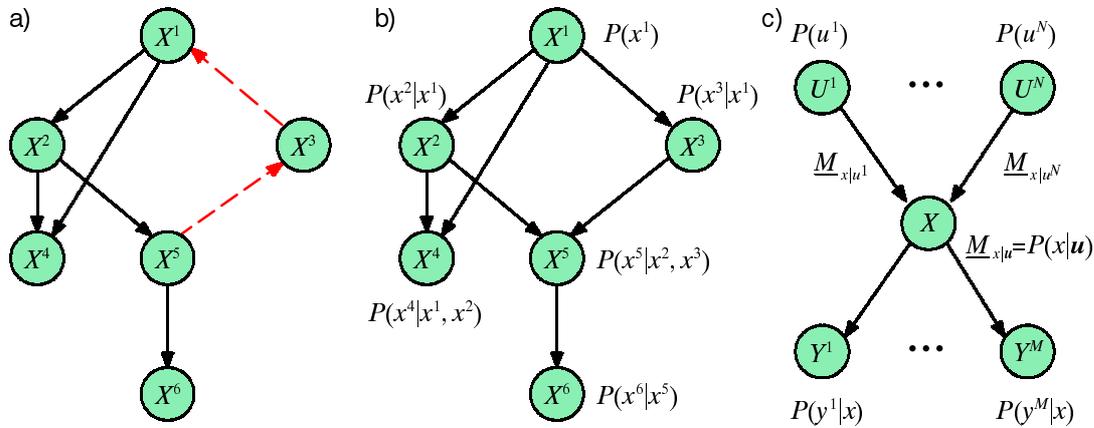


Bild 4.3: a) Gerichtetes Graph mit einem unerlaubten Zyklus. b) Bayes-Netz für eine Wahrscheinlichkeitsverteilung von sechs Zufallsvariablen. c) Bayes-Knoten mit  $N$  Eltern- und  $M$  Kindknoten.

$$\underline{M}_{x|u} = P(x|u) = \begin{matrix} \xrightarrow{n \text{ Spalten}} \\ \left[ \begin{array}{ccc} P(x_1|u_1^1, u_1^2, \dots, u_1^N) & \dots & P(x_n|u_1^1, u_1^2, \dots, u_1^N) \\ P(x_1|u_2^1, u_2^2, \dots, u_2^N) & \dots & P(x_n|u_2^1, u_2^2, \dots, u_2^N) \\ P(x_1|u_3^1, u_3^2, \dots, u_3^N) & \dots & P(x_n|u_3^1, u_3^2, \dots, u_3^N) \\ \vdots & & \vdots \\ P(x_1|u_m^1, u_m^2, \dots, u_m^N) & \dots & P(x_n|u_m^1, u_m^2, \dots, u_m^N) \end{array} \right] \downarrow m^N \text{ Zeilen} \end{matrix} \quad (4.9)$$

Die bedingten Wahrscheinlichkeiten aus Gl. (4.9) berücksichtigen die wechselseitigen Abhängigkeiten der einzelnen Vaterknoten  $U^i$ . In vielen Fällen sind jedoch diese gegenseitigen Beeinflussungen vernachlässigbar, so daß sich die Matrix aus Gl. (4.9) gemäß Gl. (4.10) näherungsweise aus dem Produkt der einzelnen Matrizen  $\underline{M}_{x|u^i}$  aus Gl. (4.11) bestimmen läßt. Der Vektor  $\alpha$  normiert die Matrix zeilenweise zu 1.

$$\underline{M}_{x|u} = \underline{M}_{x|u^1, \dots, u^N} \approx \alpha \cdot \prod_{i=1}^N \underline{M}_{x|u^i} \quad (4.10)$$

$$\underline{M}_{x|u^i} = P(x|u^i) = \begin{bmatrix} P(x_1|u_1^i) & P(x_2|u_1^i) & \dots & P(x_n|u_1^i) \\ P(x_1|u_2^i) & P(x_2|u_2^i) & \dots & P(x_n|u_2^i) \\ \vdots & \vdots & & \vdots \\ P(x_1|u_m^i) & P(x_2|u_m^i) & \dots & P(x_n|u_m^i) \end{bmatrix} \quad (4.11)$$

Die Matrix  $\underline{M}_{x|u^i}$  beschreibt nur die direkte Abhängigkeit des Knoten  $X$  vom Knoten  $U^i$ . Die Matrix kann demnach direkt der Kante von  $U^i$  nach  $X$  zugeordnet werden und wird deshalb auch *Kantenmatrix* genannt.

In der Praxis kann die vollständige Matrix aus Gl. (4.9) nur schwer bestimmt werden. Zudem ändert sie sich, sobald ein Knoten  $U^i$  hinzugefügt oder entfernt wird. Die Kantenmatrizen hingegen sind modular und voneinander unabhängig. Sie bieten deshalb Vorteile beim praktischen Einsatz. Der Fehler durch die Vernachlässigung der wechselseitigen Einflüsse der Elternknoten ist in vielen Fällen tolerierbar.

### 4.1.3 Propagierung von Evidenz in Bayes-Netzen

Nachdem das Modellwissen in einem Bayes-Netz repräsentiert wurde, wird es anschließend genutzt, um Schlußfolgerungen bei der Auswertung von Sensordaten zu ziehen. Dazu werden Beobachtungen als Evidenz in das Netz eingebracht und darin verteilt. Die aktualisierten Glaubwürdigkeiten der Knoten werden schließlich zur Entscheidungsunterstützung eingesetzt.

Der Vorgang der Evidenzverbreitung wird *Propagierung* oder auch *Belief-Update* genannt. In der Literatur wurden unterschiedliche Propagierungsalgorithmen für Bayes-Netze vorgeschlagen [66][100][107]. An dieser Stelle wird das Verfahren für einfach verbundene Netze von Pearl [107] vorgestellt.

Zunächst wird vom einfachsten Fall ausgegangen, daß das Bayes-Netz nur aus zwei Knoten und einer Kante,  $X \rightarrow Y$  besteht. Bei Beobachtung der Evidenz  $e = \{Y = y\}$  ergibt sich die Glaubwürdigkeit des Wurzelknotens  $X$  mit Gl. (4.4) und Gl. (4.5) zu:

$$BEL(x) = P(x|e) = \frac{P(e|x)P(x)}{P(e)} = \alpha P(x) \lambda(x), \quad (4.12)$$

mit dem sog. Likelihood-Vektor  $\lambda(x) = P(e|x) = P(Y=y|x)$  und dem Normierungsvektor  $\alpha = P(e)^{-1}$ .  $\lambda(x)$  ist demnach diejenige Zeile der Matrix  $\underline{M}_{y|x}$ , die für die Behauptung  $\{Y = y\}$  steht. Sie kann im Knoten  $Y$  berechnet und als Nachricht an den Vaterknoten  $X$  geschickt werden, der dann seinerseits seine Glaubwürdigkeit  $BEL(x)$  aktualisieren kann.

Wurde die Evidenz nicht direkt, sondern indirekt über einen dritten Knoten  $Z$  beobachtet, so ergibt sich für das nun entstandene Bayes-Netz  $X \rightarrow Y \rightarrow Z$  der Likelihood-Vektor  $\lambda(x)$  aus Gl. (4.13). Dabei wird ausgenutzt, daß der Knoten  $Y$  die Knoten  $X$  und  $Z$  separiert und sie somit bedingt unabhängig sind.

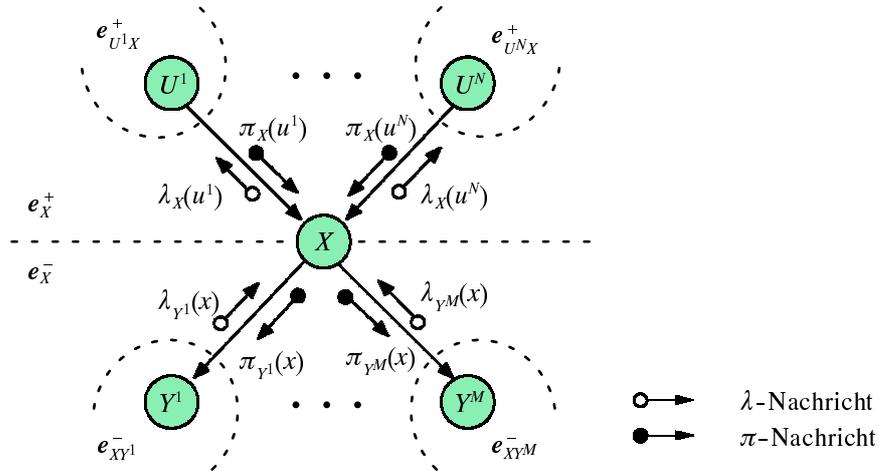


Bild 4.4: Ausschnitt aus einem einfach verbundenen Bayes-Netz

$$\lambda(x) = P(e|x) = \sum_y P(e|y, x)P(y|x) = \sum_y P(e|y)P(y|x) \quad (4.13)$$

$\lambda(x)$  wird demnach im Knoten  $Y$  aus der Matrix  $\underline{M}_{y|x} = P(y|x)$  und der sog.  $\lambda$ -Nachricht  $P(e|y)$ , die der Knoten  $Y$  vom Knoten  $Z$  erhält, berechnet werden. Das Verfahren kann für beliebig lange Ketten von Bayes-Knoten verallgemeinert werden. Jeder Knoten aktualisiert seine Glaubwürdigkeit aus seiner bedingten Wahrscheinlichkeitsmatrix und der  $\lambda$ -Nachricht vom Kindknoten.

Die Überlegungen werden von den bislang betrachteten Ketten auf einfach verbundene Netze, wie in Bild 4.4 dargestellt, erweitert. Ein Knoten  $X$  besitzt nun  $N$  Elternknoten  $U^1, \dots, U^N$  und  $M$  Kindknoten  $Y^1, \dots, Y^M$ . Es sollen die Nachrichten berechnet werden, die der Knoten  $X$  verschickt und empfängt. Mit  $e_X^-$  ( $e_X^+$ ) wird die Evidenz bezeichnet, die im Teilbaum unterhalb (oberhalb) von  $X$  enthalten ist. Für die Evidenzen gilt:

$$e_X^+ = \bigcup_i e_{U^i X}^+ \quad \text{bzw.} \quad e_X^- = \bigcup_j e_{X Y^j}^- \quad (4.14)$$

$$e = e_X^- \cup e_X^+ \quad (4.15)$$

Die Glaubwürdigkeit des Knoten  $X$  berechnet sich gemäß Gl. (4.16) durch Anwendung der Bayes-Regel Gl. (4.5) und unter Berücksichtigung der Separierung von  $e_X^-$  und  $e_X^+$  durch  $X$ :

$$\begin{aligned}
BEL(x) &= P(x|e_X^-, e_X^+) = \frac{P(x, e_X^-, e_X^+)}{P(e_X^-, e_X^+)} \\
&= P(e_X^- | x, e_X^+) P(x|e_X^+) \frac{P(e_X^+)}{P(e_X^-, e_X^+)} \\
&= \alpha P(e_X^- | x) P(x|e_X^+) \\
&= \alpha \lambda(x) \pi(x) \quad \text{mit : } \alpha = \frac{P(e_X^+)}{P(e_X^-, e_X^+)} \quad (4.16)
\end{aligned}$$

Gl. (4.16) beschreibt ein Produkt von einerseits prädiktiven und andererseits erklärenden Faktoren normiert durch den Vektor  $\alpha$ . Der Faktor  $\pi(x) = P(x|e_X^+)$  repräsentiert die Unterstützung der Behauptung  $\{X = x\}$  durch beobachtete Ursachen (z. B. "Der Patient ist Raucher") und wird deshalb auch *kausaler Support* genannt. Der Term  $\lambda(x) = P(e_X^- | x)$  entspricht dem bereits erwähnten Likelihood-Wert und repräsentiert die Unterstützung der Behauptung durch beobachtete Wirkungen (z. B. "Der Patient hat einen positiven Röntgenbefund"). Hier spricht man vom *diagnostischen Support*. Die Trennung von kausalem und diagnostischem Support ermöglicht die Formulierung des *Belief-Update* in Form eines iterativen Prozesses, bei dem jeder Bayes-Knoten seine Glaubwürdigkeit aus den empfangenen Nachrichten seiner Nachbarknoten berechnet. Anschließend bestimmt er aus seiner aktualisierten Glaubwürdigkeit die Nachrichten, die er seinerseits an seine Nachbarn verschickt.

Zunächst soll erläutert werden, wie sich die Terme  $\lambda(x)$  und  $\pi(x)$  in Gl. (4.16) bestimmen lassen. Aufgrund der wechselseitigen bedingten Unabhängigkeit der Kindknoten  $Y^j$  kann  $\lambda(x)$  aus dem Produkt der einzelnen von den Kindknoten eingehenden  $\lambda$ -Nachrichten  $\lambda_{Y^j}$  berechnet werden (vgl. Bild 4.4):

$$\lambda(x) = \prod_{j=1}^M \lambda_{Y^j}(x) \quad (4.17)$$

Der kausale Support  $\pi(x)$  ergibt sich durch Multiplikation der bedingten Wahrscheinlichkeitsmatrix von  $X$  mit dem Produkt der von den übergeordneten Knoten eingehenden  $\pi$ -Nachrichten  $\pi_X(u^i)$ :

$$\pi(x) = \sum_{u^1, \dots, u^N} P(x|u^1, \dots, u^N) \prod_{i=1}^N \pi_X(u^i) \quad (4.18)$$

Hat der Knoten  $X$  seine Glaubwürdigkeit nach Gl. (4.16) berechnet, muß er diejenigen  $\lambda$ - und  $\pi$ -Nachrichten berechnen, die er seinerseits an seine Elternknoten  $U^i$  bzw. Kindknoten  $Y^j$  sendet. Sie ergeben sich aus den Gln. (4.19) und (4.20).

$$\lambda_X(u^i) = \alpha \sum_x \lambda(x) \sum_{w^j: j \neq i} \left( P(x|u^1, \dots, u^N) \prod_{j \neq i} \pi_x(u^j) \right) \quad \forall i = 1, \dots, N \quad (4.19)$$

$$\pi_{Y^j}(x) = \alpha \frac{BEL(x)}{\lambda_{Y^j}(x)} \quad \forall j = 1, \dots, M \quad (4.20)$$

Die Matrix der bedingten Wahrscheinlichkeiten  $P(x|u^1, \dots, u^N)$  in den Gln. (4.18) und (4.19) kann bei Bedarf gemäß Gl. (4.10) aus den einzelnen Kantenmatrizen  $P(x|u^i)$  genähert werden.

Der Prozeß des *Belief-Update* und der Berechnung der Propagierungsnachrichten wiederholt sich für jeden Knoten im Netz solange, bis alle Knoten aktualisiert wurden. Die Wurzel- und die Blattknoten des Bayes-Netzes erfordern eine besondere Behandlung, da sie als Randknoten keine  $\pi$ - bzw.  $\lambda$ -Nachrichten erhalten:

- Für die *Wurzelknoten* wird die  $\pi$ -Nachricht auf die a-priori-Wahrscheinlichkeit des Knotens gesetzt. Mit dieser  $\pi$ -Nachricht wird das Bayes-Netz von den Wurzelknoten aus entsprechend der obigen Propagierungsregeln initialisiert.
- Für einen *Blattknoten*, der (noch) keine Evidenz enthält, ist die Glaubwürdigkeit gleich seiner  $\pi$ -Nachricht. Somit ergibt sich seine  $\lambda$ -Nachricht zu  $\lambda = (1, 1, \dots, 1)$ .

Bild 4.5 illustriert den Propagierungsvorgang für ein baumartiges Bayes-Netz. Zunächst befindet sich das Netz im Gleichgewicht (Bild 4.5a). Der Gleichgewichtszustand wird durch die Einbringung von Evidenz, z. B. durch eine Beobachtung in den Daten, am Knoten  $I$  gestört. Dieser sendet eine Nachricht  $\lambda_I(f)$  an den Knoten  $F$  (Bild 4.5b). Durch diese Nachricht wird der Knoten  $F$  aktiviert, der daraufhin seine Glaubwürdigkeit  $BEL(f)$  aktualisiert und die Nachrichten  $\lambda_F(c)$  und  $\pi_H(f)$  an die Knoten  $C$  bzw.  $H$  sendet. Die Propagierung wird entsprechend fortgesetzt, bis sich schließlich wieder ein Gleichgewichtszustand eingestellt hat (Bild 4.5 d, e, f). Man erkennt, daß die Evidenz sowohl *bottom-up* als auch *top-down* durch das Netz verbreitet wird, so daß eine Änderung in einem Kindknoten auch Auswirkungen auf alle seine Geschwister hat. Diese Eigenschaft unterscheidet das Bayes-Netz von vielen anderen Schlußfolgerungstechniken.

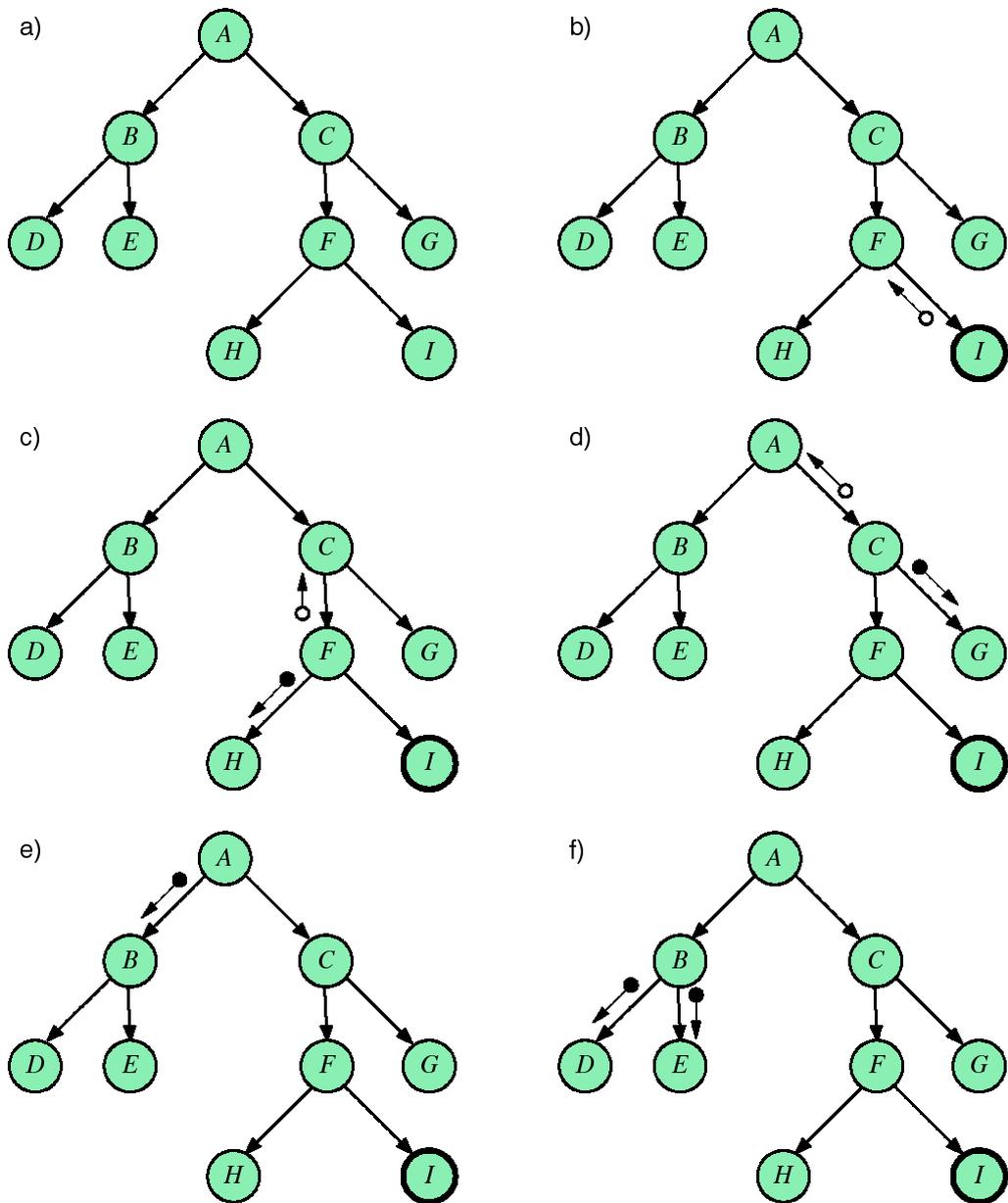


Bild 4.5: Beispiel für die Propagierung von Evidenz (im Knoten I) durch ein Bayes-Netz

$\circ \rightarrow$   $\lambda$ -Nachricht     $\bullet \rightarrow$   $\pi$ -Nachricht

### 4.1.3.1 Behandlung von Schleifen (Loops)

Die bisher betrachteten Bayes-Netze besaßen keine Schleifen (engl.: *Loop*). Eine Schleife ist immer dann vorhanden, wenn es mehr als einen Weg durch das Bayes-Netz gibt, der von einem Knoten  $X$  zu einem Knoten  $Y$  führt. Bild 4.6a zeigt ein Beispiel für ein Bayes-Netz mit Schleife. Bei dem beschriebenen Propagierungsverfahren nach Pearl ist es möglich, daß Nachrichten unbestimmt in den Schleifen umlaufen und somit entweder kein stabiler Gleichgewichtszustand erreicht wird oder bei Erreichen eines stabilen Zustands die Glaubwürdigkeitswerte nicht korrekt sind. Die bei allen Berechnungen zugrundegelegte Annahme der bedingten Unabhängigkeit ist nicht mehr gewährleistet, wenn Schleifen im Netz existieren. So wird in Gl. (4.16) vorausgesetzt, daß der Knoten  $X$  die kausale Evidenz  $e_X^+$  und die diagnostische Evidenz  $e_X^-$  voneinander separiert. Durch das Auftreten von Schleifen kann aber z. B. die Evidenz  $e_X^-$  die Elternknoten von  $X$  über einen Weg beeinflussen, der  $X$  umgeht. Dadurch wäre die bedingte Unabhängigkeit der Elternknoten  $U^i$  und der Kindknoten  $Y^j$  nicht mehr gewährleistet.

Es existieren mehrere Ansätze zur Behandlung von Schleifen wie z. B. das *Clustering-Verfahren* [107] oder die *stochastische Simulation* [20][42][66][107]. An dieser Stelle wird kurz das später eingesetzte *Conditioning-Verfahren* [107][130][131] erläutert. Es beruht darauf, die Schleife in Gedanken an ihrer Wurzel aufzuspalten und dadurch wieder ein einfach verbundenes Bayes-Netz zu erhalten. Der Wurzelknoten wird dupliziert und nacheinander mit seinen möglichen Werten belegt. In Bild 4.6b ist die Aufspaltung für das Beispielnetz dargestellt, wenn am Knoten  $E$  die Evidenz  $+e$  beobachtet wird. Es werden binäre Zufallsvariablen mit den Werten  $+x$  und  $\neg x$  vorausgesetzt. Der Knoten  $A$  wird zunächst mit dem Wert  $+a$  belegt und an beiden durch die Aufspaltung entstandenen Enden ins Netz eingebracht. Die Evidenzen  $+a$  und  $+e$  werden nach den bekannten Propagierungsregeln im Netz verbreitet. Anschließend wird der Vorgang für die Behauptung  $\{A = \neg a\}$  wiederholt. In beiden Fällen ergeben sich unterschiedliche Glaubwürdigkeiten für die Knoten. Sie basieren auf verschiedenen Annahmen - hier hängen sie von den jeweiligen Werten von  $A$  ab - und werden deshalb *bedingte Glaubwürdigkeiten* genannt. Die jeweiligen bedingten Glaubwürdigkeiten werden durch eine gewichtete Mittelung zur Glaubwürdigkeit zusammengefaßt. Für den Knoten  $B$  ergibt sich z. B.:

$$\begin{aligned} BEL(b) &= BEL(b|+a) P(+a|+e) + BEL(b|\neg a) P(\neg a|+e) \\ &= \sum_a BEL(b|a) P(a|+e) \end{aligned} \quad (4.21)$$

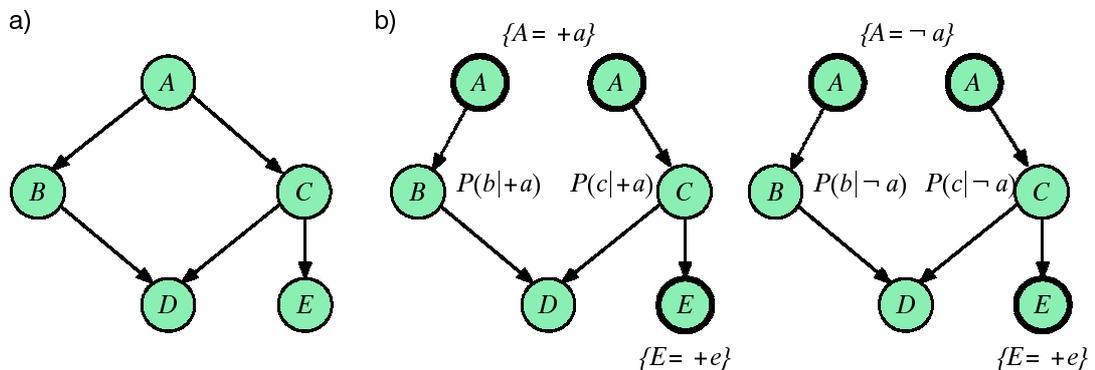


Bild 4.6: a) Bayes-Netz mit Schleife. Die Knoten seien binär.  
 b) Auftrennung der Schleife und Instanzierung der Wurzelknoten mit dem Wert  $+a$  (links) bzw. mit  $\neg a$  (rechts).

Das sog. *Mixing Weight*  $P(a|+e)$  läßt sich mit Hilfe der Bayes-Regel im Knoten  $E$  berechnen:

$$\begin{aligned}
 P(a|+e) &= \alpha \frac{P(+e|a) P(a)}{P(+e)} \\
 &= \alpha P(+e|a) P(a) \quad , \text{ da } P(+e) = 1
 \end{aligned}
 \tag{4.22}$$

Der Term  $P(+e|a)$  ist die Wahrscheinlichkeit für  $\{E = +e\}$  unter der Bedingung  $\{A = a\}$ , bevor die Evidenz beobachtet wurde. Er kann ermittelt werden, indem man den Knoten  $A$  mit dem Wert  $a$  belegt und durch das Netz propagiert.  $P(+e|a)$  ist dann die Komponente des Vektors  $BEL(e)$ , die für den Wert  $+e$  steht. Der Term  $P(a)$  entspricht der  $\pi$ -Nachricht  $\pi(a)$  des Knoten  $A$ .

Das Verfahren läßt sich für die Fälle multipler Schleifen und mehrfacher Beobachtung von Evidenz verallgemeinern [107][130][131]. Die dabei entstehenden Rechenvorschriften stellen eine Generalisierung der in Kapitel 4.1.3 vorgestellten Propagierungsregeln dar. Wendet man das Conditioning-Verfahren an, so ist der Fall der einfach verbundenen Netze automatisch mit abgedeckt. Das Verfahren zeigt allerdings bei stark vernetzten Graphen mit vielen Schleifen Nachteile, da in solchen Fällen alle Schleifen aufgetrennt und über alle Kombinationen von möglichen Wertebelegungen gemittelt werden muß. Die Zahl der Propagierungsprozesse steigt exponentiell mit der Zahl der Schleifen. Bei Bayes-Netzen mit vielen Schleifen sollte deshalb auf andere Techniken, z. B. die stochastische Simulation, zurückgegriffen werden. Die an dieser Stelle betrachteten semantischen Netze sind in der Regel jedoch baumförmig. Werden sie auf ein Bayes-Netz abgebildet, entstehen nur selten Schleifen, so daß der Nachteil des Conditioning-Verfahrens kaum in Erscheinung tritt.

## 4.2 Bayes-Netze für die Bewertung von semantischen Netzen

In den vorangegangenen Ausführungen wurde deutlich, daß die über eine Bayes-Kante verbundenen Bayes-Knoten in einer kausalen Beziehung stehen. In semantischen Netzen bestehen ebenfalls Zusammenhänge zwischen verbundenen Knoten, die je nach Kantentyp unterschiedliche Bedeutungen haben. Somit ist es prinzipiell möglich, ein semantisches Netz als Bayes-Netz zu interpretieren und umgekehrt. Gesucht wurde ein Verfahren zur Bewertung des semantischen Netzes, das sowohl die Evidenz aus den Daten als auch die Erwartungen aus der Wissensbasis berücksichtigt. In anderen Worten: Es sollen sowohl der diagnostische als auch der kausale Support für die Richtigkeit der Szenenbeschreibung in die Bewertung einfließen. Das vorgestellte Propagierungsverfahren für Bayes-Netze gewährleistet diesen bidirektionalen Informationsfluß. Dieses und die strukturelle Ähnlichkeit der beiden Netzwerke legen es nahe, Bayes-Netze zur Bewertung eines semantischen Netzes einzusetzen.

Zunächst muß jedoch das semantische Netz auf ein Bayes-Netz abgebildet werden. Es gilt zu klären, welche Knoten, Kanten und Attribute des semantischen Netzes im Bayes-Netz übernommen werden müssen. Die im semantischen Netz durch Attributwerte repräsentierten Beobachtungen in den Sensordaten dienen als Evidenz für das Bayes-Netz. Mit Hilfe der Propagierungsregeln wird die Evidenz im Netz verteilt, bis sich ein stabiler Gleichgewichtszustand einstellt. Aus den aktualisierten Glaubwürdigkeiten der Bayes-Knoten wird schließlich eine Gesamtglaubwürdigkeit für die Szenenbeschreibung abgeleitet, die als Bewertung des semantischen Netzes herangezogen wird. Auf Basis dieses Bewertungsmaßes selektiert der A\*-Algorithmus die vielversprechendste Alternative unter den konkurrierenden Lösungen.

Die folgenden Abschnitte beschreiben die Transformation des semantischen Netzes in ein Bayes-Netz und auf welche Weise dabei ungenaue Aussagen berücksichtigt werden. Schließlich werden die beiden vorgestellten Bewertungsverfahren vergleichend gegenübergestellt.

### 4.2.1 Vom semantischen Netz zum Bayes-Netz

#### 4.2.1.1 Abbildung der Knoten

Die Knoten eines Bayes-Netzes repräsentieren durch diskrete Zufallsvariablen modellierte Objekte oder Ereignisse. Im semantischen Netz stehen die Knoten ebenfalls für beliebige Objekte und deren Realisierung in der Szene. Während es im Bayes-Netz

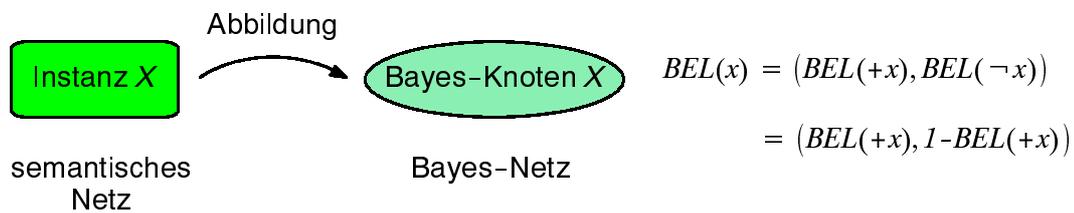


Bild 4.7: *Abbildung einer Instanz des semantischen Netzes auf einen Bayes-Knoten*

nur einen Knotentyp gibt, unterscheidet man im semantischen Netz Konzepte und Instanzen mit ihren verschiedenen Stati von der Hypothese bis zur kompletten Instanz. Die aus Konzepten bestehende Wissensbasis ist für alle Szenenbeschreibungen identisch und muß deshalb nicht beurteilt und somit auch nicht in das Bayes-Netz übernommen werden. Die Instanzen werden hingegen vollständig und unabhängig von ihrem Status auf Bayes-Knoten abgebildet, da durch sie die zu bewertende Szenenbeschreibung repräsentiert wird.

Aus jeder Instanz des semantischen Netzes entsteht demnach ein Bayes-Knoten. Fraglich ist, wieviele Zustände die durch den Bayes-Knoten modellierte Zufallsvariable haben soll. Die Instanzen nehmen im semantischen Netz keine diskreten Werte an, sondern sie repräsentieren durch ihre Existenz in der Szenenbeschreibung die Behauptung, ein entsprechendes Objekt existiere auch in der beobachteten Szene. Diese Behauptung ist entweder wahr oder falsch, wobei diese Entscheidung in der Regel nur unsicher getroffen werden kann. Aus der Sicht des Bayes-Netzes sind derartige Ereignisse durch binäre Zufallsvariablen darzustellen. Eine Instanz, die ein Objekt  $X$  repräsentiert, wird demnach auf einen Bayes-Knoten  $X$  abgebildet, der die beiden Zustände  $+x$  und  $\neg x$  annehmen kann (Bild 4.7). Der erste Zustand steht für die Richtigkeit der Behauptung “ $X$  existiert in der Szene”, der zweite für die Unrichtigkeit. Die Glaubwürdigkeit des Bayes-Knoten  $BEL(x)$  besteht somit aus einem zweidimensionalen Vektor, der die Wahrscheinlichkeiten der beiden Zustände enthält. Da sich beide Wahrscheinlichkeiten zu 1 addieren müssen, ist die Glaubwürdigkeit durch die Wahrscheinlichkeit  $BEL(+x)$  eindeutig definiert.

#### 4.2.1.2 Abbildung der Kanten

Auch bei den Netzwerkkanten unterscheidet das semantische Netz mehr Typen als das Bayes-Netz. Die Kantentypen des semantischen Netzes wurden in Kapitel 2.3.1.3 vorgestellt. Sie beschreiben strukturelle, topologische oder temporale Beziehungen zwischen den Objekten. Eine Bayes-Kante hingegen repräsentiert allgemein eine kausale Abhängigkeit zwischen zwei Knoten. Sie ist von der Ursache zur Wirkung

gerichtet. Für jeden Kantentyp des semantischen Netzes ist zu prüfen, ob die durch ihn verbundenen Knoten in einem Ursache-Wirkung-Verhältnis stehen und die Kante somit auf eine Bayes-Kante abgebildet werden kann und muß:

- ***instance-of***: Diese Kante verbindet eine Instanz der Szenenbeschreibung mit dem korrespondierenden Konzept in der Wissensbasis. Da die Konzepte nicht als Bayes-Knoten übernommen werden, entfällt auch die Abbildung der *instance-of*-Kante.
- ***is-a***: Die *is-a*-Kante beschreibt die Spezialisierung einer allgemeinen Objektklasse in spezifischere Klassen. Zusammen mit dem Vererbungsmechanismus entlang der Kante ist sie ein Hilfsmittel für eine kompaktere Wissensrepräsentationsform. Sie tritt nur im Konzeptnetz der Wissensbasis und nicht in der Szenenbeschreibung auf, da dort die allgemeinen Instanzen durch ihre spezielleren Ausprägungen ersetzt werden. Die Abbildung der Spezialisierungskante entfällt damit.
- ***concrete-of***: Die Konkretisierungskante verbindet Objekte mit ihrer Repräsentation in einer anderen Abstraktionsebene. So ist z. B. eine *Linie* die konkrete geometrische Form des Objekts *Straße*, das zur semantischen Ebene gehört. Aufgrund der Bezeichnung *concrete-of* ist die Kante vom Kind- zum Vaterknoten gerichtet. Fraglich ist, welcher der verbundenen Knoten die Ursache und welcher die Wirkung darstellt. Denkbar ist die Vorstellung, daß z. B. eine *Linie* im Bild die Ursache für die Interpretation *Straße* in der semantischen Ebene ist. Damit hätte die Bayes-Kante dieselbe Richtung wie die *concrete-of*-Kante. Weiter verbreitet ist jedoch die Variante, daß das Objekt der semantischen Ebene die Ursache für die Beobachtung in den Sensordaten bildet. Eine *Straße* in der Szene ist demnach der Grund für die *Linie* im Bild. Das entspricht auch dem Verständnis in dem medizinischen Beispiel aus Bild 4.2, wo ein beobachtetes Symptom als Auswirkung einer Krankheit verstanden worden war. Die Konkretisierungskanten des semantischen Netzes werden somit mit umgekehrter Richtung als Bayes-Kante in das Bayes-Netz übernommen. Dieses gilt auch für kontextabhängige Konkretisierungskanten, da die Kontextabhängigkeit nur Einfluß auf die Reihenfolge der Instanzierung, nicht aber auf die prinzipielle Bedeutung der Kante hat. Über optionale Kanten angebundene Knoten stellen zusätzliche Informationen dar. Sie erhöhen das Vertrauen in die Interpretation, da sie zusätzliche Evidenz liefern. Auch sie werden auf Bayes-Kanten abgebildet.
- ***part-of***: Die *part-of*-Kante zerlegt ein komplexes Objekt in seine Bestandteile. Auch hier werden die Kindknoten, also die Objektbestandteile, als Wirkung der

---

Existenz des übergeordneten Objekts gesehen. So stellt beispielsweise ein *Straßennetz* die Ursache für die Beobachtung der einzelnen *Straßen* dar. Sie liefern den diagnostischen Support für das Ereignis *Straßennetz*. Demnach werden analog zur Konkretisierungskante alle *part-of*-Kanten des semantischen Netzes unter Umkehrung ihrer Richtung auf Bayes-Kanten abgebildet. Kontextabhängige und optionale Kanten werden wiederum wie obligatorische Kanten behandelt.

- ***data-of***: Über die *data-of*-Kante werden segmentierte Bildprimitive an ihre korrespondierenden Instanzen der Szenenbeschreibung gebunden. Sie dient hauptsächlich dazu, eine Referenz zu den Daten herzustellen. Die relevanten Informationen aus den Datenknoten werden in den Attributwerten der Instanzen abgelegt. So wird z. B. die Länge eines segmentierten Liniendatums berechnet und in einem entsprechenden Attribut des Knoten *Linie* abgespeichert. Die Instanzen der Szenenbeschreibung werden bereits als Knoten im Bayes-Netz repräsentiert. Die Abbildung der *data-of*-Kante und die der darüber angebindenen Datenknoten entfällt, da dadurch keine zusätzliche Information ins Bayes-Netz gelangen würde.
- ***attributierte Relation***: Diese Kante beschreibt eine topologische Beziehung zwischen zwei Objekten einer Abstraktionsebene, die durch die enthaltenen Attribute quantifiziert ist. Eine attributierte Relation namens *benachbart* wird beispielsweise ein Attribut enthalten, das die Entfernung der benachbarten Objekte beschreibt. Bei topologischen Relationen läßt sich allerdings keine eindeutige kausale Beziehung im Sinne von Ursache und Wirkung bestimmen. Aus diesem Grund wird die Kante nicht ins Bayes-Netz übernommen. Um jedoch die in der Kante enthaltene Information, z. B. über die gemessene und erwartete Entfernung zweier Objekte, bei der Bewertung berücksichtigen zu können, wird aus den Attributen der Kante eine Evidenz berechnet. Diese Evidenz wird an beiden durch die Kante ursprünglich verbundenen Knoten in das Bayes-Netz eingebracht und durch das Netz propagiert. Auf welche Weise aus Attributwert und -wertebereich eine Evidenz in Form einer  $\lambda$ -Nachricht berechnet wird, erläutert der Abschnitt 4.2.1.3.
- ***temporale Relation***: Die in Kapitel 3.3.1 neu eingeführte temporale Relation beschreibt die zeitliche Veränderung eines Objekts. Durch sie erhält eine Region der Szene über die Zeit unterschiedliche Bedeutungen. Ein Flurstück ändert beispielsweise seine Zustand von *bewachsenes Feld* zu *geerntetes Feld* (vgl. Bild 3.8) oder ein Industriegebiet wird zunächst als inaktives Messegelände und später als Messegelände während einer Aufbauphase (vgl. Bild 3.9) erkannt. Der zeitlich frühere Zustand kann als Ursache für den Folgezustand betrachtet werden,

so daß eine kausale Beziehung zwischen beiden Zuständen besteht. Somit läßt sich die temporale Relation auf eine Bayes-Kante abbilden. Da die Kante bereits von der Ursache zur Wirkung gerichtet ist, muß hier die Richtung nicht wie bei der Konkretisierungs- und Bestandteils-kante umgedreht werden.

Somit werden alle Bestandteils- und Konkretisierungskanten und alle temporalen Relationen des semantischen Netzes im Bayes-Netz übernommen. Jede Bayes-Kante besitzt eine  $2 \times 2$ -Kantenmatrix, da die durch sie verbundenen Bayes-Knoten binär sind und jeweils zwei Zustände einnehmen können. Die Matrix  $\underline{M}_{x|u}$  einer Bayes-Kante vom Knoten  $U$  zum Knoten  $X$  hat allgemein die Form:

$$\underline{M}_{x|u} = \begin{bmatrix} P(+x|+u) & P(\neg x|+u) \\ P(+x|\neg u) & P(\neg x|\neg u) \end{bmatrix} \quad (4.23)$$

Die bedingte Wahrscheinlichkeit  $P(+x|+u)$  beschreibt die Wahrscheinlichkeit für die Richtigkeit der Behauptung  $\{X = +x\}$  unter der Voraussetzung  $\{U = +u\}$ . Mit anderen Worten: Sie steht für die Wahrscheinlichkeit, daß der Knoten  $X$  die richtige Interpretation für sein korrespondierendes Szenenobjekt ist, sofern der Knoten  $U$  ebenfalls die korrekte Beschreibung für sein Szenenobjekt darstellt. Diese Wahrscheinlichkeit läßt sich oftmals aus Expertenwissen ableiten oder zumindest tendenziell schätzen. Die Wahrscheinlichkeit des Gegenereignisses  $P(\neg x|+u)$  ergibt sich aus der Differenz zu 1. Die beiden restlichen Wahrscheinlichkeiten  $P(+x|\neg u)$  und  $P(\neg x|\neg u)$  sind zumeist sehr schwer zu bestimmen. Sie beschreiben die Auftretenswahrscheinlichkeit des Ereignisses  $X$ , wenn  $U$  nicht vorliegt, d. h. wenn die Bedeutungszuweisung für  $U$  nicht korrekt war. Darüber liegt in vielen Fällen kein Vorwissen vor, so daß beide Wahrscheinlichkeiten zweckmäßigerweise auf 0.5 zu setzen sind.

Alle bedingten Wahrscheinlichkeiten der Netzwerkkanten können vom Benutzer in der Wissensbasis vordefiniert werden. Die Angaben werden während der Instanzierung in die Kanten der Szenenbeschreibung und somit in die Bayes-Kanten übernommen. Während sich in der Regel die bedingten Wahrscheinlichkeiten für Bestandteils- und Konkretisierungskanten nur schwer angeben lassen, sind die für temporale Relationen durch die Zustandsübergangswahrscheinlichkeiten des temporalen Wissens gegeben. In der Wissensbasis aus Bild 3.9 wurde beispielsweise die Wahrscheinlichkeit, daß der Zustand *Messeaufbau* vorliegt, wenn vorher der Zustand *Messe inaktiv* erkannt wurde, mit 0.3 angenommen. Die Zustandsübergangswahrscheinlichkeit der temporalen Relationen entspricht der bedingten Wahrscheinlichkeit  $P(+x|+u)$  der Bayes-Kante. Somit ergibt sich für das genannte Beispiel die folgende Kantenmatrix:

$$\underline{M}_{\text{Messeaufbau} | \text{Messe inaktiv}} = \begin{bmatrix} 0.3 & 0.7 \\ 0.5 & 0.5 \end{bmatrix} \quad (4.24)$$

Werden vom Benutzer keine expliziten Angaben über die bedingten Wahrscheinlichkeiten einer Kante gemacht, so wird die Kantenmatrix derart initialisiert, daß die Evidenzen unverändert über die Bayes-Kante weitergereicht wird. Dadurch läßt sich ein neutrales Verhalten der Bayes-Kanten erreichen.

Besitzt ein Bayes-Knoten mehrere Elternknoten  $U$ , so werden die bedingten Wahrscheinlichkeiten  $\underline{M}_{x|u}$  des Bayes-Knoten gemäß Gl. (4.10) aus den einzelnen Kantenmatrizen  $\underline{M}_{x|u_i}$  bestimmt. Dabei werden, wie bereits erwähnt, die Abhängigkeiten der Elternknoten untereinander vernachlässigt. Diese Näherung ist dann zulässig, wenn diese Abhängigkeiten gering sind. In der vorliegenden Anwendung tritt der beschriebene Fall auf, wenn eine Instanz des semantischen Netzes von mehreren übergeordneten Knoten gebunden wird. In Bild 4.9 besitzt beispielsweise die Instanz *Linie-1* die beiden Elternknoten *Asphalt-1* aus der Materialebene und *3D-Streifen-1* aus der Geometrieebene der Wissensbasis. Die beiden Elternknoten stammen aus unterschiedlichen Begriffswelten und haben keinerlei kausale Abhängigkeiten zueinander. Deshalb ist die erwähnte Näherung zur Berechnung der bedingten Wahrscheinlichkeiten  $\underline{M}_{x|u}$  zulässig.

#### 4.2.1.3 Abbildung der Attribute

Die Attribute der Instanzen enthalten in ihren Attributwerten und -wertebereichen die zum Bewertungszeitpunkt vorhandene Evidenz. Die Attributwerte werden durch Aktivierung der Attributberechnungsfunktion aus den Sensordaten oder aus den Attributen benachbarter Instanzen bestimmt. Die Wertebereiche entsprechen den erwarteten Attributwerten und sind aus der Wissensbasis übernommen oder während der Laufzeit durch die entsprechenden Wertebereichsfunktionen ermittelt worden. Gegenstand der Bewertung ist, wie gut die Messungen, also die Attributwerte, mit den Erwartungen, also den Wertebereichen, übereinstimmen.

Sowohl die Attributwerte als auch ihre Erwartungsbereiche sind nur ungenau zu bestimmen. Sie werden, wie in Kapitel 2.4.3 erläutert, durch trapezförmige Fuzzy-Sets repräsentiert. Als Maß für die Übereinstimmung von Attributwert  $E$  (für: Evidenz) und -wertebereich  $H$  (für: Hypothese) wird die Schnittmenge der Fuzzy-Sets genutzt (Bild 4.8). Der Überlappungsbereich der beiden Fuzzy-Mengen ist der Teil der Messung, der zu der Erwartung kompatibel ist und positiv zu werten ist. Meßwerte außerhalb des Erwartungsbereichs sind negativ zu beurteilen. Die beiden Anteile werden durch die Flächen  $A_1$  und  $A_2$  (Bild 4.8) quantifiziert und entsprechen der

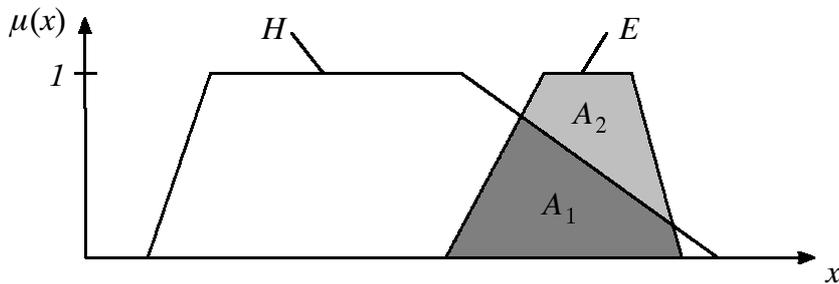


Bild 4.8: Bestimmung der Schnittmenge von Attributwert  $E$  und Erwartungsbereich  $H$  zur Bewertung eines Attributs

Evidenz, die für bzw. gegen die Hypothese sprechen. Durch Normierung der Flächen auf die Gesamtfläche erhält man Verhältniszahlen, die als Wahrscheinlichkeiten interpretiert und als Evidenzvektoren für das Bayes-Netz genutzt werden.

Ein Attribut ist einem Knoten  $X$  im semantischen Netz und somit auch dem Bayes-Knoten  $X$  eindeutig zugeordnet. Um die im Attribut enthaltenen Evidenz in das Bayes-Netz einbringen zu können, werden die Attribute auf sog. *Dummy-Knoten* abgebildet. Diese Knoten haben die Eigenschaft, keine  $\lambda$ - und  $\pi$ -Nachrichten zu empfangen, sondern lediglich eine  $\lambda_D$ -Nachricht an ihren Vaterknoten  $X$  zu verschicken. Dadurch wird gewährleistet, daß die gemessene Evidenz nicht durch Nachrichten der übergeordneten Knoten modifiziert wird. Die Nachricht des Dummy-Knotens  $D$  an den Knoten  $X$  entspricht dem o. a. Evidenzvektor des Attributs:

$$\lambda_D(x) = \left( \frac{A_1}{A_1 + A_2}, \frac{A_2}{A_1 + A_2} \right) = \left( \frac{A_1}{A_1 + A_2}, 1 - \frac{A_1}{A_1 + A_2} \right) \quad (4.25)$$

Besitzt ein Knoten mehrere Attribute, so werden die jeweiligen Evidenzen  $\lambda_{D_i}$  durch eine geometrische Mittelung zu einer  $\lambda_D$ -Nachricht verrechnet. Dabei gehen die einzelnen Attribute entsprechend ihres vordefinierten Attributgewichts  $w_i$  in die Mittelung ein (Gl. (4.26)). Analog werden die Evidenzen aus attributierten Relationen über einen zweiten Dummy-Knoten in das Bayes-Netz eingefügt.

$$\lambda_D(+x) = \sqrt[I]{\prod_{i=1}^I \lambda_{D_i}^{w_i}}(+x) \quad \text{mit } W = \frac{1}{I} \sum_{i=1}^I w_i \quad (4.26)$$

$$\lambda_D(\neg x) = 1 - \lambda_D(+x)$$

Ein Sonderfall für die Bewertung von Instanzen ergibt sich, wenn der Knoten den Status *fehlende Instanz*  $I_F$  besitzt. Dieses tritt immer dann auf, wenn eine Hypothese

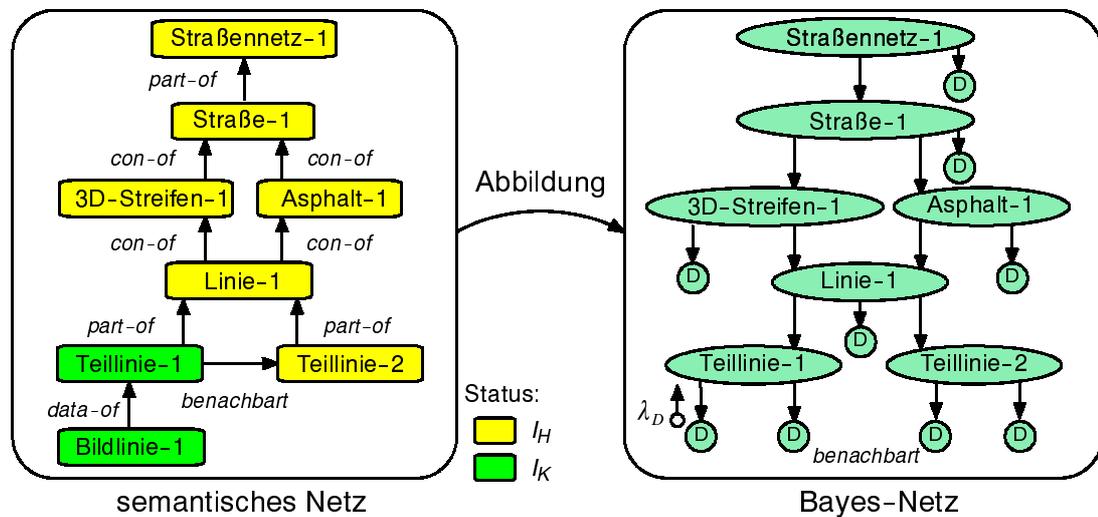


Bild 4.9: Abbildung eines semantischen Netzes auf ein Bayes-Netz für die Bewertung

nicht verifiziert werden kann, da z. B. in den segmentierten Daten kein zu den Erwartungen passender Kandidat gefunden wurde. Sofern die Instanz ein obligatorischer Bestandteil der Szenenbeschreibung ist, ist ihr Fehlen schlecht zu bewerten. Dies geschieht durch eine  $\lambda_D$ -Nachricht, die völlige Unglaubwürdigkeit beschreibt:

$$\lambda_D(x_{I_F}) = (0, 1) \quad (4.27)$$

In Bild 4.9 ist ein Beispiel für die Abbildung eines semantischen Netzes auf ein Bayes-Netz dargestellt. Die Szenenbeschreibung repräsentiert ein Zwischenergebnis für die Extraktion eines Straßennetzes aus einem Luftbild gemäß der Wissensbasis aus Bild 2.4. Alle Instanzen werden ungeachtet ihres Status auf einen Bayes-Knoten abgebildet. Die Bestandteils- und Konkretisierungskanten werden in umgekehrter Richtung als Bayes-Kanten übernommen. Lediglich der Knoten *Bildlinie-1* wird nicht im Bayes-Netz repräsentiert, da er über eine *data-of*-Kante angebunden ist. Jeder Bayes-Knoten erhält einen Dummy-Knoten *D*, über den die Evidenz seiner Attribute in Form von  $\lambda_D$ -Nachrichten in das Netzwerk eingebracht werden kann. Die topologische Relation *benachbart* wird nicht im Bayes-Netz übernommen. Stattdessen wird die Evidenz der in ihr enthaltenen Attribute durch zwei identische Dummy-Knoten an den Knoten *Teillinie-1* und *Teillinie-2* repräsentiert.

#### 4.2.2 Bewertung eines Suchbaumknotens

Um eine durch einen Suchbaumknoten repräsentierte Szenenbeschreibung zu bewerten, wird folgendermaßen vorgegangen:

- 
- Das semantische Netz wird nach den in den vorangegangenen Abschnitten erläuterten Regeln auf ein Bayes-Netz abgebildet.
  - Das Bayes-Netz wird zunächst von seinen Wurzelknoten aus mit  $\pi$ -Nachrichten initialisiert. Diese entsprechen den in der Wissensbasis definierten a-priori-Wahrscheinlichkeiten der Knoten. Insbesondere sind dies die Zustandswahrscheinlichkeiten, die im Zuge des temporalen Wissens (Kapitel 3.3.1) angegeben wurden. Auf diese Weise gelangt das Expertenwissen in Form eines kausalen Supports in das Bayes-Netz. Die Bayes-Knoten wahrscheinlicherer Ereignisse erhalten somit eine höhere Glaubwürdigkeit. Wurde für einen Knoten keine a-priori-Wahrscheinlichkeit definiert, so wird die  $\pi$ -Nachricht für Unwissenheit  $\pi = (0.5, 0.5)$  genutzt.
  - Die  $\pi$ -Nachrichten werden mit Hilfe des Propagierungsverfahrens von Pearl (Kapitel 4.1.3) im Bayes-Netz verteilt, bis alle Glaubwürdigkeiten der Bayes-Knoten aktualisiert sind.
  - Nach der Initialisierung wird jeder Bayes-Knoten auf Evidenz untersucht. Die Attribute der Knoten und attributierten Relationen werden, wie in Kapitel 4.2.1.3 beschrieben, bewertet. Die Einzelbewertungen werden gemäß Gl. (4.26) zu einer  $\lambda_D$ -Nachricht zusammengefaßt. Anschließend wird die Evidenz nach den Regeln von Pearl durch das Netz propagiert. Falls das Bayes-Netz Schleifen (Loops) enthält, wird das in Kapitel 4.1.3.1 beschriebene Conditioning-Verfahren eingesetzt. Nach Erreichen eines Gleichgewichtszustands enthält jeder Bayes-Knoten eine aktualisierte Glaubwürdigkeit  $BEL(x)$ , die eine Bewertung des Knotens repräsentiert. Aufgrund des Propagierungsverfahrens wurden sowohl der kausale als auch der diagnostische Support bei der Bewertung berücksichtigt.
  - Schließlich muß aus den einzelnen Knotenbewertungen eine Gesamtbewertung für die Szenenbeschreibung berechnet werden, die die Systemsteuerung nutzen kann, um sie mit den Bewertungen konkurrierender Szenenbeschreibungen zu vergleichen. Relevant für die Qualität der Szenenbeschreibung sind letztlich die Knoten der semantischen Ebene, die die symbolischen Bedeutungen der Szenenobjekte enthalten. In den Glaubwürdigkeiten dieser Knoten sind die Bewertungen ihrer Konkretisierungen durch die empfangenen  $\lambda$ -Nachrichten bereits enthalten. Die Gesamtbewertung für den Suchbaumknoten  $\mathcal{N}$  wird deshalb nur aus den Glaubwürdigkeiten der  $K$  Knoten der semantischen Ebene durch eine gewichtete arithmetische Mittelung berechnet. Jeder Knoten  $X_k$  geht mit seinem Knotengewicht  $w_k$  in die Bewertung ein:

$$BEL(\mathcal{N}) = \frac{\sum_{k=1}^K w_k BEL(x_k)}{\sum_{k=1}^K w_k} \quad (4.28)$$

Um eine optimistische Bewertung der Szenenbeschreibung zu gewährleisten, werden solche Knoten der semantischen Ebene, die keinen diagnostischen Support in Form einer  $\lambda$ -Nachricht erhalten, mit einer Glaubwürdigkeit von  $BEL(x_k) = (1, 0)$  berücksichtigt. Dadurch können unverifizierte modellgetriebene Hypothesen, die eventuell durch schlechte  $\pi$ -Nachrichten eine geringe Glaubwürdigkeit haben, nicht die Gesamtbewertung des Suchbaumknotens negativ beeinflussen. Ansonsten könnte es passieren, daß allein aufgrund der Erzeugung einer neuen modellgetriebenen Hypothese die Bewertung der Szenenbeschreibung verschlechtert und die Analyse daraufhin in einem konkurrierenden Suchbaumknoten fortgesetzt wird.

Die Güte der Szenenbeschreibung wird durch die Komponente  $BEL(+\mathcal{N})$  der Suchbaumknotenbewertung  $BEL(\mathcal{N})$  ausgedrückt. Dieser Wert wird dem A\*-Algorithmus zugeführt, der aus allen konkurrierenden Alternativen die bestbewertete für die weitere Analyse auswählt.

Das vorgestellte probabilistische Verfahren zur Bewertung von semantischen Netzen liefert letztlich für jede Szenenbeschreibung ein Qualitätsmaß zwischen 0 und 1. Damit ist es kompatibel zu dem in Kapitel 2.4.3 beschriebenen possibilistischen Ansatz, bei dem die Möglichkeit  $Po(\mathcal{N})$  als Entscheidungskriterium für die Steuerung der Bildanalyse herangezogen wurde. Die beiden Verfahren können wahlweise als Bewertungssystem eingesetzt werden, ohne daß Änderungen an der Wissensbasis nötig werden. Der folgende Abschnitt stellt die beiden Ansätze vergleichend gegenüber und diskutiert Vor- und Nachteile.

---

### 4.3 Vergleich mit dem possibilistischen Bewertungssystem

Das possibilistische Bewertungsverfahren aus Kapitel 2.4.3 unterscheidet Notwendigkeit und Möglichkeit einer Behauptung, was einer pessimistischen bzw. optimistischen Abschätzung der Kompatibilität von Meßwert und Erwartung entspricht. Die Differenz wird als Unwissenheit interpretiert, die durch die fortschreitende Analyse der Daten reduziert wird. Die Unwissenheit kann als optimistischer Wert der Restgüte in der Bewertungsfunktion des  $A^*$ -Algorithmus (Gl. (2.7) und (2.8)) genutzt werden.

Um aus den Einzelbewertungen der Knoten und Attribute eine Gesamtbewertung der Szenenbeschreibung zu erhalten, werden einfache Operationen wie der geometrische Mittelwert oder die Kombinationsregel von Dempster verwendet. Aus diesem Grund benötigt das Bewertungsverfahren auch für große semantische Netze wenig Rechenkapazität.

In die possibilistische Bewertung fließen ausschließlich die Evidenzen aus den Attributen, also der diagnostische Support ein. Das führt zu dem erwähnten Effekt, daß konkurrierende Szenenbeschreibungen, die sich bei ansonsten identischer Evidenz nur in einer modellgetriebenen Hypothese unterscheiden, gleich gut bewertet werden, so daß die Analyse in einem zufällig ausgesuchten Suchbaumknoten fortgesetzt wird. Selbst wenn zusätzliches Wissen über Auftretenswahrscheinlichkeiten der Ereignisse vorliegt, kann es bei dem possibilistischen Verfahren prinzipbedingt nicht ausgenutzt werden. In der Possibility-Theorie lassen sich keine bedingten Wahrscheinlichkeiten, wie z. B. die Zustandsübergangswahrscheinlichkeiten repräsentieren [107]. Denkbar wäre lediglich die Berücksichtigung von Verbundwahrscheinlichkeiten, die jedoch in der Regel nur sehr schwer bestimmbar sind. Die Possibility-Theorie ist nicht geeignet, sowohl diagnostischen als auch kausalen Support zu verarbeiten.

Bayes-Netze unterstützen einen bidirektionalen Informationsfluß. Der diagnostische Support wird über die  $\lambda$ -Nachrichten von unten nach oben, der kausale Support mittels der  $\pi$ -Nachrichten von oben nach unten im Netz verteilt. Dadurch werden sowohl die Beobachtungen in den Daten als auch die Erwartungen aus dem Vorwissen für die Bewertung genutzt. In der vorliegenden Anwendungen werden die Zustands- und Zustandsübergangswahrscheinlichkeiten des temporalen Wissens als  $\pi$ -Nachrichten bzw. als bedingte Wahrscheinlichkeiten in den Kantenmatrizen verwendet. Bei konkurrierenden Szenenbeschreibungen mit gleichem diagnostischen Support entscheidet der höhere kausale Support und führt zu einer besseren Gesamtbewertung. Auf diese Weise kann die a priori wahrscheinlichere Alternative ausgewählt und bevorzugt

---

untersucht werden. Die Zufälligkeit bei der Analyse, wie sie beim possibilistischen Ansatz auftritt, entfällt.

Beim vorgeschlagenen Bewertungssystem basierend auf Bayes-Netzen wird eine gemittelte Glaubwürdigkeit als Entscheidungskriterium für die Steuerung der Analyse genutzt. Ein Maß für die Unwissenheit existiert nicht, da lediglich aktualisierte Glaubwürdigkeiten für die einzelnen Knoten berechnet werden.

Ein Bayes-Netz stellt ein geschlossenes probabilistisches Modell einer gegebenen Problemstellung dar. Dadurch wird es erforderlich, daß alle Modellgrößen vollständig bekannt sind. Neben der Struktur des Netzes wird für jeden Bayes-Knoten eine komplette Matrix von bedingten Wahrscheinlichkeiten benötigt, die in der Praxis zumeist nicht explizit bekannt sind. Dieses wird oft als Nachteil der Bayes-Netze gewertet.

In der vorliegenden Anwendung ist die Struktur des Bayes-Netzes durch die des zu bewertenden semantischen Netzes vorgegeben, so daß kein zusätzlicher Modellierungsaufwand entsteht. Alle Bayes-Knoten repräsentieren binäre Zufallsvariablen, wodurch lediglich zweidimensionale Vektoren für Glaubwürdigkeit sowie  $\lambda$ - und  $\pi$ -Nachricht zu berechnen sind. In den meisten Fällen sind die Netze baumartig, so daß jeder Bayes-Knoten in der Regel nur einen Vaterknoten besitzt. Die bedingten Wahrscheinlichkeiten des Knotens entsprechen dann einer  $2 \times 2$ -Kantenmatrix. In den seltenen Fällen, in denen ein Knoten mehrere Eltern hat, wird die Matrix der bedingten Wahrscheinlichkeiten durch das Produkt der einzelnen Kantenmatrizen genähert (Gl. (4.10)). Die Kantenmatrizen werden derart initialisiert, daß die Evidenz unverändert über die Kanten propagiert wird. Der diagnostische Support wird automatisch aus den aktuellen Attributwerten und -wertebereichen berechnet. Somit kann die Bewertung mittels Bayes-Netzen auch ohne die explizite Angabe von Wahrscheinlichkeiten erfolgen. In diesem Fall weist das Bayes-Netz ein vergleichbares Bewertungsverhalten wie der possibilistische Ansatz auf.

Durch die Definition einzelner Wahrscheinlichkeiten hat der Anwender die Möglichkeit, gezielt Expertenwissen und damit kausalen Support einzubringen. Wird temporales Wissen in Form eines Zustandsübergangsdiagramms definiert, werden die darin enthaltenen Wahrscheinlichkeiten für Zustände und Zustandsübergänge bei der Bewertung auf Basis von Bayes-Netzen ausgewertet. Nach der Prädiktion konkurrierender Folgezustände während einer multitemporalen Bildanalyse (s. Kapitel 3.3.2) wird dadurch stets der wahrscheinlichste Zustandsübergang am besten bewertet und deshalb als erster verfolgt. Somit wird in vielen Fällen die Effizienz des Analyseprozesses gesteigert, wie im folgenden Kapitel an konkreten Anwendungsbeispielen

aufgezeigt werden wird. Da letztlich nicht die absolute Bewertung einer Szenenbeschreibung, sondern die relativen Unterschiede der Bewertungen konkurrierender Szenenbeschreibung für den Analyseverlauf entscheidend sind, ist auch keine hohe absolute Genauigkeit für die a-priori-Wahrscheinlichkeiten erforderlich. Vielmehr sollen durch ihre Definition die wahrscheinlichen Ereignisse von den unwahrscheinlichen unterschieden werden.

Das vorgestellte Propagierungsverfahren verteilt die Nachrichten im Bayes-Netz solange, bis die Glaubwürdigkeiten aller Knoten aktualisiert worden sind. Hierbei handelt es sich um einen rekursiven Prozeß. Treten Schleifen im Bayes-Netz auf, so wird der Vorgang durch den Einsatz des Conditioning-Verfahrens mehrfach wiederholt. Der Rechenaufwand für die Bewertung mit Bayes-Netzen ist deshalb prinzipiell höher als der für die Bewertung mit dem possibilistischen Ansatz, bei dem die Einzelbewertungen lediglich *bottom-up* zu einer Gesamtbewertung aggregiert werden.

Zusammenfassend läßt sich feststellen, daß das probabilistische Bewertungssystem durch die Berücksichtigung von a-priori-Wahrscheinlichkeiten zu einem deterministischen und oftmals effizienteren Analyseverlauf führt (s. Kapitel 5), da die wahrscheinlicheren Lösungen stets bevorzugt untersucht werden. Enthält die Wissensbasis keine a-priori-Wahrscheinlichkeiten, so entfällt dieser Vorteil. In solchen Fällen stellt das possibilistische Bewertungssystem eine robuste Alternative mit besserem Laufzeitverhalten dar.

---

## 5. Anwendungsbeispiele

An zwei konkreten Anwendungsbeispielen soll im folgenden die Funktionalität des Systems zur Interpretation multitemporaler Luftbilder aufgezeigt werden. In beiden Beispielen wird temporales Wissen für die Analyse der Bilder eingesetzt. Kapitel 5.1 greift das Beispiel aus Kapitel 3.3 auf und behandelt die Detektion eines Messegeländes aufgrund der Beobachtung eines charakteristischen zeitlichen Verhaltens. Der Einsatz des Bildinterpretationssystem zur Detektion baulicher Veränderungen wird in Kapitel 5.2 geschildert.

Für beide Anwendungsbeispiele werden überlappende Luftbilder des Messegeländes Hannover aus den Jahren 1993, 1995, 1997 und 1998 genutzt<sup>1</sup>. Einige der Senkrechtaufnahmen sind Farbbilder, bei der Mehrzahl handelt es sich jedoch um Grauwertbilder. Die Maßstäbe variieren zwischen 1:6000 und 1:40000. Die Bilder einer Befliegung wurden zu jeweils einem Bildmosaik verschnitten und in geographische Koordinaten (Gauß-Krüger-Koordinaten) umgerechnet. Dieser Vorgang wird auch *Geokodierung* genannt. Die so entstandenen sog. *Orthophotos* erleichtern die Herstellung eines Bezugs zu anderen Orthophotos des Beobachtungsgebiets und zu einer Karte oder den Daten eines GIS-Systems. Für die vorliegenden Anwendungsbeispiele wurden für jeden Aufnahmezeitpunkt Orthophotos mit Auflösungen von 2 m, 1 m und 0,5 m pro Bildpunkt hergestellt. Bild 5.1 zeigt exemplarisch das Orthophoto aus dem Jahr 1997.

### 5.1 Detektion eines Messegeländes

Wie bereits in vorherigen Ausführungen erläutert, läßt sich ein Messegelände nicht aus einem einzelnen Luftbild erkennen. Die Detektion der Hallen und Parkplätze allein reicht nicht zur Klassifikation aus. Ein Messegelände zeichnet sich vielmehr durch den Zyklus von Inaktivität, Messeaufbau, Messezeit und Messeabbau (vgl. Bild 3.9) aus, was eine multitemporale Beobachtung des Geländes erforderlich macht.

Die folgenden Abschnitte behandeln den Ablauf der Bildinterpretation und die dabei eingesetzte Wissensbasis sowie die genutzten Daten und Bildverarbeitungsverfahren. Darüber hinaus erfolgt für das vorliegende Beispiel ein Vergleich der beiden vorgestellten Bewertungsverfahren bezüglich ihrer Effizienz.

---

1. Mit freundlicher Genehmigung der Firma Scout-Systems, des Landesbetriebs Landesvermessung & Geobasisinformation Niedersachsen (LGN) und des Stadtvermessungsamts Hannover



Bild 5.1: Luftbild des Messegelände Hannover aus dem Jahr 1997 (Orthophoto)

### 5.1.1 Datenmaterial und Wissensbasis

Die Wissensbasis für die Detektion eines Messegeländes baut auf dem Zustandsübergangsdiagramm aus Bild 3.9 auf (Bild 5.2). Jeder Zustand des Messegeländes ist als Spezialisierung des Konzepts *Industriegebiet* modelliert. Ein Industriegebiet besteht

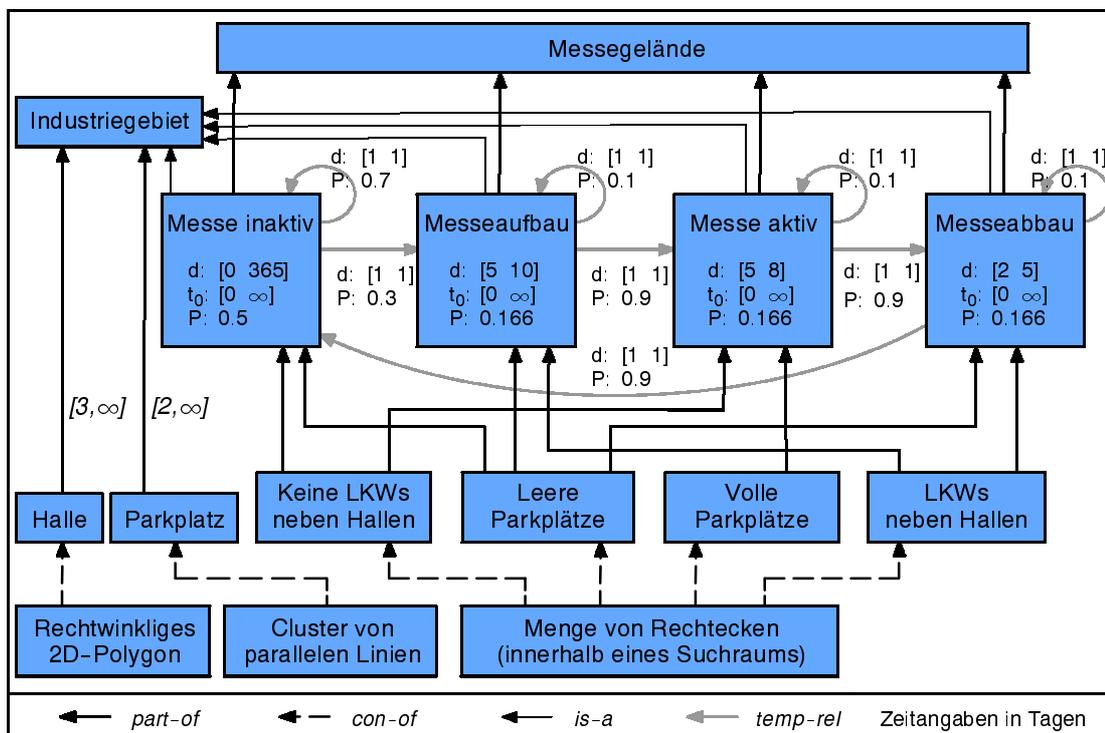


Bild 5.2: Vereinfachtes semantisches Netz für die Erkennung eines Messegeländes

hier aus mindestens drei *Hallen* und zwei *Parkplätzen*. Hallen sind zumeist rechtwinklige Gebäude, so daß sie im Bild als *rechtwinklige 2D-Polygone* abgebildet werden. Parkplätze lassen sich in den vorliegenden Luftbildern an den einzelnen Fahrspuren erkennen und werden daher als *Cluster von parallelen Linien* in der Wissensbasis repräsentiert. Jeder der vier Zustände des Messegeländes erbt über die *is-a*-Kante die Bestandteile *Halle* und *Parkplatz* vom Konzept *Industriegebiet*. Zusätzlich sind sie - wie bereits in Kapitel 3.3.1.1 erläutert - durch eine mehr oder weniger große Anzahl von Fahrzeugen auf den Parkplätzen und zwischen den Hallen gekennzeichnet.

Um entscheiden zu können, ob Parkplätze voll oder leer sind oder ob sich LKWs zwischen den Hallen befinden oder nicht, sind die Fahrzeuge in den entsprechenden Bildregionen zu detektieren und zu zählen. Allgemein sind Fahrzeuge als kleine Rechtecke im Bild zu erkennen. Der jeweilige Suchraum, wo im Bild nach Fahrzeugen zu suchen ist, ergibt sich erst zur Laufzeit der Bildanalyse. Wird z. B. das aktuelle Bild auf volle Parkplätze untersucht, so werden innerhalb der zuvor als Parkplatz klassifizierten Bildregionen alle Fahrzeugkandidaten gezählt. Übersteigt die Zahl einem vorgegebenen Schwellwert, so gelten die Parkplätze als voll, ansonsten als leer. Die notwendigen Parameter wie Größe der zu detektierenden Rechtecke, interessieren-

de Bildregion und Schwellwert für die Fahrzeugzahl werden während der Interpretation modellgetrieben zum Knoten *Menge von Rechtecken* propagiert.

Ein Messegelände kann laut Wissensbasis dann instanziiert werden, wenn in einer Reihe von multitemporalen Luftbildern alle vier Zustände erkannt werden konnten, da die Zustände als obligatorische Bestandteile des Konzepts *Messegelände* definiert sind. Somit wird ein Datensatz benötigt, in dem alle Phasen eines Messezyklus in der richtigen Reihenfolge und mit passenden zeitlichen Abständen zwischen den Aufnahmezeitpunkten abgebildet sind. Da die Befliegungen der Vermessungsämter in der Regel nicht in solch kurzen Zeitabständen erfolgen, ist es nicht gelungen, einen Datensatz zu beschaffen, der ein und dieselbe Messe mit all ihren Phasen darstellt. Daher ist es erforderlich, einen derartigen Datensatz zu simulieren: Die vorliegenden Luftbilder zeigen das Messegelände Hannover inaktiv (1993, 1995), während einer Auf- bzw. Abbauphase (1997) und während einer Messe (1998). Durch die Manipulation der Aufnahmezeitpunkte läßt sich eine Bildreihe der gewünschten Art zusammensetzen, auch wenn nicht die identische Messe abgebildet ist. Den folgenden Ausführungen liegt die Bildreihe 1993-1997-1998-1997 mit den manipulierten Zeitstempeln 5.3., 10.3., 15.3. bzw. 25.3. zu Grunde.

### 5.1.2 Bildverarbeitungsverfahren

Bevor den Szenenobjekten eine symbolische Bedeutung zugeordnet werden kann, müssen mit Hilfe geeigneter Bildverarbeitungsverfahren Kandidaten für die einzelnen Objektklassen segmentiert werden. Im vorliegenden Beispiel sind Hallen, Parkplätze und Fahrzeuge zu detektieren. Zur Detektion von Hallen sind laut Wissensbasis rechtwinklige Polygone zu extrahieren. Eine kantenbasierte Vorgehensweise direkt im Luftbild liefert aufgrund des sehr unterschiedlichen Aussehens der Hallen nur unbefriedigende Ergebnisse. Aus diesem Grund erfolgt die Segmentierung von Hallen in einem digitalen Höhenmodell (DHM), das in einem regelmäßigen Raster Höhenwerte für die Geländepunkte enthält. Derartige Modelle werden entweder durch eine Befliegung mit einem Laserscanner aufgenommen oder - wie im vorliegenden Beispiel - aus überlappenden Luftbildern mittels Stereokorrespondenzanalyse [54][88] berechnet. In derartigen Höhenmodellen können Gebäude aufgrund des signifikanten Höhensprungs gut segmentiert werden. Die extrahierten Regionen werden anhand ihrer Kontur polygonal beschrieben (s. Anhang A-1) und dem Bildinterpretationssystem als Hallenkandidaten zur Verfügung gestellt.

Die Parkplätze werden aufgrund der Anhäufung von parallelen Linien im Bild detektiert, die von den regelmäßig angeordneten Fahrspuren herrühren (s. Anhang A-2). Zunächst werden Linien einer gewissen Breite und Länge mit Hilfe eines

---

Liniendetektors [144] extrahiert. Die Linien werden auf ihre Länge und Richtung hin untersucht und in Gruppen (nahezu) paralleler Linien eingeteilt. Benachbarte parallele Linien werden zu Clustern zusammengefaßt, deren Hülle als Parkplatzkandidat übernommen wird.

Fahrzeuge sind, wie bereits erwähnt, als kleine Rechtecke im Bild erkennbar. Das Bildinterpretationssystem gibt aufgrund der aktuellen Zwischenergebnisse den Suchraum im Bild und die erwartete Größe der Rechtecke vor. Entweder werden LKWs zwischen den Hallen oder PKWs auf den Parkplätzen gesucht. Über Schwellwertoperatoren mit anschließender Elimination zu großer Regionen werden Fahrzeugkandidaten segmentiert und gezählt. Die Zahl der Kandidaten wird dem Interpretationssystem als Ergebnis zurückgeliefert, das seinerseits eine Bewertung bezüglich der erwarteten Fahrzeugzahl durchführt.

Die Segmentierung von Hallen und Parkplätzen kann vor Beginn der Bildanalyse erfolgen. Die Fahrzeuge hingegen können erst während des Interpretationsprozesses segmentiert werden, da der Suchraum im Bild aus den aktuellen Interpretationsergebnissen abgeleitet wird. Die Randbedingungen der Segmentierung werden somit durch das Interpretationssystem situationsabhängig vorgegeben.

### 5.1.3 Interpretationsprozeß

Die Interpretation der multitemporalen Bildfolge beginnt mit der Suche nach einem *Industriegebiet*. Aus diesem Grund erzeugt das System eine Hypothese des gleichnamigen Konzepts. Zur Instanzierung fehlen die obligatorischen Bestandteile *Halle* und *Parkplatz*. Zunächst werden Hallen durch wiederholte Generierung der Hypothesen  $I_H(\text{Halle})$  und  $I_H(\text{rechtwinkliges } 2D\text{-Polygon})$  gesucht. Aus den zuvor segmentierten Hallenkandidaten werden nach und nach diejenigen Polygone ausgewählt, die den erwarteten Eigenschaften entsprechen. Hierbei werden Formmerkmale aus [152] wie Fläche, Länglichkeit und Kompaktheit des Polygons genutzt. Desweiteren werden nur solche Polygone als Halle akzeptiert, die in der Nachbarschaft einer anderen, bereits detektierten Halle liegen. Hierbei wird angenommen, daß semantisch zusammengehörige Gebäudekomplexe auch räumlich beieinander liegen. Bei der Suche nach der ersten Halle, für die die Nachbarschaftsbedingung nicht angewandt werden kann, wird das größte in Frage kommende Polygon ausgewählt, das mit großer Sicherheit eine Halle darstellt. Bild 5.3 zeigt die segmentierten und davon akzeptierten Hallen für das Luftbild aus dem Jahr 1997.

Nach der Instanzierung aller Hallen erfolgt analog die Suche nach Parkplätzen aus der Menge der zuvor segmentierten Parkplatzkandidaten. Auch hier entscheiden

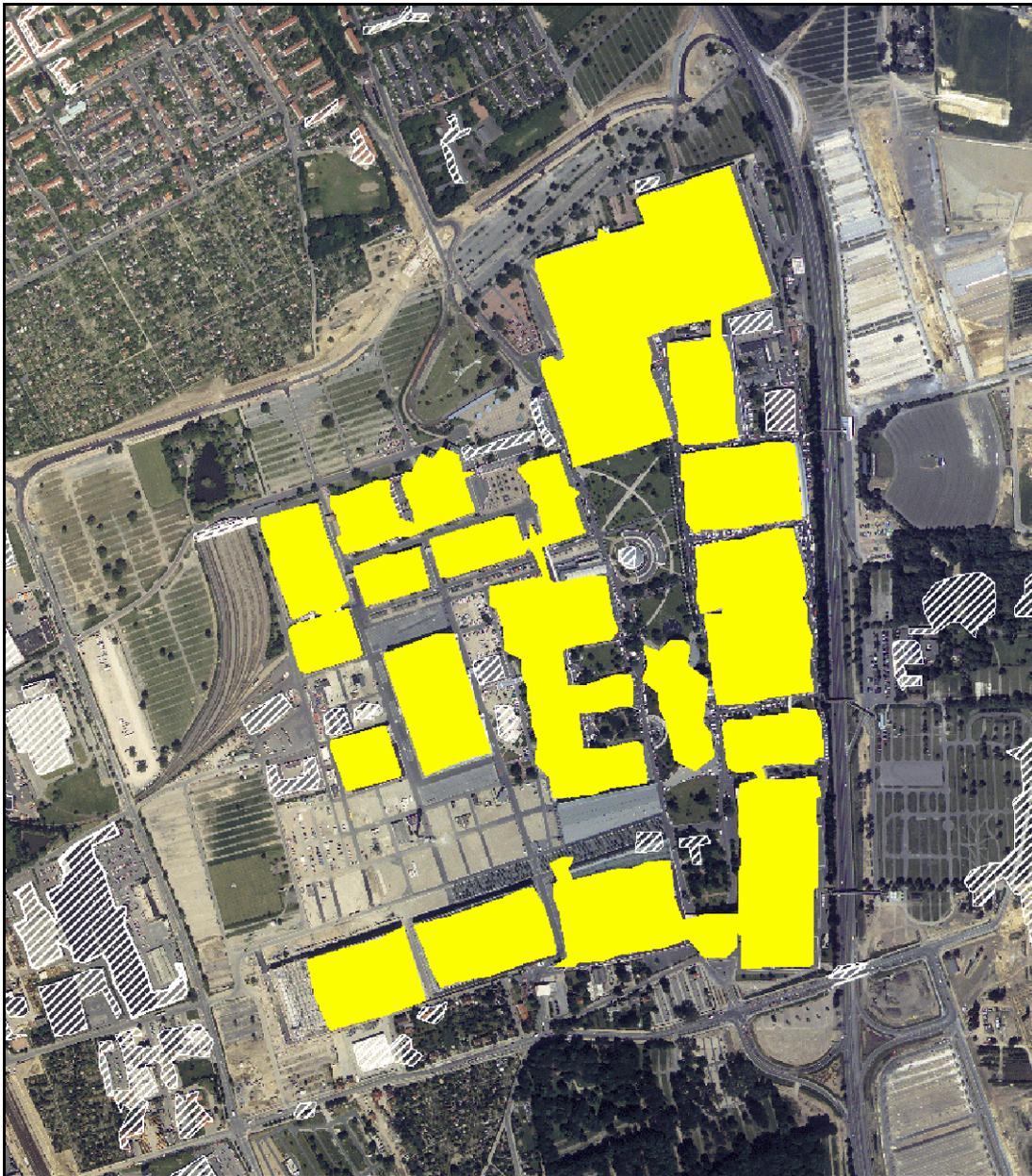


Bild 5.3: *Abgelehnte (schraffiert) und akzeptierte (gefüllt) Hallenkandidaten für das Luftbild des Jahres 1997.*

Formmerkmale über die Akzeptanz oder Ablehnung eines Liniencusters. Zusätzlich wird die Lage der detektierten Hallen ausgenutzt. Liniencuster, die innerhalb des durch die Hallen umgebenen Gebiets liegen oder mit einer Halle deckungsgleich sind, werden

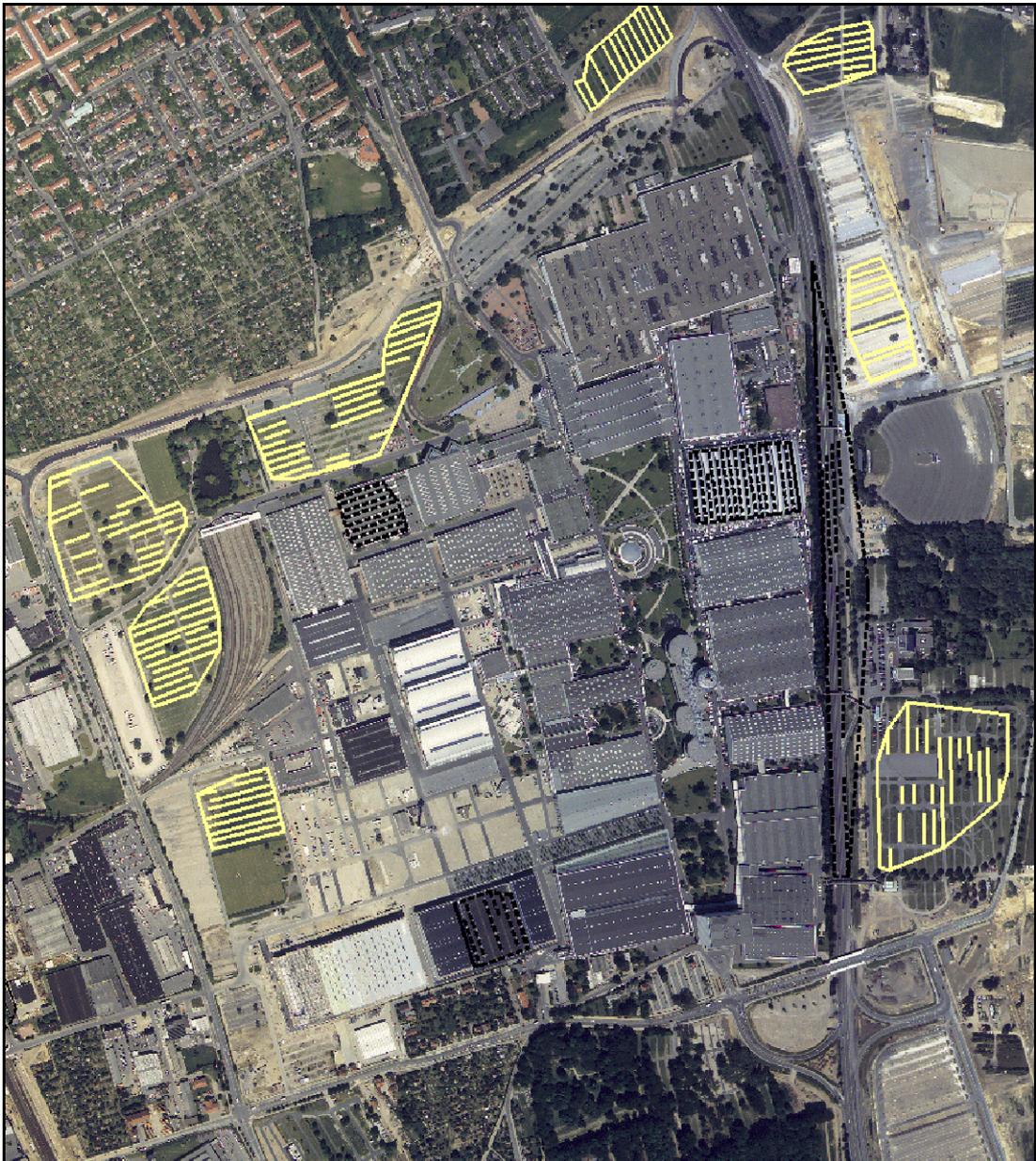


Bild 5.4: Als Parkplätze abgelehnte (dunkel gestrichelt) und akzeptierte Cluster paralleler Linien (hell) für das Luftbild des Jahres 1997.

nicht akzeptiert, da Parkplätze in der Regel außerhalb eines Messegeländes liegen. Das Ergebnis für das Luftbild des Jahres 1997 ist in Bild 5.4 dargestellt.

Wurden genügend Hallen und Parkplätze detektiert, so kann die Hypothese  $I_H$ (Industriegebiet) komplett instanziiert werden. Das Zielkonzept *Messegelände* ist



---

LKWs neben den Hallen stehen. Für die restlichen Alternativen ergeben sich aus der Wissensbasis andere Kombinationen. Das System erzeugt entsprechende Hypothesen und aktiviert das zugehörige Bildverarbeitungsverfahren, das in den vorgegebenen Suchgebieten Fahrzeuge detektiert. Liegt die Zahl der gefundenen Fahrzeuge innerhalb des jeweiligen Erwartungsbereichs, gilt die Hypothese als bestätigt, anderenfalls erhält sie eine schlechte Bewertung, so daß das System die Analyse in einem anderen, besser bewerteten Suchbaumknoten fortsetzt. Auf diese Weise wird (je nach Bewertungsverfahren) mehr oder weniger zielstrebig die korrekte Interpretation für das erste Bild des Datensatzes erreicht. Es handelt sich um ein Messegelände während einer Inaktivitätsphase.

Für die Analyse des zweiten Bildes werden mittels temporaler Inferenz, wie in Kapitel 3.3.2 beschrieben, mögliche Folgezustände für den gefundenen Zustand *Messe inaktiv* prognostiziert. Gemäß der Wissensbasis aus Bild 5.2 kommen bei dem hier angenommenen zeitlichen Abstand der Aufnahmezeitpunkte von fünf Tagen zwei Folgezustände in Frage. Entweder das Gelände befindet sich unverändert im Zustand *Messe inaktiv*, oder eine Messeaufbauphase hat begonnen. Anstatt das zweite Bild auf Hinweise für alle vier möglichen Zustände des Geländes zu prüfen, brauchen aufgrund des zeitlichen Wissens nur zwei Alternativen untersucht werden. Die beiden anderen Zustände können von vornherein ausgeschlossen werden, was die Effizienz der Bildanalyse erhöht.

Analog zu Bild 3.11 erzeugt das System über die temporale Relation Hypothesen für jeden der beiden möglichen Folgezustände und legt sie in zwei neuen Suchbaumknoten ab ( $\mathcal{N}_7$  und  $\mathcal{N}_8$  in Bild 5.7). Anschließend wird versucht, die Hypothesen in den Daten zu verifizieren, indem wie zuvor Hallen, Parkplätze und Fahrzeuge detektiert werden. Welche der beiden Alternativen als erste untersucht wird, hängt wiederum von der Wahl des Bewertungssystems ab. Während das possibilistische Verfahren zufällig einen Suchbaumknoten auswählt, selektiert das Verfahren auf Basis von Bayes-Netzen stets die a priori wahrscheinlichere Alternative. In diesem Fall beinhaltet der Knoten  $\mathcal{N}_7$  die wahrscheinlichere Lösung, da gemäß Wissensbasis die Wahrscheinlichkeit, daß der Zustand *Messe inaktiv* unverändert bleibt, höher ist als der Übergang in den Zustand *Messeaufbau*. Letztlich wird im zweiten Luftbild der Zustand *Messeaufbau* verifiziert, da ausreichend viele Fahrzeuge zwischen den Hallen bei gleichzeitig leeren Parkplätzen detektiert werden können. Bild 5.6 zeigt exemplarisch die extrahierten Fahrzeuge für einen Luftbildausschnitt aus dem Jahr 1997. Der aus den detektierten Hallen (Bild 5.3) abgeleitete Suchraum ist dunkel, die erkannten Fahrzeuge sind weiß eingefärbt. Man

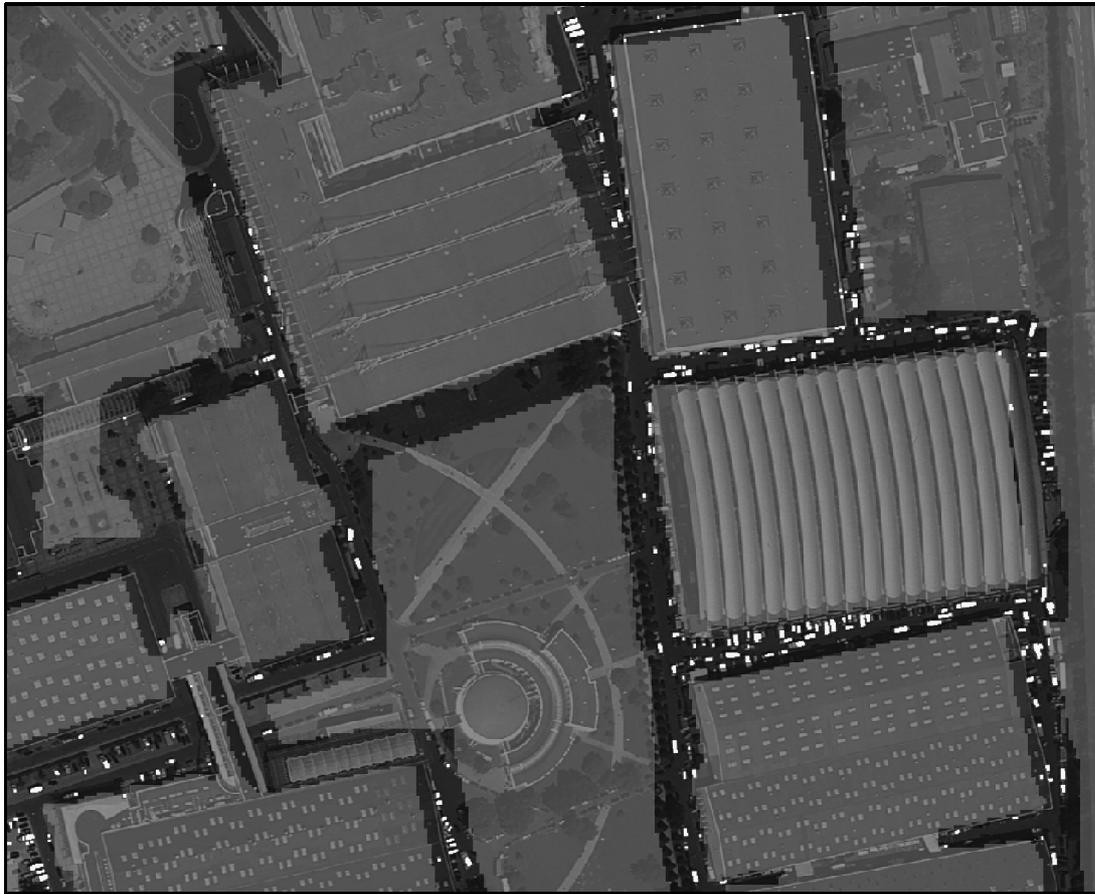


Bild 5.6: *Detektierte Fahrzeuge (weiß) zwischen den Messehallen für einen Ausschnitt des Luftbilds des Jahres 1997. Die dunklen Bildregionen stellen den aus den zuvor detektierten Hallen abgeleiteten Suchraum für das Segmentierungsverfahren dar.*

erkennt, daß aufgrund des Wissens um die Lage der Hallen die LKWs von den visuell nahezu identischen Dachfenstern der Hallen differenziert werden können.

Die Analyse des Datensatzes wird für die letzten zwei Bilder analog fortgesetzt. Für das gemäß (manipuliertem) Zeitstempel fünf Tage später aufgenommene, dritte Bild der Reihe werden die Zustände *Messeaufbau* und *Messe aktiv* prognostiziert. Die entsprechenden Hypothesen werden durch die Suchbaumknoten  $\mathcal{N}_9$  und  $\mathcal{N}_{10}$  dokumentiert. Da die Parkplätze des Geländes als gefüllt klassifiziert werden und die Fahrzeuge zwischen den Hallen verschwunden sind, läßt sich der Zustand *Messe aktiv* bestätigen. Als mögliche Folgezustände nach 10 Tagen Zeitdifferenz sind für das vierte und letzte Bild die Zustände *Messeabbau*, *Messe inaktiv* und der erneute *Messeaufbau*

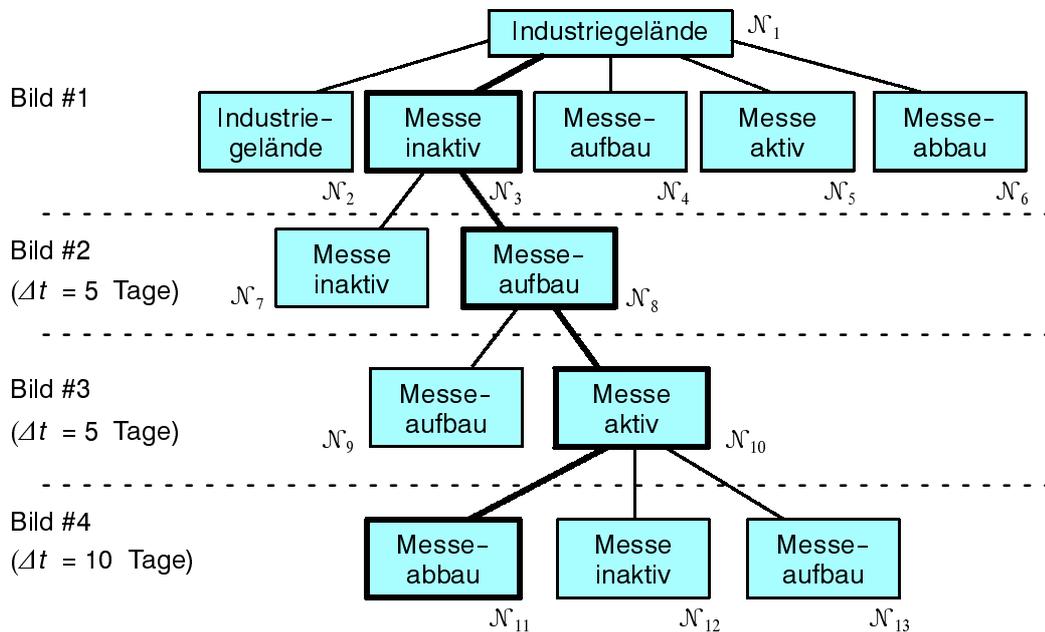


Bild 5.7: Suchbaum für die Detektion eines Messegeländes aus vier multitemporalen Luftbildern (korrekter Pfad fett gedruckt).

denkbar. Die drei Alternativen werden nacheinander untersucht bis schließlich eine Abbauphase anhand der zwischen den Hallen detektierten Fahrzeugen verifiziert werden kann. Damit wurden über die Zeitreihe hinweg der gesamte Messezyklus beobachtet, was die Instanzierung des Konzepts *Messegelände* rechtfertigt. Somit ist das Analyseziel erreicht und das Messegelände erkannt.

Insgesamt erzeugt das Interpretationssystem für das beschriebene Beispiel 13 Suchbaumknoten, die in Bild 5.7 mitsamt dem korrekten Pfad zum Analyseziel dargestellt sind. Das semantische Netz der resultierenden symbolischen Szenenbeschreibung repräsentiert sämtliche Hallen und Parkplätze der vier Bilder in 832 Instanzen. Werden die Bilder mit anderen Zeitstempeln versehen, die nicht in das Schema eines Messezyklus passen, so wird (korrekterweise) kein Messegelände instanziiert.

Das Beispiel zeigt, daß auch komplexe Muster wie ein Messegelände aus einer Reihe multitemporaler Luftbilder durch die Nutzung von temporalem Wissen automatisch extrahiert werden kann. Insbesondere die quantitativen Angaben wie Zustands- und Zustandsübergangsdauer ermöglichen, daß unmögliche Folgezustände a priori ausgeschlossen werden können. Ohne diese Informationen müßte im vorliegenden Fall jedes der Bilder auf alle vier möglichen Zustände untersucht werden, wodurch ein Suchbaum mit insgesamt 17 Suchbaumknoten entstehen würde. Darüber hinaus können die visuell

identisch erscheinenden Zustände *Messeaufbau* und *Messeabbau* nur deshalb voneinander differenziert werden, weil die absoluten Zeitangaben eine Einordnung in den Messezyklus mit seinen zeitlichen Restriktionen gestattet. Der Einsatz von temporalem Wissen erhöht demnach die Effizienz der Bildanalyse und kann dazu beitragen, daß die Szenenbeschreibung qualitativ verbessert wird.

#### 5.1.4 Vergleich der Bewertungssysteme

Im folgenden sollen die beiden vorgestellten Bewertungsverfahren des Bildinterpretationssystems vergleichend gegenübergestellt werden. Die Wahl des Bewertungssystems hat für das obige Anwendungsbeispiel keinen Einfluß auf die Erreichung der korrekten Szenenbeschreibung. Allerdings unterscheiden sich die Verfahren in der resultierenden Effizienz der Analyse:

Das **possibilistische Verfahren** selektiert immer dann per Zufall einen zu bearbeitenden Suchbaumknoten, wenn die verschiedenen Alternativen identischen diagnostischen Support beinhalten und sich nur im kausalen Support, z. B. durch eine modellgetriebene Hypothese unterscheiden. Diese Situation tritt im vorliegenden Beispiel zum einen nach der Spezialisierung des Konzepts  $I_K(\text{Industriegebiet})$  (vgl. Bild 5.5) und zum anderen nach der Schätzung von zeitlichen Nachfolgezuständen auf. Dieses Verhalten führt beim hier untersuchten Bildmaterial oftmals dazu, daß unwahrscheinliche Hypothesen vor den naheliegenden Lösungen untersucht werden. Durch den Zufälligkeitscharakter ist der Analyseverlauf nicht deterministisch und benötigt je nach gewähltem Pfad durch den Suchbaum unterschiedlich viele Inferenzschritte. In Bild 5.8 ist die akkumulierte Zahl der aktivierten Inferenzregeln für das Messebeispiel über die vier untersuchten Luftbilder aufgetragen. Die Werte mehrerer Analysedurchläufe mit possibilistischem Bewertungsverfahren sind als graue Kurvenschar erkennbar. Zur vollständigen Interpretation des ersten Bildes werden zwischen 298 und 333 Inferenzregeln aktiviert, je nachdem welcher der Suchbaumknoten  $\mathcal{N}_3$  bis  $\mathcal{N}_6$  als erster untersucht wird. Die Analyse des zweiten Luftbilds benötigt 309 Inferenzschritte, wenn der Knoten  $\mathcal{N}_8$  bevorzugt wird, ansonsten 614 Inferenzschritte. Beim dritten Bild werden entweder 469 oder 921 Regeln gefeuert, bis letztlich der Zustand *Messe aktiv* erkannt wird. Für das letzte Luftbild benötigt das System mit possibilistischer Bewertung zwischen 623 und 937 Inferenzschritte. Dadurch variiert die Gesamtzahl der aktivierten Regeln zwischen 2028 und 2766, bis das Messegelände komplett detektiert ist.

Das **probabilistische Verfahren** zeigt bei wiederholtem Einsatz ein deterministisches Verhalten (schwarzer Graph in Bild 5.8). Es folgt der Heuristik, bei gleichem

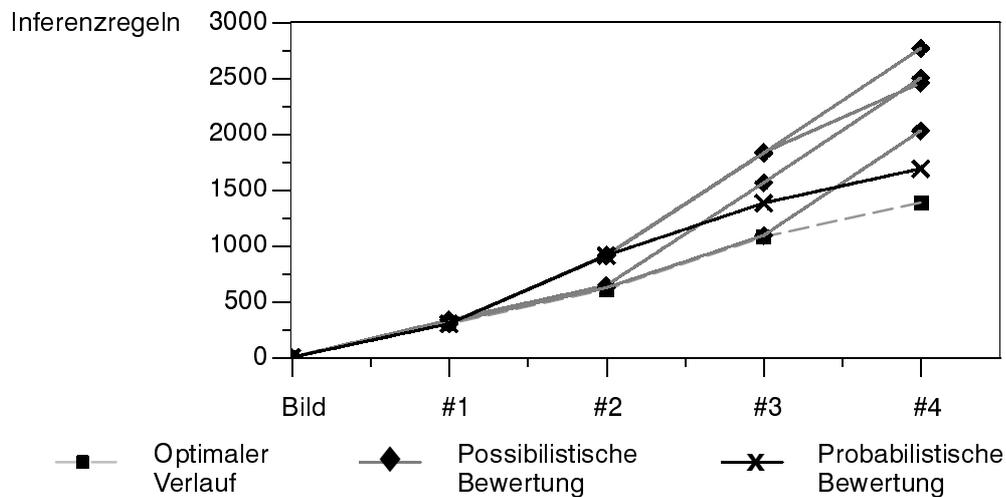


Bild 5.8: Vergleich der beiden Bewertungssysteme bezüglich der benötigten Inferenzschritte für die Detektion eines Messegeländes aus vier Luftbildern [48]

diagnostischem Support die Alternative mit größtem kausalen Support auszuwählen. Da der kausale Support in der Wissensbasis durch die Zustands- und Zustandsübergangswahrscheinlichkeiten definiert ist, ergibt sich stets derselbe Pfad durch den Suchbaum. Die Analyse des ersten Bildes erfordert konstant 298 Inferenzschritte. Nach der Spezialisierung der Instanz  $I_K(\text{Industriegebiet})$  wird stets der Knoten  $\mathcal{N}_3$  mit der wahrscheinlichsten Lösung (*Messe inaktiv*) zuerst untersucht und sofort bestätigt. Da es laut Wissensbasis naheliegender ist, daß dieser Zustand unverändert bleibt, wählt das System bei der Analyse des zweiten Bildes zunächst den (falschen) Suchbaumknoten  $\mathcal{N}_7$  aus. Nach 305 Inferenzschritten wird die Hypothese, daß es sich noch immer um ein inaktives Messegelände handelt, widerlegt und das System wechselt zum Knoten  $\mathcal{N}_8$ . Nach insgesamt 614 Regelaktivierungen kann der Zustand *Messeaufbau* verifiziert werden. In diesem Fall hat die Heuristik, die wahrscheinlichere Alternative zu bevorzugen, nicht gegriffen. Bei den beiden letzten Bildern stimmen hingegen die naheliegenden Hypothesen mit den korrekten Lösungen überein, so daß das System sofort die Knoten  $\mathcal{N}_{10}$  bzw.  $\mathcal{N}_{11}$  selektiert und damit unmittelbar dem korrekten Pfad durch den Suchbaum folgt. Es werden nur 469 bzw. 309 Inferenzschritte zur Detektion der Zustände *Messe aktiv* bzw. *Messeabbau* benötigt. Insgesamt werden 1690 Inferenzregeln bis zur Erreichung des Analyseziels aktiviert, was einer Einsparung von 17% bis 39% gegenüber der Analyse mit possibilistischer Bewertung entspricht. Da bei der Analyse des zweiten Bildes nicht sofort der korrekte Suchpfad verfolgt wurde, weicht das probabilistische Verfahren um insgesamt 305 Inferenzschritte oder um 22% vom Optimalverlauf (gestrichelte Linie in Bild 5.8) ab.

---

Die Reduktion der benötigten Inferenzschritte gegenüber dem possibilistischen Verfahren stellt eine Steigerung der Analyseeffizienz dar. Die an dieser Stelle erzielte Einsparung an aktivierten Inferenzregeln läßt sich allerdings nicht verallgemeinern. Sie wurde deswegen erreicht, weil die Strategie, die wahrscheinlichste Alternative zu bevorzugen, beim untersuchten Datensatz mehrmals zum Erfolg führte. Bei anderen Datensätzen und Anwendungsbeispielen ist es möglich, daß diese Heuristik keinen Vorteil bringt. Die Analyse mit probabilistischer Bewertung kann jedoch schlechtestenfalls so ineffizient werden wie die ungünstigste Analyse mit possibilistischer Bewertung. In solchen Fällen ist es möglich, daß das possibilistische Verfahren zufällig einen günstigeren Pfad durch den Suchbaum wählt als das probabilistische und deshalb mit weniger Inferenzschritten das Ziel erreicht.

## 5.2 Erkennung baulicher Veränderungen in Industriegebieten

Das zweite Anwendungsbeispiel aus dem Bereich der Änderungsdetektion fokussiert auf die Erkennung baulicher Veränderungen in Industriegebieten z. B. für die Aktualisierung von GIS-Daten. Es soll vor allem demonstriert werden, wie mit dem beschriebenen Bildinterpretationssystem Änderungen in der Landnutzung, beispielsweise der Abriß oder Neubau von Gebäuden, dokumentiert werden können. Zu diesem Zweck erfolgt eine Klassifikation der Bilder, indem den segmentierten Regionen durch das semantische Netz eine symbolische Bedeutung zugewiesen wird. Die Entwicklung der Klassenzugehörigkeit der einzelnen Regionen wird über die Zeit verfolgt, wodurch Bedeutungsänderungen erkennbar werden. Temporales Wissen wird dazu genutzt, Hypothesen für solche Änderungen von vornherein auszuschließen, die unmöglich während der zwischen zwei Bildern liegenden Zeitspanne aufgetreten sein können.

### 5.2.1 Datenmaterial und Wissensbasis

Als Datensatz dienen die Luftbilder des hannoverschen Messegeländes aus den Jahren 1993, 1995 und 1997. Im abgedeckten Zeitraum wurden Messehallen neu gebaut und zusätzliche Parkplätze eingerichtet. Die Bildanalyse beschränkt sich auf die Erkennung der Objektklassen *Halle*, *Parkplatz* und *Wald*. Regionen, die keiner der drei genannten Klassen angehören, werden als *Sonstiges* klassifiziert. Die Klassenzugehörigkeit einer Region wird im vorliegenden Beispiel als ein (mehr oder weniger stabiler) temporaler Zustand des Flurstücks betrachtet. Bleibt die Klassenzugehörigkeit unverändert, so liegt ein stabiler Zustand vor. Prinzipiell sind für jedes Flurstück beliebige Zustandsübergänge von einer Klasse zur einer anderen denkbar. Dennoch ist es unwahrscheinlich, daß

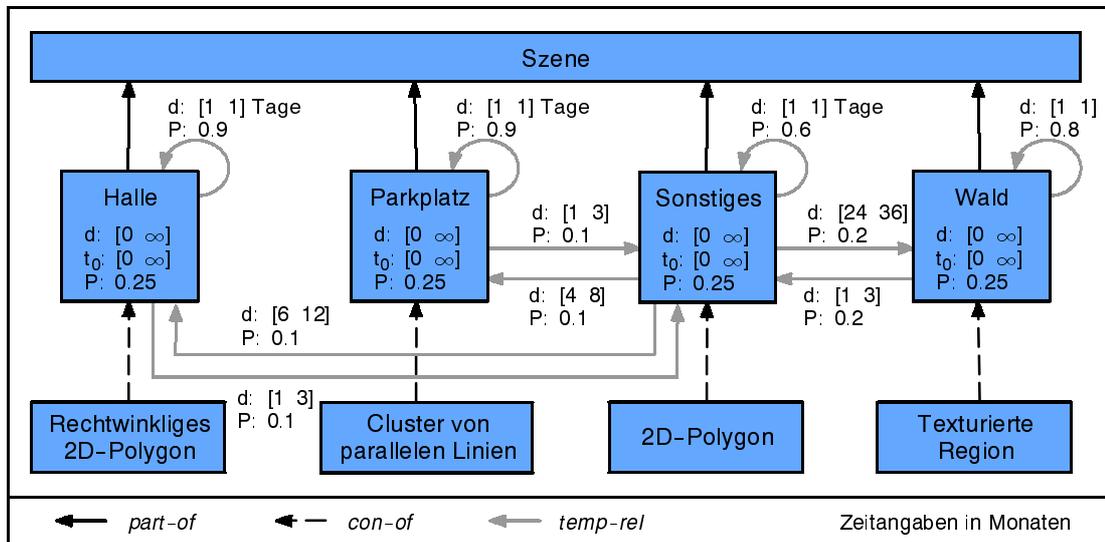


Bild 5.9: Wissensbasis für die Erkennung der Objektklassen Halle, Parkplatz und Wald sowie ihrer zeitlichen Veränderungen.

z. B. dort, wo eine Halle detektiert wurde, kurze Zeit später ein Wald zu finden ist. Dieses temporale Wissen kann wie in dem vorigen Anwendungsbeispiel in einem semantischen Netz repräsentiert und für die Bildauswertung genutzt werden.

Bild 5.9 zeigt die verwendete Wissensbasis, in der alle Zeitangaben in Monaten angegeben sind. Die *Szene* wird aus beliebig vielen Instanzen der vier genannten Klassen zusammengesetzt. Jede Klasse für sich stellt einen stabilen Zustand dar, was durch eine unendliche Zustandsdauer  $[0 \infty]$  modelliert wird. Sämtliche Zustandsübergänge sind über das Konzept *Sonstiges* geführt, da jede hier betrachtete Landschaftsveränderung mit einer baulichen Maßnahme verbunden ist. Während dieser Phasen ist das betreffende Flurstück nicht als Halle, Parkplatz oder Wald klassifizierbar. Ist der Zeitabstand zwischen zwei aufeinanderfolgenden Bildern groß genug, so erzeugt das System mit Hilfe der beschriebenen temporalen Inferenz (s. Kapitel 3.3.2) auch Hypothesen für den direkten Übergang zwischen zwei Klassen, wie z. B. von *Halle* nach *Parkplatz*. Somit sind je nach Zeitabstand alle Kombinationen von Zustandsübergängen abgedeckt. Die Zustandsübergangsdauern variieren zwischen drei und 24 Monaten. Die rückgekoppelten Übergänge sind hingegen mit einer Dauer von einem Tag angegeben. Bei der Angabe der Zustandsübergangswahrscheinlichkeiten ist kein empirisches Wissen, sondern nur die allgemeine Erfahrung eingeflossen, daß ein Zustand eher unverändert bleibt, als daß er sich ändert. Aus der Symmetrie des Zustandsübergangsdiagramms ergeben sich daraus gleichwahrscheinliche Zustände mit  $P(S_i) = 0.25$ .

---

Die Klassen *Halle* und *Parkplatz* werden wie in Kapitel 5.1.1 beschrieben als *rechtwinklige 2D-Polygone* bzw. als *Cluster paralleler Linien* in den Bildern erkannt. Hier kommen die identischen Bildverarbeitungsverfahren wie in der zuvor beschriebenen Anwendung zum Einsatz. Die Klasse *Wald* prägt sich im Bild als *texturierte Region* mit hohem Grünanteil aus. Zur Segmentierung wird ein Texturanalyse-Verfahren basierend auf Gibbs-Zufallsfeldern [43][44] eingesetzt. Alle Segmentierungsverfahren liefern Kandidaten für ihre jeweilige Objektklasse in Form von Polygonen, die die umschließende Kontur der segmentierten Bildregion approximieren. Da die Verfahren unabhängig voneinander arbeiten, kann es sein, daß eine Bildregion beispielsweise sowohl als Hallen- als auch als Waldkandidat segmentiert wird. Die Auflösung dieser Widersprüche ist die Aufgabe des Interpretationssystems.

## 5.2.2 Interpretationsprozeß

Bevor bauliche Veränderungen detektiert werden können, muß eine initiale Szenenbeschreibung erzeugt werden. Diese kann prinzipiell aus bereits vorhandenen GIS-Daten abgeleitet werden. Im vorliegenden Fall wird sie aus dem ersten Bild des Datensatzes (hier: aus dem Jahr 1993) erstellt. Analog zum vorigen Beispiel werden dazu alle akzeptablen Hallen, Parkplätze und Wälder aus den segmentierten Kandidaten ausgewählt und im semantischen Netz als Instanzen repräsentiert.

Anschließend beginnt die Analyse des nachfolgenden Bildes (hier: 1995). Für jedes im ersten Bild detektierte Objekt erzeugt das System über die temporale Inferenz Hypothesen dafür, welchen Zustand das korrespondierende Flurstück im zweiten Bild des Datensatzes haben könnte. Dadurch entstehen ggf. mehrere konkurrierende Lösungen. Beispielsweise kann eine Instanz der Klasse *Halle* bei der hier vorliegenden Zeitdifferenz von zwei Jahren entweder unverändert bleiben oder aber zu den Klassen *Parkplatz* oder *Sonstiges* übergehen. Ein Übergang zur Klasse *Wald* nimmt laut Wissensbasis eine längere Zeit in Anspruch.

Die über die temporale Inferenz erzeugten Hypothesen beinhalten für das jeweils betrachtete Flurstück eine erwartete Klassenzugehörigkeit, die es in den Daten zu überprüfen gilt. Dabei ist zu beachten, daß die im aktuellen Bild segmentierten Regionen sich sowohl in Form als auch in der Klassenzugehörigkeit von den aus dem vorangegangenen Bild abgeleiteten Erwartungen unterscheiden können. Folgende Fälle lassen sich differenzieren und werden während der Interpretation berücksichtigt:

- Eine Bildregion mit beliebiger Klassenzugehörigkeit aus Zeitpunkt  $t_k$  entspricht genau einer segmentierten Bildregion mit beliebiger Klassenzugehörigkeit zum Zeitpunkt  $t_{k+1}$ . Dabei sind geringe Abweichungen in Geometrie und Lage zu

---

tolerieren, da stets mit Bild- und somit mit Segmentierungsunterschieden zu rechnen ist. Gehören die Regionen unterschiedlichen Klassen an, so liegt eine Landnutzungsänderung vor, ansonsten ist die Nutzung unverändert geblieben.

- Eine Bildregion mit beliebiger Klassenzugehörigkeit aus Zeitpunkt  $t_k$  entspricht keiner segmentierten Bildregion zum Zeitpunkt  $t_{k+1}$ , da sie im Folgebild nicht als Halle, Parkplatz oder Wald klassifiziert werden konnte. Die Region wird deshalb der Klasse *Sonstiges* zugeordnet.
- Eine Bildregion mit beliebiger Klassenzugehörigkeit aus Zeitpunkt  $t_k$  wird zum Zeitpunkt  $t_{k+1}$  durch mehrere segmentierte Bildregionen einer einzigen oder mehrerer unterschiedlicher Klassen abgedeckt. Es liegt nur dann eine Landnutzungsänderung vor, wenn zumindest eine Teilregion in  $t_{k+1}$  einer anderen Klasse angehört als zum Zeitpunkt  $t_k$ . Handelt es sich zu beiden Zeitpunkten um Instanzen derselben Klasse, so hat lediglich das Segmentierungsverfahren ein zerstückeltes Ergebnis als für das vorherige Bild geliefert. Die Nutzung des gesamten Flurstücks hat sich jedoch nicht verändert.
- Die Umkehrung des zuvor genannten Falls ist gegeben, wenn mehrere Regionen mit beliebiger Klassenzugehörigkeit aus Zeitpunkt  $t_k$  durch eine Bildregion aus  $t_{k+1}$  abgedeckt wird. Auch hier handelt es sich nur dann um eine Landnutzungsänderung, wenn die Regionen zu unterschiedlichen Klassen gehören.
- Schließlich kann der Fall auftreten, daß in Regionen, die im Bild vom Zeitpunkt  $t_k$  nicht segmentiert wurden, im Bild aus  $t_{k+1}$  als relevantes Objekt detektiert werden. Diese neu hinzukommenden Objekte werden zusätzlich als Instanz in der Szenenbeschreibung aufgenommen.

Das System prüft nach und nach alle Hypothesen für mögliche temporale Nachfolger der im vorherigen Bild erkannten Objekte und klassifiziert sie entsprechend der Segmentierungsergebnisse des aktuellen Bildes. Liegt für ein Flurstück kein Segmentierungsergebnis vor, erfolgt die Zuordnung zur Klasse *Sonstiges*. Anschließend werden die noch nicht interpretierten Flächen des aktuellen Bildes in die Szenenbeschreibung aufgenommen. Die resultierende Szenenbeschreibung für das aktuelle Bild dient wiederum als Grundlage für die Interpretation des nachfolgenden Bildes. Dieser Vorgang wird solange wiederholt, bis das letzte Bild der Reihe interpretiert wurde.

Bild 5.10 zeigt beispielhaft das Ergebnis der Änderungsdetektion für einige Messehallen. Im Bild von 1993 werden in dem betrachteten Bildausschnitt korrekterweise vier Hallen erkannt. Für jede Halle werden anschließend die jeweils möglichen Folgezustände *Halle*, *Parkplatz* und *Sonstiges* prognostiziert und im Bild aus dem Jahre

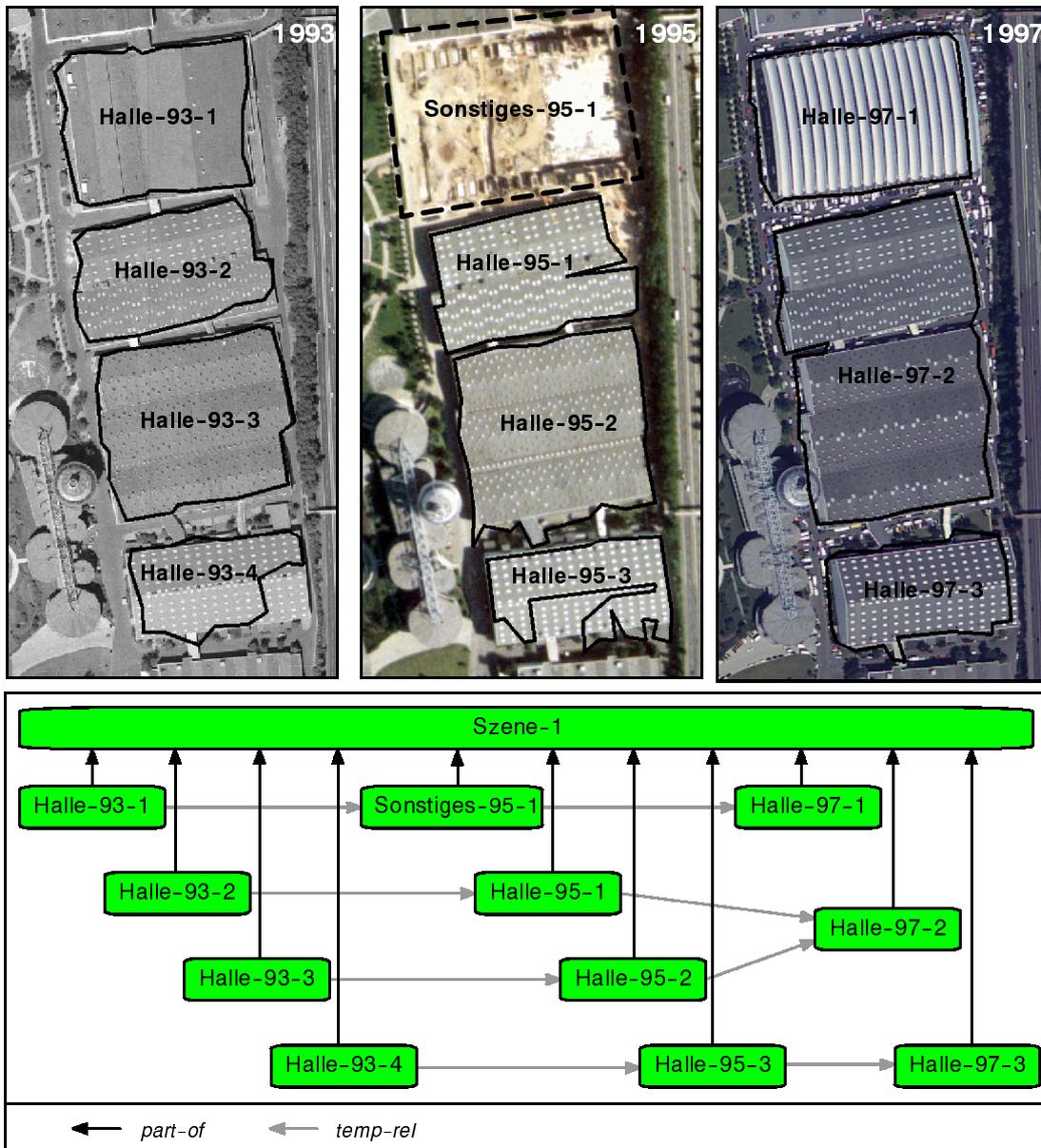


Bild 5.10: Veränderungen von Messehallen zwischen 1993 und 1997 und deren Repräsentation als Instanzen im semantischen Netz.

1995 verifiziert. Drei Hallen werden als solche wiedererkannt, die vierte Halle wurde jedoch abgerissen. Deshalb wird dem entsprechenden Flurstück die Klasse *Sonstiges* zugewiesen. Im letzten Bild ist die Halle neu errichtet und wird auch korrekt erkannt. Hier hat über den Beobachtungszeitraum hinweg eine Landnutzungsänderung stattgefunden, die auch im semantischen Netz dokumentiert ist. Im Segmentierungsergebnis

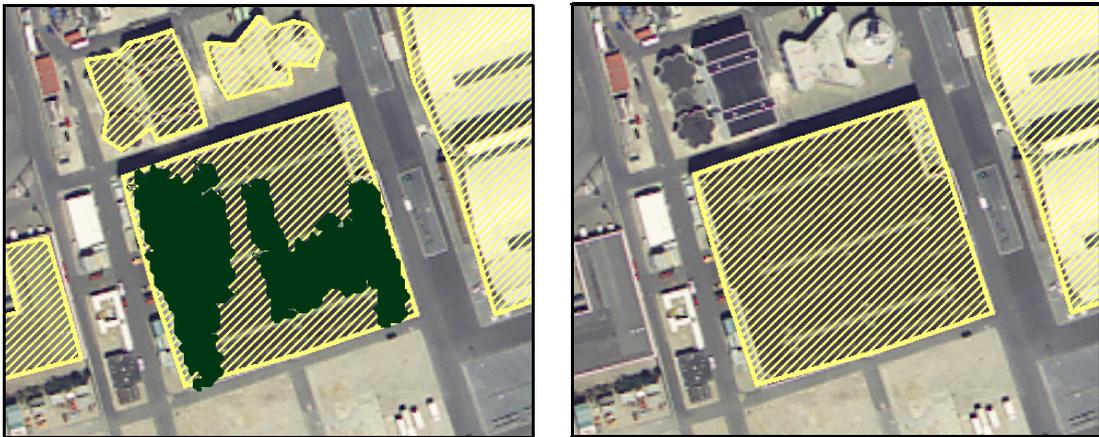


Bild 5.11: *Widersprüchliche Segmentierungsergebnisse für das Luftbild von 1997: Ein Bildbereich wird sowohl als Wald- (gefüllt) als auch als Hallenkandidat (schraffiert) segmentiert (links). Aufgrund des temporalen Vorwissens und der Szenenbeschreibung für 1995 wird die Region korrekt als Halle klassifiziert (rechts).*

für das Jahr 1997 werden fälschlicherweise zwei Hallen zu einer verschmolzen. Diese Verschmelzung wird im semantischen Netz nachvollzogen, indem die Instanz *Halle-97-2* als gemeinsamer Nachfolger der Instanzen *Halle-95-1* und *-2* repräsentiert wird. Gemäß der Segmentierungsergebnisse wird somit eine korrekte Szenenbeschreibung erzeugt.

Das temporale Wissen wird während der Analyse genutzt, um unmögliche Folgezustände für die einzelnen Flurstücke auszuschließen. Bild 5.11 zeigt einen Bildausschnitt aus dem Jahre 1997, in dem eine Region sowohl als potentieller Wald als auch als Halle segmentiert wurde. Da die entsprechende Region im Jahr 1995 als Halle klassifiziert worden war, kann sie gemäß Wissensbasis innerhalb von zwei Jahren nicht bewaldet sein. Das System klassifiziert die Region deshalb korrekterweise als Halle. Einige benachbarte Hallenkandidaten werden wegen zu geringer Größe abgewiesen.

Insgesamt werden in den Bildern 41 Hallen, 25 Parkplätze, 159 Wälder und 83 Objekte der Klasse *Sonstiges* erkannt. Die zeitlichen Veränderungen können korrekt dokumentiert werden, sofern die Segmentierungsergebnisse dieses gestatten. Das System erzeugt dabei einen Suchbaum bestehend aus 651 Suchbaumknoten. Eine detailliertere Darstellung der Ergebnisse ist in [12] zu finden.

Der vorgestellte Ansatz zur Änderungsdetektion läßt sich in die Klasse der Klassifikationsvergleiche (s. Kapitel 1.2) einordnen. Er unterscheidet sich jedoch darin, daß die Klassifikation eines Bildes von der Interpretation des vorhergehenden Bildes

abhängig gemacht wird. Unrealistische Landschaftsveränderungen können durch das temporale Wissen a priori ausgeschlossen werden, wodurch die Klassifikationsgüte erhöht wird. Der Ansatz ist prinzipiell für die Aktualisierung von GIS-Daten geeignet. Denn das System dokumentiert zum einen Änderungen in der Landnutzung, indem die Flurstücke unterschiedlichen Klassen zugeordnet werden. Zum anderen werden auch rein geometrische Veränderungen in den jeweiligen Attributwerten der Objekte erfaßt. Letzteres setzt allerdings eine möglichst exakte Segmentierung der Bilder voraus.

### 5.2.3 Vergleich der Bewertungssysteme

Auch die beschriebene Änderungsdetektion kann wahlweise mit dem possibilistischen oder dem probabilistischen Bewertungssystem durchgeführt werden. Beide Verfahren führen letztlich zu identischen Szenenbeschreibungen. Allerdings nutzt das auf Bayes-Netzen basierende Verfahren die Zustandsübergangswahrscheinlichkeiten des temporalen Wissens aus, so daß bei der Prädiktion von Folgezuständen stets die wahrscheinlichste Alternative, hier die Hypothese der unveränderten Szene, als erste untersucht wird.

Die folgende Tabelle zeigt, daß aus dem genannten Grund das probabilistische Verfahren für diesen Datensatz weniger Inferenzschritte bis zum Analyseziel benötigt. Während hier stets 2306 Regeln aktiviert werden, sind beim probabilistischen Ansatz bis zu 2831 Schritte notwendig, was einer Differenz von 19% entspricht.

Tabelle 5.1 *Kumulierte Zahl der aktivierten Inferenzregeln bei Verwendung des possibilistischen bzw. des probabilistischen Bewertungssystems für die Detektion baulicher Veränderungen auf dem Messegelände Hannover von 1993 bis 1997.*

Kumulierte Zahl der Inferenzregeln	Bild 1 (1993)	Bild 2 (1995)	Bild 3 (1997)
Possibility	474	1411-1496	2750-2831
Bayes-Netz	474	1252	2306

## 5.3 Diskussion der Ergebnisse

Bei der wissensbasierten Bildinterpretation wurden bislang strukturelle, topologische, objekt- und sensorspezifische Informationen sowie Karten- bzw. GIS-Daten als Wissensquelle genutzt. Das System AIDA bietet die Möglichkeit, alle genannten Informationen strukturiert in einem semantischen Netz zu repräsentieren und automatisch für die Erstellung einer symbolischen Szenenbeschreibung auszuwerten. Die

---

vorliegende Arbeit erweitert das Spektrum der Wissensquellen erstmals um temporale Informationen, so daß auch zeitlich nacheinander aufgenommene Bilder automatisch interpretiert werden können.

Die beiden Anwendungsbeispiele zeigen, daß eine wissensbasierte Interpretation multitemporaler Luftbilder möglich ist und dabei das explizit formulierte temporale Wissen vorteilhaft genutzt werden kann. Im Gegensatz zu vielen aus der Literatur bekannten Ansätzen zur Auswertung multitemporaler Bilder erfolgt hier die Detektion und Deutung der zeitlichen Veränderungen nicht auf Bildpunktebene, sondern auf einer semantischen Ebene. Erst durch die Bedeutungszuweisung wird die Einbringung des temporalen Vorwissens in den Analyseprozeß möglich. Auf dem Gebiet der Änderungsdetektion können dadurch unrealistische Landschaftsveränderungen von vornherein ausgeschlossen werden, da die Landnutzung eines Flurstücks zum vorhergehenden Beobachtungszeitpunkt bekannt ist. Darüber hinaus lassen sich komplexe Objekte, wie beispielsweise das Messegelände aus Kapitel 5.1, automatisch erkennen, indem nach einer charakteristischen Abfolge von Ereignissen über den gesamten Beobachtungszeitraum gesucht wird.

Die gewählte Repräsentationsform für das temporale Wissen erscheint für eine automatische Auswertung geeignet. Sie beschränkt sich auf diskrete Objektzustände und Übergänge zwischen diesen. Dadurch ist die Integration in das semantische Netz leicht möglich. Die zeitlichen Angaben über Zustands- und Zustandsübergangsdauern werden für die optimistische Prognose von Objektveränderungen genutzt. Durch die Elimination unmöglicher Landschaftsänderungen wird der Suchraum so klein wie möglich gehalten. Diese Vorgehensweise bewirkt einen effizienteren Analyseverlauf.

Die in der Wissensbasis abgelegten Auftretenswahrscheinlichkeiten für Objektzustände und -zustandsübergänge nutzt das auf Bayes-Netzen basierende Bewertungsverfahren. Dieses sorgt dafür, daß wahrscheinlichere Lösungsalternativen bevorzugt untersucht werden, was in vielen Fällen zu einer zusätzlichen Effizienzsteigerung führt. In den beschriebenen Beispielen konnte die Zahl der Inferenzschritte um bis zu 39% gegenüber dem bisherigen Bewertungsverfahren reduziert werden (vgl. Kapitel 5.1.4). Das Einsparungspotential dieser Heuristik ist umso größer, je mehr Inferenzschritte zur Widerlegung einer falschen Lösung benötigt werden. Können falsche Alternativen nach wenigen Inferenzschritten von der richtigen unterschieden werden, ist der Gewinn eher unbedeutend. Bei der temporalen Inferenz werden jedoch auf hoher Ebene im semantischen Netz Hypothesen erzeugt, deren Verifikation sehr umfangreich sein kann. So erfordert beispielsweise die Vermutung, daß es sich bei dem Beobachtungsgebiet um ein Messegelände während einer bestimmten Phase handelt, die Untersuchung des

---

gesamten Bildes mit allen Hallen und Parkplätzen. Die Wahl eines falschen Pfades im Suchbaum wird schnell mit mehreren hundert vergeblich aktivierten Inferenzregeln bestraft. Somit erscheint der probabilistische Bewertungsansatz gerade für die multitemporale Bildanalyse gut geeignet zu sein. Für den Fall, daß die wahrscheinlichste Alternative auch die korrekte ist, erreicht das System das Analyseziel auf direktem Weg. Ansonsten werden maximal gleich viele Inferenzregeln aktiviert, wie das possibilistische Verfahren schlechtestenfalls zur Erreichung des Ziels benötigt.

Durch den Einsatz von Bayes-Netzen für die Bewertung von Szenenbeschreibungen wird es erstmals möglich, sowohl den kausalen als auch den diagnostischen Support für eine Hypothese zu berücksichtigen, da der Informationsfluß sowohl *bottom-up* als auch *top-down* im Netzwerk erfolgt. Die berechneten Glaubwürdigkeitswerte für die einzelnen Instanzen des semantischen Netzes stellen ein Maß für die Kompatibilität von Erwartungen und Messungen dar. Diese werden bislang ausschließlich zur Ableitung eines Bewertungsmaßes für die gesamte Szenenbeschreibung genutzt. Sie lassen sich jedoch darüber hinaus als Qualitätsmaß des korrespondierenden Szenenobjekts z. B. im Rahmen einer Verifikation von GIS-Daten verwenden. Die zu prüfenden GIS-Daten repräsentieren die Erwartungen des Systems, die durch die Bewertungsmethoden mit den aktuellen Messungen im Bild verglichen werden. Am Ende eines Bewertungsvorgangs kann für jedes Szenenobjekt eine Aussage über den Grad der Übereinstimmung zwischen GIS- und Bilddaten gemacht werden. Da jedoch sowohl die GIS-Daten als auch die Segmentierungsergebnisse als unsicher und ungenau anzusehen sind, muß eine abschließende qualitative Beurteilung der zu prüfenden GIS-Daten mit besonderer Vorsicht erfolgen. Diesbezüglich bedarf es gesonderter Untersuchungen.

Die Formulierung von temporalen Wissensinhalten für eine multitemporale Bildinterpretation bleibt einem Experten vorbehalten, der sich entweder auf seine Erfahrungen oder gegebenenfalls auf empirische Messungen stützt. Ein automatisches Erlernen der relevanten zeitlichen Informationen erscheint zum heutigen Stand der Forschung unrealistisch. Nur ein Experte kann relevante Objektzustände oder Ereignisse identifizieren, die sich charakteristisch in den Luftbildern ausprägen und sich somit automatisch detektieren lassen. Die absoluten und relativen Zeitangaben dürfen allerdings mit einer Unsicherheit behaftet sein, da das Interpretationssystem in der Lage ist, Zeitintervalle zu verarbeiten. Auch die Angabe von Auftretenswahrscheinlichkeiten bedarf keiner hohen Genauigkeit, da sie über das Bewertungssystem lediglich zur Steuerung der Bearbeitungsreihenfolge genutzt werden. Somit reicht oftmals die Repräsentation hinlänglich bekannter Erfahrungen aus, um eine Effizienzsteigerung bei der multitemporalen Luftbildinterpretation zu erzielen.

---

## 6. Zusammenfassung

Die vorliegende Arbeit hat die automatische Interpretation von Luftbildern zum Thema. Insbesondere behandelt sie die Analyse multitemporaler Bilder, d. h. solcher Aufnahmen, die ein Beobachtungsgebiet zu unterschiedlichen Zeitpunkten abbilden. Unter Ausnutzung von strukturellen, topologischen und vor allem temporalen Wissensinhalten wird automatisch eine symbolische Beschreibung der Szene für den gesamten Beobachtungszeitraum erstellt. Dadurch können zum einen bauliche Veränderungen erkannt und erklärt werden. Zum anderen lassen sich komplexe Muster, die sich durch eine zeitliche Abfolge charakteristischer Ereignisse auszeichnen, automatisch detektieren. Die resultierende symbolische Szenenbeschreibung kann zur Erfassung von Geoinformationen genutzt werden, da sie sowohl die Semantik als auch die geographische Lage der Szenenobjekte beinhaltet.

Bislang werden Luftbilder zumeist manuell von menschlichen Operateuren ausgewertet. Eine Automatisierung wird durch die Komplexität der Szene und die Vielfältigkeit der möglichen Analyseziele erschwert. Besonders die Interpretation multitemporaler Bildreihen erfordert in einem hohen Maße Expertenwissen, um relevante Veränderungen in der Landschaft von irrelevanten zu unterscheiden oder um aus einer Reihe von beobachteten Ereignissen auf ein dadurch charakterisiertes Objekt zu schließen. Bisherige automatische Verfahren zur Luftbildauswertung beschränken sich entweder auf eine bildpunktbasierende Änderungsdetektion, ohne die zeitlichen Veränderungen zu deuten, oder sind lediglich in der Lage, monotemporale Bilder zu interpretieren. Kernidee der vorliegenden Arbeit ist es, Vorwissen über zeitliche Zusammenhänge für die automatische Auswertung multitemporaler Luftbilder nutzbar zu machen. Durch die Interpretation der gesamten Szene wird es möglich, eine Änderungsdetektion nicht allein aufgrund veränderter Bildpunkte, sondern aufgrund veränderter Objektbedeutungen durchzuführen. Einzelbeobachtungen können in einen zeitlichen Kontext gestellt und übergreifend gedeutet werden.

Der vorgestellte Ansatz basiert auf dem Bildinterpretationssystem AIDA [152], welches explizit formuliertes Wissen über die Struktur, die Geometrie, das Material und die Topologie der erwarteten Szenenobjekte für die Bildinterpretation nutzt. Als Wissensrepräsentationsform kommen semantische Netze zum Einsatz, in denen die relevanten Szenenobjekte mitsamt ihrer Attribute als Knoten und deren wechselseitigen Beziehungen als Kanten modelliert werden. Eine allgemeingültige Netzwerksprache ermöglicht die Definition von Klassen- und Bestandteilshierarchien, von topologischen Beziehungen sowie die Unterteilung der Wissensinhalte in konzeptionelle Ebenen wie z. B. eine symbolische 3D-Szenen- und eine bildbezogene 2D-Sensorebene. Dadurch

---

können Landschaftsobjekte wie Straßen und Flüsse aber auch abstrakte Muster wie z. B. ein Klärwerk beschrieben werden. Mit Hilfe der Wissensbasis versucht das System, den segmentierten Bildprimitiven eine Semantik zuzuweisen und sie zu immer komplexeren Strukturen zu gruppieren, um so schrittweise eine symbolische Beschreibung der gesamten Szene in Form eines instanziierten semantischen Netzes aufzubauen. Erwartungen und Zwischenergebnisse werden genutzt, um gezielt Hypothesen bezüglich Lage und Aussehen von gesuchten Objekten zu generieren. Das System ist in der Lage, multisensorielle Datensätze auszuwerten und dabei gegebenenfalls die szenenspezifischen Informationen eines Geoinformationssystems zu nutzen.

Um multitemporale Bildreihen analysieren zu können, müssen über die bislang verwendeten Wissensinhalte hinaus zeitliche Aspekte berücksichtigt werden. Bei der sequentiellen Auswertung der Bilder sollen die Erkenntnisse aus den vorherigen, bereits interpretierten Aufnahmen genutzt werden. Dabei soll nicht ausschließlich angenommen werden, daß die Objekte über die Zeit unverändert bleiben, sondern es wird eine gezielte Prädiktion von Landschaftsveränderungen angestrebt. Für eine derartige Vorgehensweise ist neben der Information über die verstrichene Zeit zwischen aufeinanderfolgenden Bildern zusätzliches Wissen über zeitliche Zusammenhänge erforderlich. Diese temporalen Wissensinhalte müssen derart repräsentiert werden, daß sie während der Analyse automatisch für eine Prädiktion genutzt werden können.

Da die einzelnen Luftbilder jeweils Momentaufnahmen einer dynamischen Welt darstellen, werden die Landschaftsveränderungen durch Übergänge zwischen diskreten Objektzuständen oder -klassen modelliert. Aufgrund der strukturellen Ähnlichkeit zu semantischen Netzen eignen sich Zustandsübergangsdiagramme als Repräsentationsform für das temporale Wissen. Die Zustände werden durch Knoten des semantischen Netzes dargestellt, die zusätzlich Informationen über Zustandsdauer und Auftretenswahrscheinlichkeit beinhalten. Eine neu eingeführte temporale Kante im semantischen Netz beschreibt die möglichen Zustandsübergänge, d. h. die zeitlichen Veränderungen, mitsamt der Übergangsdauer und Wahrscheinlichkeit. Durch die Definition von Schleifen im Zustandsübergangsdiagramm ist auch der Fall abgedeckt, daß ein Zustand sein eigener Nachfolger ist und somit über die Zeit unverändert bleibt. Unsichere Zeitangaben werden in Form von Intervallen beschrieben. Die gewählte Repräsentationsform vereint Eigenschaften von Zustandsübergangsdiagrammen, Markov-Ketten und Netzplänen, da sie neben der Definition von Zuständen auch explizite Wahrscheinlichkeits- und Zeitangaben zuläßt.

Nach vollständiger Interpretation eines Bildes der Zeitreihe wird das temporale Wissen genutzt, um Landschaftsveränderungen für das nächste Bild vorherzusagen. Für

---

die einzelnen Szenenobjekte werden alle Zustandsübergänge bestimmt, die laut Wissensbasis in der Zeitspanne bis zum nächsten Aufnahmezeitpunkt auftreten können. Die temporale Inferenz gewährleistet durch eine optimistische Vorgehensweise, daß alle denkbaren Zustandsübergänge erfaßt werden. Unmögliche Zustandsübergänge lassen sich jedoch a priori ausschließen, wodurch der Suchraum verkleinert und die Analyse effizienter wird.

Für jeden ermittelten Zustandsübergang werden Hypothesen im semantischen Netz erzeugt, die es in den Bilddaten zu verifizieren gilt. Die einzelnen Hypothesen sind als konkurrierende Alternativen zu betrachten und werden deshalb getrennt voneinander repräsentiert. Da die Analyse nur sequentiell erfolgen kann, werden mit Hilfe eines Bewertungssystems die verschiedenen Alternativen hinsichtlich ihrer Qualität beurteilt, um die vielversprechendste von ihnen bevorzugt bearbeiten zu können.

Die bisherigen Ansätze zur Bewertung konkurrierender Szenenbeschreibungen basieren auf dem Vergleich von den konkreten Messungen in den Bilddaten mit den aus dem Vorwissen abgeleiteten, erwarteten Meßwerten. Die beschriebene Prädiktion von zeitlichen Veränderungen erfolgt jedoch ausschließlich auf Basis des temporalen Wissens. Die neu erzeugten Hypothesen sind zum Zeitpunkt der Bewertung noch nicht in den Daten verifiziert worden, so daß für alle Alternativen gleich viele Messungen vorliegen. Bewertungsverfahren, die ausschließlich datengetrieben vorgehen, liefern deshalb identische Beurteilungen für die verschiedenen Lösungen. Das Analysesystem wählt somit zufällig einen der Kandidaten für die weitere Bearbeitung aus. Dadurch ist es oftmals der Fall, daß zunächst abwegige Hypothesen verfolgt und widerlegt werden, bevor sich das System den naheliegenden Alternativen zuwendet. Dieser unerwünschte Effekt ist besonders gravierend, wenn die Falsifizierung einer Hypothese umfangreiche Untersuchungen erfordert, was besonders bei der temporalen Inferenz häufig vorkommt. Zur Vermeidung dieses Effekts ist eine Strategie "vom Wahrscheinlichen zum Unwahrscheinlichen" anzustreben, so daß die wahrscheinlicheren Lösungen zuerst untersucht werden.

Zu diesem Zweck wurde ein neuartiges Bewertungsverfahren entwickelt, das sowohl die Hinweise aus den Daten als auch a priori bekannte Auftretenswahrscheinlichkeiten (insbesondere von Objektzuständen und Zustandsübergängen) bei der Beurteilung einer Szenenbeschreibung berücksichtigt. Dazu wird das semantische Netz auf ein Bayes-Netz abgebildet. In Bayes-Netzen lassen sich kausale Abhängigkeiten von Ereignissen durch die Kanten eines Graphen modellieren und durch bedingte Wahrscheinlichkeiten quantifizieren. Die Messungen aus den Daten dienen als diagnostische und die Auftretenswahrscheinlichkeiten der Objekte als kausale Unter-

---

stützung für die zu bewertende Szenenbeschreibung. Mit Hilfe eines Propagierungsverfahrens nach Pearl [107] werden Glaubwürdigkeiten für die einzelnen Szenenobjekte berechnet. Dabei werden die Unsicherheit der Daten und die o.g. Wahrscheinlichkeiten berücksichtigt. Aus einer Menge von konkurrierenden Szenenbeschreibungen mit identischer diagnostischer Unterstützung wird deshalb diejenige mit der höchsten Auftretenswahrscheinlichkeit am besten bewertet. Die zufällige Auswahl einer Alternative kann so vermieden werden.

Der vorgestellte Ansatz bietet somit erstmals die Möglichkeit, multitemporale Luftbilder automatisch zu analysieren. Der wissensbasierte Ansatz erleichtert die flexible Anpassung des Systems an variierende Aufgabenstellungen. Temporale Wissensinhalte lassen sich in Form eines Zustandsübergangsdiagramms in das semantische Netz integrieren und während der Analyse für eine Prädiktion von Landschaftsveränderungen nutzen. Das neuartige Bewertungsverfahren ermöglicht, unter vergleichbaren Alternativen stets die wahrscheinlichste für die weitere Analyse auszuwählen.

Die Funktionalität des Bildinterpretationssystems wurde anhand zweier Anwendungsbeispiele nachgewiesen. Aus einer Reihe von multitemporalen Luftbildern konnten die baulichen Veränderungen in einem Industriegebiet automatisch detektiert werden. Dabei wird der Abriß und der Neubau von Hallen und Parkplätzen korrekt in einem semantischen Netz dokumentiert. In einem zweiten Beispiel wurde gezeigt, daß sich komplexe Objektklassen wie ein Messegelände anhand eines typischen zeitlichen Verhaltens aus einer Bildreihe erkennen lassen. Ein Messegelände ist durch den Zyklus Messeaufbau, Messezeit, Messeabbau und Inaktivität charakterisiert. Für die einzelnen Phasen lassen sich sowohl die Dauer als auch die Auftretenswahrscheinlichkeiten abschätzen. Die Auf- und Abbauarbeiten können anhand von neben den Hallen parkenden LKWs der Messebauer erkannt werden, während gefüllte Parkplätze darauf hindeuten, daß eine Messe stattfindet. Aus einer geeigneten Bildreihe konnten die entsprechenden Ereignisse innerhalb des erwarteten Zeitrahmens detektiert und somit die Existenz eines Messegeländes bestätigt werden. Das neuartige Bewertungssystem bewirkt in beiden Beispielen eine Steigerung der Analyseeffizienz. Im Vergleich zu einem existierenden, rein datengetriebenen Bewertungsansatz wird die Zahl der benötigten Inferenzschritte bis zum Analyseziel deutlich reduziert.

Die vorliegende Arbeit zeigt auf, daß sich Vorwissen über zeitliche Zusammenhänge gewinnbringend für die automatische Analyse multitemporaler Luftbilder nutzen läßt. Sie liefert somit einen Beitrag zur automatisierten Landschaftsüberwachung sowie zur Erfassung und Aktualisierung von Geo-Daten und Karten.

---

## Anhang

### A-1 Approximation von Polygonen

Die Verarbeitung von Bildprimitiven innerhalb des Bildinterpretationssystems erfolgt auf Basis vektorisierter Daten, d.h. Liniensegmente werden beispielsweise durch einen Start- und Endpunkt und Regionen durch die Stützpunkte eines die Regionenkontur beschreibenden Polygons definiert. Da die Bildverarbeitungsverfahren jedoch mit Rasterdaten arbeiten, müssen die Segmentierungsergebnisse einer Raster-zu-Vektor-Umformung unterzogen werden.

Das Verfahren zur Segmentierung von Hallen extrahiert mit Hilfe adaptiver Schwellwerte aus Höhendaten erhabene Regionen, die aufgrund des Höhensprungs als Kandidaten für Gebäude in Frage kommen. Um eine erste vektorielle Beschreibung einer segmentierten Region zu erhalten, wird ein Polygon erzeugt, das für jeden Bildpunkt am Außenrand der Region einen Stützpunkt enthält. Dadurch entsteht zwar eine formerhaltende polygonale Beschreibung der Bildregion, die allerdings oftmals aus unnötig vielen Stützpunkten besteht (s. Bild A-1.2a). Aus diesem Grund wird ein Verfahren zur Nachbearbeitung der Polygone eingesetzt, das eine Approximation der Region mit möglichst wenigen Stützpunkten liefert. Hierbei muß ein Kompromiß zwischen Minimierung der Stützpunktzahl und möglichst guter Formerhaltung gefunden werden.

Die Reduzierung der Stützpunktzahl eines Polygons basiert auf einem Verfahren zur Detektion der kritischen Punkte einer Kontur [168]. Als kritische Punkte werden Orte maximaler Krümmung entlang einer Kontur bezeichnet. Sie liefern einen hohen Informationsgehalt bezüglich der Form der Region und sind deshalb zu erhalten, während Konturpunkte an Stellen geringer Krümmung unter geringem Informationsverlust eliminiert werden können. Für eine diskrete Kurve wie ein Polygon ist der Begriff der Krümmung mathematisch nicht exakt definiert. Deshalb erfolgt die Entscheidung, ob ein Stützpunkt des Polygons ein kritischer Punkt ist, anhand des Approximationsfehlers, der bei Elimination dieses Punktes in Kauf genommen werden müßte:

Ein Polygon wird als zyklische Liste seiner Stützpunkte betrachtet, so daß jeder Punkt zwei Nachbarn, seinen Vorgänger und seinen Nachfolger in der Stützpunktliste, besitzt. Zusammen mit seinen direkten Nachbarn spannt jeder Punkt ein Dreieck mehr oder weniger großer Fläche auf. Kleine Dreiecksflächen deuten auf redundante Punkte hin, da bei Elimination des zugehörigen Stützpunkts nur geringe Änderungen in Form

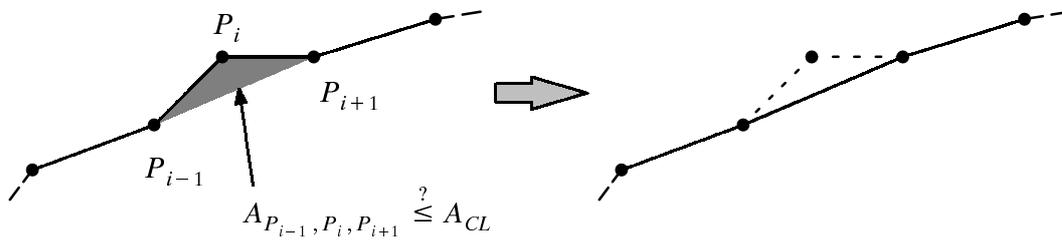


Bild A-1.1: Ein Punkt der Stützpunktliste wird dann entfernt, wenn der dadurch entstehende Fehler (Fläche des aufgespannten Dreiecks) eine Schwelle (Critical Level) unterschreitet.

und Fläche des Polygons entstehen (Bild A-1.1). Im wesentlichen funktioniert das Verfahren wie folgt:

1. Berechne die Fehlerflächen für jeden Stützpunkt des Polygons
2. Suche den Punkt/die Punkte  $P_i$  mit minimaler Fehlerfläche  $A_{P_{i-1}, P_i, P_{i+1}}$ .
3. Falls die Fehlerfläche  $A_{P_{i-1}, P_i, P_{i+1}}$  einen vorgegebenen Schwellwert  $A_{CL}$  (Critical Level) unterschreitet, eliminiere  $P_i$  und aktualisiere die Fehlerflächen für  $P_{i-1}$  und  $P_{i+1}$ . Gehe zu Schritt 2.

Neben den Fehlerflächen wird auch das Maß der Richtungsänderung in einem Konturpunkt berücksichtigt. Der eingesetzte Algorithmus ist detailliert in [168] beschrieben und wird dort mit anderen Verfahren zur Konturapproximation verglichen.

Ein weiterer unerwünschter Effekt der Segmentierung ist, daß eng beieinander liegende Gebäude oftmals als eine zusammenhängende Region segmentiert werden, so daß zerklüftete Regionen mit Konkavitäten entstehen. Das beschriebene Verfahren zerlegt deshalb die Polygone an besonders engen Stellen in Teilpolygone, so daß mehrere, kompakte Einzelregionen entstehen. Die Zergliederung in Teilpolygone erfolgt dort, wo der Abstand zwischen zwei Konturpunkten unter eine Schwelle fällt und die Konturpunkte nicht benachbart im Sinne der Konturtopologie sind. Es entstehen zwei Teilpolygone mit einem gemeinsamen Polygonsegment. Zu kleine Teilpolygone werden ganz eliminiert.

In Bild A-1.2 ist beispielhaft das Ergebnis der Polygonapproximation und -zerlegung für einen Ausschnitt des Luftbilds von 1997 gezeigt. Die erste polygonale Beschreibung des segmentierten Gebäudekomplexes enthält 505 Stützpunkte. Nach der Bearbeitung wird der Komplex durch zwei Polygone mit 34 bzw. 15 Stützpunkten beschrieben, ohne die ursprüngliche Form erheblich zu verfälschen.

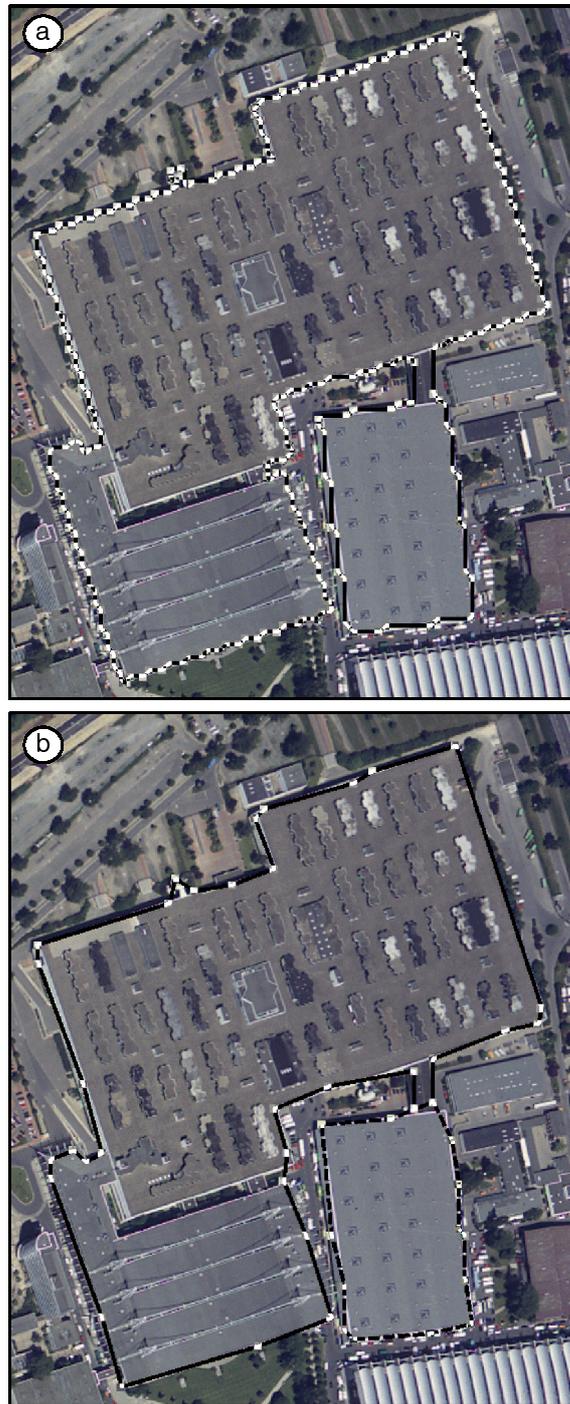


Bild A-1.2: *a) Erste polygonale Beschreibung eines segmentierten Hallenkandidats  
b) Vereinfachte Beschreibung der Region und Zerteilung in zwei Polygone  
(durchgezogene bzw. gestrichelte Kontur)*

## A-2 Detektion von Parkplätzen

Bei beiden in Kapitel 5 genannten Anwendungsbeispielen wird ein Verfahren zur automatischen Detektion von Parkplätzen benötigt. Im verwendeten Bildmaterial fallen die Parkplätze aufgrund der regelmäßigen Anordnung von Fahrspuren als eine Anhäufung von parallelen Linien auf. Aus diesem Grund basiert das verwendete Verfahren auf einer Liniendetektion mit anschließender Gruppierung paralleler, benachbarter Linien.

In Bild A-1.3 sind die einzelnen Verarbeitungsschritte mit den korrespondierenden Zwischenergebnissen für einen Luftbildausschnitt dargestellt. Zunächst erfolgt die Extraktion und Vektorisierung linienförmiger Objekte. Als Liniendetektor wird das Verfahren nach Steger [144] eingesetzt. Nach der Verschmelzung benachbarter, kollinearere Linien werden kurze Linien eliminiert, wodurch sich die Zahl der Linienkandidaten deutlich reduzieren läßt (Bild A-1.3b). Anschließend erfolgt eine Richtungsanalyse der übrig gebliebenen Linien, um sie anhand ihrer Orientierung gruppieren zu können. Die Richtungen der Linien werden in 16 Stufen quantisiert, wodurch geringe Orientierungsabweichungen toleriert werden. In Bild A-1.3c sind die verschiedenen Richtungsklassen durch unterschiedliche Punktierungsarten markiert. Schließlich erfolgt die Zusammenfassung benachbarter, parallel verschobener Linien derselben Richtungsklasse zu sog. *Clustern* und die Verschmelzung dicht beieinanderliegender Liniencuster. Solche Cluster, die eine ausreichende Zahl von Linien enthalten, werden als Parkplatzkandidaten übernommen und durch ihre umgebende Hülle polygonal beschrieben (Bild A-1.3d). Die Kandidaten werden anschließend durch das Bildinterpretationssystem bezüglich ihrer Größe, Form und Lage untersucht und schließlich als Parkplatz akzeptiert oder abgelehnt.

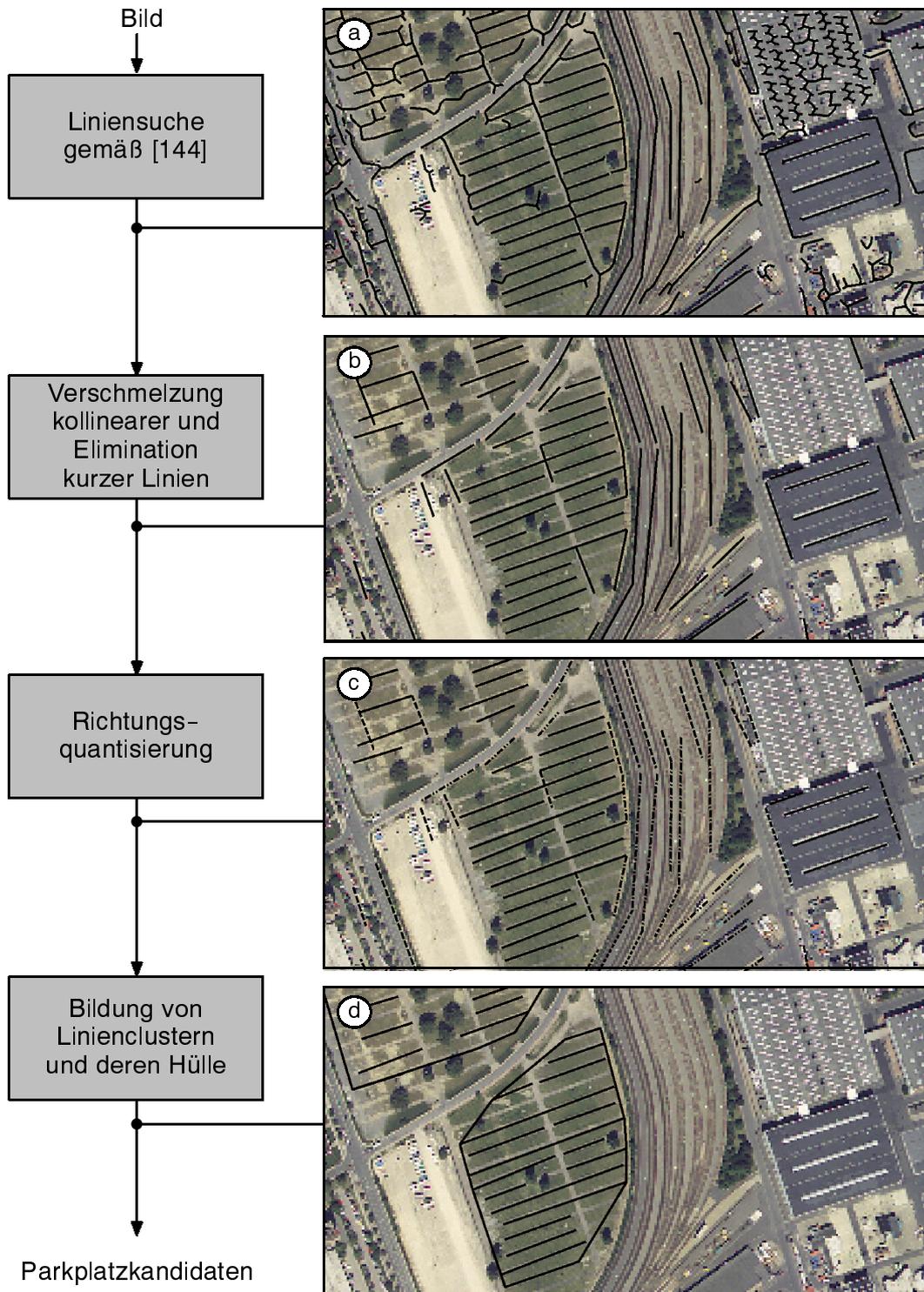


Bild A-1.3: Detektion von Parkplätzen als Anhäufung benachbarter, paralleler Linien

## A-3 Graphische Benutzerschnittstelle

Das Bildinterpretationssystem AIDA ist über eine graphische Benutzerschnittstelle bedienbar, mit der das semantische Netz visualisiert und inspiziert werden kann

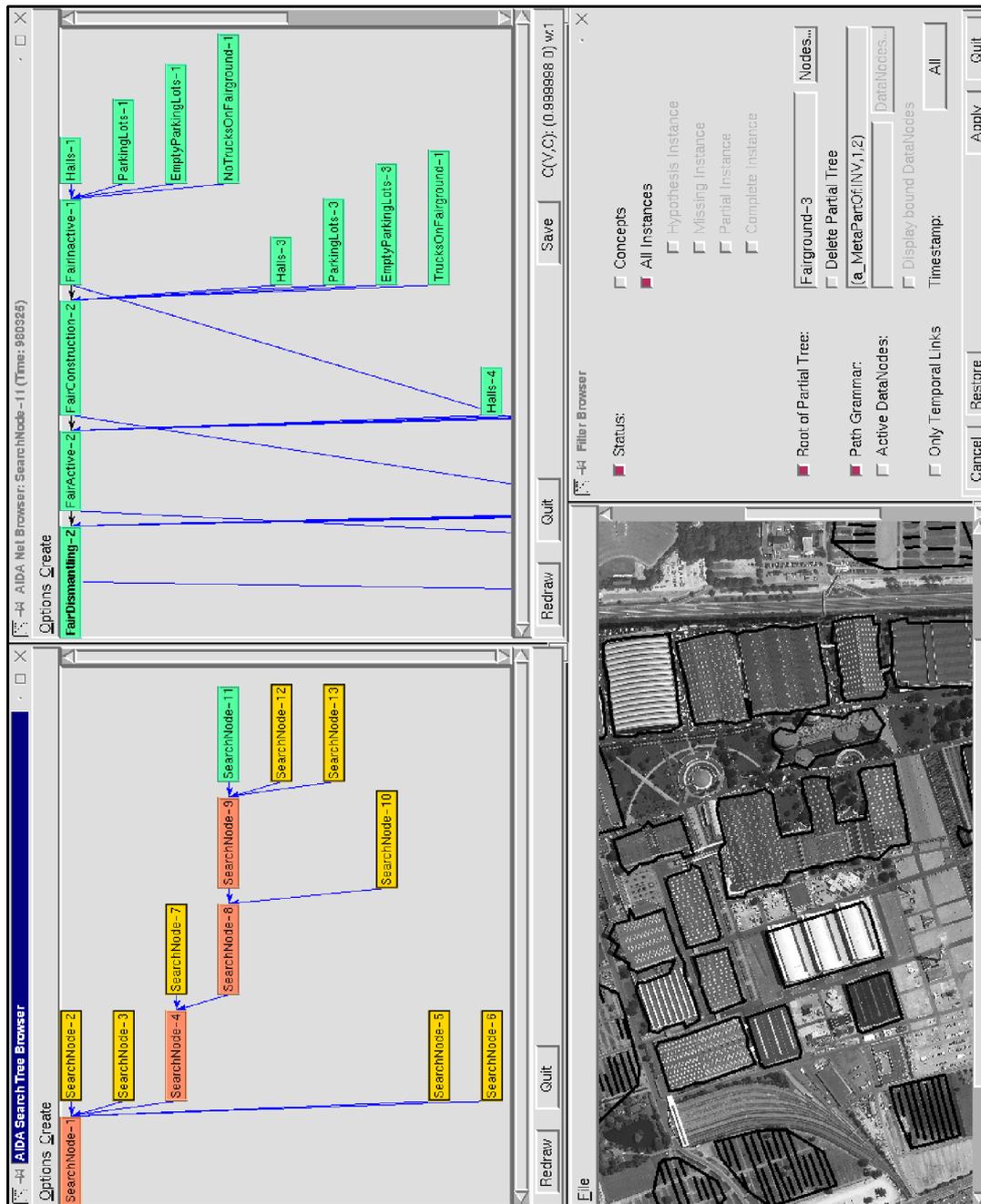


Bild A-1.4: Graphische Benutzerschnittstelle des Bildinterpretationssystems AIDA

---

(Bild A-1.4). Darüberhinaus verfügt die Oberfläche über Editierfunktionalitäten, so daß sich eine Wissensbasis interaktiv erstellen oder während der Analyse modifizieren läßt. Neben dem Suchbaum und der aktuellen Szenenbeschreibung können einzelne Knoten, Kanten und Attribute mitsamt ihrer aktuellen Wertebelegung visualisiert und editiert werden.

Umfangreiche semantische Netze werden schnell unübersichtlich. So läßt sich die resultierende Szenenbeschreibung aus Kapitel 5.1 mit ihren 832 Instanzen nur schwer visuell erfassen. Aus diesem Grund wurde die graphische Benutzerschnittstelle um die Möglichkeit erweitert, das semantische Netz nach unterschiedlichen Kriterien zu filtern, um das Netz partiell zu visualisieren (Bild A-1.4 rechts unten). Neben der Auswahl, ob Konzepte oder Instanzen angezeigt werden sollen, läßt sich ein Knoten als Wurzel eines darzustellenden bzw. zu eliminierenden Teilnetzes definieren. Soll das Netz nur bis zu einer gewissen Tiefe visualisiert werden, so kann der entsprechende Bereich über eine Pfadgrammatik spezifiziert werden. Dadurch wird es auch möglich, einzelne Ebenen des semantischen Netzes oder nur Bestandteilshierarchien zu selektieren. Der Benutzer hat die Wahl, ob die über eine *data-of*-Kante angebotenen Datenknoten dargestellt werden sollen. Im Rahmen einer multitemporalen Bildanalyse kann er durch eine Filterung bezüglich des Zeitstempels diejenigen Teilnetze inspizieren, die das korrespondierende Bild der Zeitreihe beschreiben. Läßt er sich ausschließlich die durch temporale Relationen verbundenen Knoten anzeigen, können die zeitlichen Veränderungen der Szenenobjekte besonders gut veranschaulicht werden.

---

## Literatur

- [1] AdV: "ATKIS-Gesamtdokumentation, Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV)", *Landesvermessungsamt Nordrhein-Westfalen*, Bonn, 1989.
- [2] Aho, A. V., Ullman, J. D.: "The Theory of Parsing, Translation and Compiling. Vol. 1", *Prentice Hall*, 1972.
- [3] Allen, J.: "Maintaining knowledge about temporal intervals", *Communications of the ACM*, Vol. 26, S. 832-843, 1983.
- [4] Allen, J.: "Towards a general theory of action and time", *Artificial Intelligence*, Vol. 23, S. 123-154, 1984.
- [5] Allen, J., Hayes, P.: "A common-sense theory of time", *Proceedings of IJCAI'85*, S. 528-531, 1985.
- [6] Allen, J.: "Planning as temporal reasoning", *Proceedings KR'91*, S. 3-14, 1991.
- [7] Ament, R.: "Orthobilder in der ATKIS-Fortführung", In: Fritsch, D., Hobbie, D. (Hrsg.): *Photogrammetric Week '97*, Wichmann Verlag, Heidelberg, S. 261-272, 1997.
- [8] Andreassen, S., Jensen, F. V., Andersen, K., Falck, B., Kjærulff, U., Woldbye, M., Sørensen, A., Rosenfalck, A., Jensen, F.: "MUNIN - an expert EMG assistant", In: Desmedt, J. E. (Hrsg.): *Computer-aided electromyography and expert systems*, Elsevier Science, Amsterdam, S. 255-277, 1989.
- [9] Azimi-Sadjadi, M. R., Ghaloum, S., Zoughi, R.: "Terrain Classification in SAR images using Principal Component Analysis and Neural Networks", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 31 (2), S. 511-515, 1993.
- [10] Baum, Petrie, T.: "Statistical inference for probabilistic functions of finite state Markov chains", *Ann. Math. Stat.*, Vol. 37, S. 1554-1563, 1966.
- [11] Baumgartner, A., Steger, C., Mayer, H., Eckstein, W., Ebner, H.: "Automatic Road Extraction based on Multi-Scale, Grouping, and Context", *Photogrammetric Engineering & Remote Sensing*, Vol. 65 (7), S. 777-786, 1999.
- [12] Bigalke, J.: "Detektion von baulichen Veränderungen aus multitemporalen Luftbildern mit Hilfe eines wissensbasierten Bildinterpretationssystems", *Diplomarbeit Universität Hannover*, Universität Hannover, 2000.

- 
- [13] Binford, T., Levitt, T.: "Model-based recognition of objects in complex scenes", *Proc. of the 1994 ARPA Image Understanding Workshop*, Morgan Kaufmann, San Mateo, 1994.
- [14] Blonda, P. N., Pasquariello, G., Losito, S., Mori, A., Posa, F., Ragno, D.: "An Experiment for the Interpretation of Multitemporal remotely sensed Images based on a Fuzzy Logic Approach", *International Journal of Remote Sensing*, Vol. 12, S. 463-476, 1991.
- [15] Bruzzone, L., Pietro, D. F., Serpico, S. B.: "A Neural-Statistical Approach to Multitemporal and Multisource Remote-Sensing Image Classification", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 37 (3), S. 1350-1359, 1999.
- [16] Buxton, H., Gong, S.: "Visual Surveillance in a dynamic and uncertain world", *Artificial Intelligence*, Vol. 78, S. 431-459, 1995.
- [17] Brooks, R. A.: "Model-Based Three-Dimensional Interpretations of Two-Dimensional Images", *IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 5, S. 140-150, 1983.
- [18] Bruce, B.: "A model for temporal references and its application in a question answering program", *Artificial Intelligence*, Vol. 4, S. 1-25, 1972.
- [19] Buntine, W. L.: "Operations for Learning with Graphical Models", *Journal of Artificial Intelligence Research*, S. 159-225, 1994.
- [20] Chin, H. L., Cooper, G. F.: "Bayesian Belief Network Inference using Simulation", *Uncertainty in Artificial Intelligence 3*, North Holland, Amsterdam, 1989.
- [21] Clement, V., Giraudon, G., Houzelle, S., Sadakly, F.: "Interpretation of Remotely Sensed Images in a Context of Multisensor Fusion Using a Multispecialist Architecture", *IEEE Trans. on Geoscience and Remote Sensing*, Vol. 31, No 4, S. 779-791, 1993.
- [22] Collins, R. T., Jayens, C. O., Cheng, Y.-Q., Wang, X., Stolle, F., Riseman, E. M., Hanson, A. R.: "The Ascender System: Automated Site Modeling from Multiple Aerial Images", *Computer Vision and Image Understanding*, Vol. 72 (2), S. 143-162, 1998.
- [23] Coiner J.C.: "Using Landsat to Monitor Changes in Vegetation Cover Induced by Desertification Processes", *Proceedings of the 14th Int. Symposium on Remote Sensing of Environment*, Ann Arbor, Michigan, 1980.

- 
- [24] Corr, D. G., Taylor, A. M., Cross, A., Hogg, D. C., Lawrence, D. H., Mason, D. C., Petrou, M.: "Progress in Automatic Analysis of Multi-Temporal Remotely-Sensed Data", *International Journal of Remote Sensing*, Vol. 10 (7), S. 1175-1195, Juli 1989.
- [25] Dai, X., Khorram, S.: "Data Fusion using Artificial Neural Networks: A Case Study on Multitemporal Change Analysis", *Computers, Environment and Urban Systems*, Vol. 23 (1), S. 19-31, 1999.
- [26] Dai, X., Khorram, S.: "Requirements and Techniques for an Automated Change Detection System", *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS'98)*, Vol. 5, S. 2752-2754, 1998.
- [27] Dean, T., McDermott, D.: "Temporal data base management", *Artificial Intelligence*, Vol. 32, S. 1-55, 1987.
- [28] Dean, T., Kanazawa, K.: "Probabilistic Temporal Reasoning", *Proceedings of AAAI'88*, S. 524-528, 1988.
- [29] Dechter, R., Meiri, I., Pearl, J.: "Temporal constraint networks", *Artificial Intelligence*, Vol. 49, S. 61-95, 1991.
- [30] Deer, P.: "Digital Change Detection: Civilian and Military Applications", *Proceedings of the International Symposium on Spectral Sensing Research (ISSSR '95)*, Melbourne, 1995.
- [31] Dobson, M. C., Pierce, L. E., Ulaby, F. T.: "Knowledge-Based Land-Cover Classification using ERS-1/JERS-1 SAR Composites", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 34 (1), S. 83-99, 1996.
- [32] Dreschler-Fischer, L., Drewniok, C., Lange, H., Schröder, C.: "A Knowledge-Based Approach to the Detection and Interpretation of Changes in Aerial Images", *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS'93)*, Vol. 1, S. 159-161, 1993.
- [33] Drewniok, C., Schröder, C., Lange, H., Dreschler-Fischer, L.: "Automatisierte Bildauswertung als Hilfsmittel zur Verifikation", in: Kronfeld, U., Baus, W., Ebbesen, B., Jathe, M. (Hrsg.): *Naturwissenschaft und Abrüstung - Forschungsprojekte an deutschen Hochschulen*, Kieler Schriften zur Friedenswirtschaft, Bd. 2, Lit.-Verlag, Hamburg, S. 46-63, 1993.
- [34] Dubois, D., Prade, H.: "Possibility Theory: An Approach to Computerized Processing of Uncertainty", *Plenum Press*, New York/London, 263 S., 1988.
- [35] Dubois, D., Prade, H.: "Processing fuzzy temporal knowledge", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 19 (4), Juli/August 1989.

- 
- [36] Dubois, D., Lang, J., Prade, H.: "Timed possibilistic logic", *Fundamentae Informaticae*, Special Issue on Artificial Intelligence, Vol. 15 (3,4), S. 211-234, November/Dezember 1991.
- [37] Dutta, S.: "An event-based fuzzy temporal logic", *18th IEEE Int. Symposium on Multi-Valued Logics*, S. 64-71, 1988.
- [38] Englisch, A., Heipke, C.: "Erfassung und Aktualisierung topographischer Geo-Daten mit Hilfe analoger und digitaler Luftbilder", *Photogrammetrie Fernerkundung Geoinformation*, Bd. 3, S. 133-149, DGPF, Stuttgart, 1998.
- [39] Fischer, A., Kolbe, T. H., Lang, F., Cremers, A.B., Förstner, W., Plümer, L., Steinhage, V.: "Extracting Buildings from Aerial Images Using Hierarchical Aggregation in 2D and 3D" in: *Computer Vision and Image Understanding*, Vol. 72 (2), S. 185-203, November 1998.
- [40] Fiset, R., Cavayas, F., "Automatic Comparison of a Topographic Map with Remotely-Sensed Images in a Map Updating Perspective: The Road Network Case", *International Journal of Remote Sensing*, Vol. 18 (4), S. 991-1006, 1997.
- [41] Food, G. M., Boyd, D. S.: "Detection of partial Land Cover Change associated with the Migration of Inter-Class Transition Zones", *International Journal of Remote Sensing*, Vol. 20 (14), S. 2723-2740, 1999.
- [42] Fung, R., Chang: "Weighting and Integrating Evidence for Stochastic Simulation in Bayesian Networks", *Uncertainty in Artificial Intelligence 5*, North Holland, Amsterdam, 1990.
- [43] Gimel'farb, G. L., Zalesny, A. V.: "Low-Level Bayesian Segmentation of Piecewise-Homogenous Noisy and Textured Images", *International Journal of Imaging Systems and Technology*, Nr. 3, S. 227-243, 1991.
- [44] Gimel'farb, G. L.: "Image Textures and Gibbs Random Fields", *Kluwer Academic Publishers*, 250 S., Dordrecht/Boston/London, 1999.
- [45] Grasso, E., Lesmo, L., Lombardo, V., Maccario, P., Salato, R., Terenziani, P.: "Semantic interpretation of tense, actionality and aspect", *Proceedings of ECAI'90*, S. 320-325, 1990.
- [46] Grau, O.: "Wissensbasierte 3D-Analyse von Gebäudeszenen aus mehreren frei gewählten Stereofotos", *Dissertation Universität Hannover*, Hannover, 2000.
- [47] Growe, S.: "Knowledge Based Interpretation of Multisensor and Multitemporal Remote Sensing Images", *International Archives of Photogrammetry and Remote Sensing*, Vol. 32, Part 7-4-3W6, Valladolid, Spanien, S. 130-138, 1999.

- 
- [48] Growe, S., Schröder, T., Liedtke, C.-E.: "Use of Bayesian Networks as Judgement Calculus in a Knowledge Based Image Interpretation System", *International Archives of Photogrammetry and Remote Sensing*, Vol. 33, Part B3/1, Com. III, Amsterdam, S. 342-350, 2000.
- [49] Grünreich, D.: "ATKIS - A Topographic Information System as a Basis for a GIS and Digital Cartography in West Germany", *Geol. Jb.*, Vol. A122, S. 207-215, Hannover, 1992.
- [50] Haala, N., Brenner, C.: "Interpretation of Urban Surface Models Using 2D Building Information", *Computer Vision and Image Understanding*, Vol. 72 (2), S. 204-214, 1998.
- [51] Habermeyer, M., Schullius, C. C.: "Development of an Algorithm for the Knowledge-Based Classification of Multi-Temporal SAR Images.", *Proceedings of the SPIE - The international Society for Optical Engineering*, Vol. 3161, S. 148-158, 1997.
- [52] Hame, T., Heiler, I., San Miguel-Ayanz, J.: "Unsupervised Change Detection and Recognition System for Forestry", *International Journal of Remote Sensing*, Vol. 19 (6), S. 1079-1099, 1998.
- [53] Hanson, A. R., Riseman, E. M.: "VISIONS: A Computer System for Interpreting Scenes", In: Hanson, A. R., Riseman, E. M. (Hrsg.): *Computer Visions Systems*, Academic Press, New York, S. 303-334, 1978.
- [54] Haralick, R. M., Shapiro, L. G.: "Computer and Robot Vision", Addison-Wesley, Massachusetts, 1993.
- [55] Heckerman, D., Geiger, D., Chickering, D. M.: "Learning Bayesian networks: the combination of knowledge and statistical data.", In: de Mantaras, R. L., Poole, D. (Hrsg.): *Proceedings of the 10th Conference on Uncertainty in Artificial Intelligence*, Morgan Kaufman, San Francisco, S. 293-301, 1994.
- [56] Heckerman, D.: "A Tutorial on learning with Bayesian Networks", *Microsoft Research technical report*, MSR-TR-95-06, <ftp://ftp.research.microsoft.com/pub/tr/TR-95-06.ps>, 1995.
- [57] Heipke C., Beutner S., Straub B.-M., Wegmann H., Wiedemann C.: "Acquisition and Updating of ATKIS using Satellite Remote Sensing Imagery", *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS'99)*, Hamburg, 1999.

- 
- [58] Henricsson, O.: "The Role of Color Attributes and Similarity Grouping in 3-D Building Reconstruction", *Computer Vision and Image Understanding*, Vol. 72 (2), S. 163-184, 1998.
- [59] Holzschuh, G.: "Was ist Netzplantechnik ?", *Hüthig Verlag*, Heidelberg, 1989.
- [60] Hopcraft, J. E., Ullman, J. D.: "Introduction to Automata Theory, Languages and Computation", *Addison-Wesley*, Massachusetts, 1991.
- [61] Hornsby, K., Egenhofer, M.: "Identity-Based Change: A Foundation for Spatio-Temporal Knowledge Representation", *International Journal of Geographical Information Science*, Vol. 14 (3), S. 207-224, 2000.
- [62] Huang, T., Koller, D., Malik, J., Ogasawara, G., Rao, B., Russell, S., Weber, J.: "Automatic Symbolic Traffic Scene Analysis Using Belief Networks", In: *Proceedings of AAAI'94*, Seattle, S. 966-972, 1994.
- [63] Illert, A.: "Konzepte zur Fortführung des ATKIS", *Nachrichten aus dem Karten- und Vermessungswesen*, Reihe 1, Heft 111, S. 61-74, 1994.
- [64] Jaakkola, T., Jordan, M.: "Variational Probabilistic Inference and the QMR-DT Network", *Journal of Artificial Intelligence Research*, Vol. 1, 1999.
- [65] Jensen, F. V., Christensen, H. I., Nielsen, J.: "Bayesian methods for interpretation and control in multi-agent vision systems", In: Bowyer, K. W. (Hrsg.): *Proceedings of SPIE - Applications in Artificial Intelligence X*, SPIE - The International Society for Optical Engineering, Orlando, USA, S. 536-548, 1992.
- [66] Jensen, F. V.: "An Introduction to Bayesian Networks", *UCL Press*, London, 1996.
- [67] Jeon, B., Landgrebe, D.: "Classification with spatio-temporal interpixel class dependency contexts", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 30, S. 663-672, 1992.
- [68] Jha, C. S., Unni, N. V. M.: "Digital Change Detection of Forest Conversion of a Dry Tropical Forest Region", *International Journal of Remote Sensing*, Vol. 15, S. 2543-2553, 1994.
- [69] Jiang-Ming, L., Christensen, H. I., Jensen, F. V.: "Qualitative Recognition using Bayesian Reasoning", *Proceedings of Pattern Recognition in Practice IV*, 1994.
- [70] Jürgens, C.: "Change Detection - Erfahrungen bei der vergleichenden multitemporalen Satellitenbilddauswertung in Mitteleuropa", *Photogrammetrie Fernerkundung Geoinformation*, Bd. 1, S. 5-18, 2000.

- 
- [71] Kahn, K., Gorry, G.: "Mechanizing temporal knowledge", *Artificial Intelligence*, Vol. 9, S. 87-108, 1977.
- [72] Kohane, I.S. (Hrsg.): "Temporal Reasoning in Medical Expert Systems", *Massachusetts Institute of Technology*, 1987.
- [73] Kofler, M., Gruber, M.: "Verwaltung und Visualisierung dreidimensionaler Stadtmodelle". *Zeitschrift für Photogrammetrie und Fernerkundung*, Heidelberg, 1998.
- [74] Kohl, H. G., King, C., De Goof, H.: "Agricultural statistics: Comparison of ERS-1 and SPOT for the crop acreage estimation of the MARS project", *Proceedings of the 2nd ERS-1 Symposium*, Vol. 1, S. 87-92, Hamburg, 1993.
- [75] Konecny, G.: "Hochauflösende Fernerkundungssensoren für kartographische Anwendungen in Entwicklungsländern", *Zeitschrift für Photogrammetrie und Fernerkundung*, Bd. 64 (2), Wichmann Verlag, Heidelberg, S. 39-51, 1996.
- [76] Kraus, K.: "Photogrammetrie", *Dümmler*, Bonn, 1996.
- [77] Krebs, B., Burkhardt, M., Wahl, F. M.: "A Bayesian Network for 3d Object Recognition in Range Data", *Proceedings of the Int. Conference of Computer Analysis of Images and Patterns (CAIP'97)*, Kiel, S. 361-368, 1997.
- [78] Krebs, B., Wahl, F. M.: "Automatic Generation of Bayesian Nets for 3d Object Recognition", *Proceedings of the Int. Conference on Pattern Recognition (ICPR'98)*, Brisbane, Australien, S. 126-128, 1998.
- [79] Kruse, R., Schwecke, E., Heinsohn, J.: "Uncertainty and Vagueness in Knowledge Based Systems", *Springer-Verlag*, Berlin, 491 S., 1991.
- [80] Kulschewski, K.: "Building Recognition with Bayesian Networks", In: Förstner, W., Plümer, L. (Hrsg.): *Semantic Modeling for the Acquisition of Topographic Information from Images and Maps (SMATI'97)*, Birkhäuser, Basel, S. 196-210, 1997.
- [81] Kulschewski, K., Koch, K. R.: "Recognition of Buildings using a dynamic Bayesian network", *Proceeding of the Workshop "Semantic Modeling for the Acquisition of Topographic Information from Images and Maps" (SMATI'99)*, München, S. 121-132, 1999.
- [82] Kumar V. P., Desai U. B.: "Image Interpretation Using Bayesian Networks", *Trans. on Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 18 (1), S. 74-77, 1996.

- 
- [83] Kunz, D., Schilling, K.-J., Vögtle, T.: "A new approach for Satellite Image Analysis by means of Semantic Network", In: Förstner, W., Plümer, L. (Hrsg.): *Semantic Modeling for the Acquisition of Topographic Information from Images and Maps (SMATI'97)*, Birkhäuser, Basel, S. 20-36, 1997.
- [84] Lambin, E. F., Strahler, A. H.: "Change Vector Analysis in Multitemporal Space: a Tool to Detect and Categorize Land-Cover Changes Processes Using High Resolution Satellite Data", *Remote Sensing of Environment*, Vol. 48, S. 231-244, 1994.
- [85] Lang, F.: "Geometrische und semantische Rekonstruktion von Gebäuden durch Ableitung von 3D Gebäudeecken", Dissertation, Universität Bonn, *Reihe Informatik*, Shaker Verlag, 1999.
- [86] Lange, H., Schröder, C.: "Analysis and Interpretation of Changes in Aerial Images: Knowledge Representation and Spatial Reasoning", In: Ebner, H., Heipke, C., Eder, K. (Hrsg.): *Proceedings of the ISPRS Commission III Symposium "Spatial Information from Digital Photogrammetry and Computer Vision"*, Vol. 30 of IAPRS, S. 475-482, München, 1994.
- [87] Lauritzen, S. L., Spiegelhalter, D. J.: "Local computations with probabilities on graphical structures and their application to expert systems.", *Journal of the Royal Statistical Society, Series B*, Vol. 50 (2), S. 157-224, 1988.
- [88] Lemmens, M.: "A Survey on Stereo Matching Techniques", *International Archives for Photogrammetry and Remote Sensing*, Vol. 27, Com. V/6, Kyoto, S. 11-23, 1988.
- [89] Liedtke, C.-E., Bückner, J., Grau, O., Growe, S., Tönjes, R.: "AIDA: A System for the Knowledge Based Interpretation of Remote Sensing Data", *3rd Int. Airborne Remote Sensing Conference and Exhibition*, Kopenhagen, Dänemark, S. 313-320, 1997.
- [90] Liedtke, C.-E., Ender, M.: "Wissensbasierte Bildverarbeitung", *Springer-Verlag*, Berlin, 230 S., 1989.
- [91] Loffet, A.: "3D-Models for Telecommunication - Methods and Experiences", *Proceedings of the OEEPE Workshop on 3D-City Models*, Institut für Photogrammetrie, Universität Bonn, Oktober 1996.
- [92] Lütjen, K.: "BPI: Ein Blackboard-basiertes Produktionssystem für die automatische Bildauswertung", In: Hartmann, G. (Hrsg.), *Mustererkennung 1986*, 8. DAGM-Symposium, S. 164-168, Springer-Verlag, Berlin, 1986.

- 
- [93] Mas, J.-F.: "Monitoring Land-Cover Changes: A Comparison of Change Detection Techniques", *International Journal of Remote Sensing*, Vol. 20 (1), S. 139-152, 1999.
- [94] Matsuyama, T., Hwang, V.S.-S.: "SIGMA: A Knowledge-Based Aerial Image Understanding System", *Plenum Press*, New York/London, 277 S., 1990.
- [95] McDermott, D.: "A temporal knowledge for reasoning about processes and plans", *Cognitive Science*, Vol. 6, S. 101-155, 1982.
- [96] McKeown, D., Harvey, W. A., McDermott, J.: "Rule-Based Interpretation of Aerial Imagery", *IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 7 (5), S. 570-585, 1985.
- [97] Mees, W., Perneel, C.: "Advances in computer assisted image interpretation", *Informatica - International Journal of Computing and Informatics*, Vol. 22(2), S. 231-243, 1998.
- [98] Middelkoop, J., Jansen, L. L. F.: "Implementation of temporal relationships in knowledge based classification of satellite images", *Photogrammetric Engineering and Remote Sensing*, Vol. 57, S. 937-945, 1991.
- [99] Munck-Fairwood, R.: "Recognition of geometric primitives using logic program and probabilistic network reasoning methods", *Proceedings of SPIE - Applications in Artificial Intelligence X; Machine Vision and Robotics*, Orlando, USA, S. 589-600, 1992.
- [100] Neapolitan, R. E.: "Probabilistic Reasoning in Expert Systems", *Wiley*, New York, 1990.
- [101] Niemann, H., Sagerer, G., Schröder, S., Kummert, F.: "ERNEST: A Semantic Network System for Pattern Understanding", *IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 12 (9), S. 883-905, 1990.
- [102] Nilsson, N. J.: "Principles of Artificial Intelligence", *Springer-Verlag*, Berlin, 1982.
- [103] Okoniewski, P.: "Kontrolle der flächengebundenen Beihilfen durch Fernerkundung", *Informationen für die Agrarberatung*, Bd. 9, S. 6-8, 1996.
- [104] Pakzad, K., Bückner, J., Growe, S.: "Knowledge Based Moorland Interpretation using a Hybrid System for Image Analysis", In: *International Archives of Photogrammetry and Remote Sensing*, Vol. 32, Part 3-2W5, München, 1999.
- [105] Papadimitriou, C., Steiglitz, K.: "Combinatorial Optimization: Algorithms and Complexity", *Prentice Hall*, Englewood Cliffs, NJ, 1982.

- 
- [106] Parra, G. A., Mouchot, M.-C., Roux, C.: "Multitemporal Land-Cover Change Analysis Tool using Change Vector Analysis and Principal Component Analysis", *Proceedings of the 1996 International Geoscience and Remote Sensing Symposium (IGARSS '96)*, Vol. 3, S. 1753-1755, 1996.
- [107] Pearl, J.: "Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference", *Morgan Kaufmann Publishers*, San Francisco, 552 S., 1988.
- [108] Pearl, J.: "Heuristic Intelligent Search Strategies for Computer Problem Solving", *Addison-Wesley*, Massachusetts, 1984.
- [109] Perkins, W., Austin, A.: "Adding temporal reasoning to expert system building environments", *IEEE Expert*, S. 23-30, 1990.
- [110] Pohl, I.: "The Avoidance of (relative) Catastrophe, Heuristic Competence, Genuine Dynamic, Weighting and Computational Issue in Heuristic Problem Solving", *Proc. of IJCAI 3*, Stanford, S. 20-23, 1993.
- [111] Prade, H.: "Operations Research with Fuzzy Data", in: Wang, P., Chang, S., (Hrsg.): *Fuzzy Sets. Theory and Applications to Policy Analysis and Information Systems*, New York/London, S. 155-169, 1980
- [112] Prantl, M., Ganster, H., Pinz, A.: "Active Fusion Using Bayesian Networks Applied to Multi-Temporal Remote Sensing Imagery", *Proceedings of the 13th International Conference on Pattern Recognition (ICPR '96)*, Vol. 3, S. 890-894, Wien, 1996.
- [113] Quint, F.: "Kartengestützte Interpretation monokularer Luftbilder", Dissertation Universität Karlsruhe, *Verlag der Bayerischen Akademie der Wissenschaften*, 105 S., 1997.
- [114] Quint, F.: "MOSES: A structural approach to aerial image understanding", In: Gruen, A., Baltsavias, E., Henricsson, O. P. (Hrsg.): *Automatic Extraction of Man-Made Objects from Aerial and Space Images (II)*, Birkhäuser, Basel, S. 323-332., 1997.
- [115] Rabetke, C.: "Fuzzy Sets in der Netzplantechnik", *Deutscher Universitäts-Verlag*, Wiesbaden, 273 S., 1991.
- [116] Rasmussen, L. K.: "BOBLO: an expert system based on Bayesian networks to blood group determination of cattle", *Research report 16*, Research Center Foulum, Denmark, 1995.
- [117] Rehg, J., Murphy, K., Russell, S.: "Vision-based Speaker Detection using Bayesian Networks", *Proceedings of the Int. Conference on Computer Vision and Pattern Recognition (CVPR'99)*, 1999.

- 
- [118] Renolen, A.: "History graphs: Conceptual modelling of spatiotemporal data". *GIS Frontiers in Business and Science*, Brno, Tschechische Republik, 1996..
- [119] Reichgelt, H., Shadbolt, N.: "A specification tool for planning systems", *Proceedings of ECAI'90*, S. 541-546, 1990.
- [120] Ridd, M. K., Liu, J.: "Comparison of four algorithms for Change Detection in an Urban Environment", *Remote Sensing of Environment*, Vol. 63 (2), S. 95-100. 1998.
- [121] Rimey, R., Brown, C.: "Control of selective perception using Bayes nets and decision theory", *International Journal of Computer Vision*, Special issue in Active Vision, Vol. 12, S. 173-207, 1994.
- [122] Rucker, D., Maron, D, Shortlife, E.: Temporal representation of clinical algorithms using expert systems", *Computer and Biomedical Research*, Vol. 23, S. 222-239, 1990.
- [123] Russ, T.: "Using hindsight in medical decision making", *Computer Methods and Programs in Biomedicine*, Vol. 32 (1), S. 81-90, 1990.
- [124] Sagerer, G., Niemann, H.: "Semantic Networks for Understanding Scences", *Advances in Computer Vision and Machine Intelligence*, Plenum Press, New York/London, 1997.
- [125] Sagerer, G.: "Automatisches Verstehen gesprochener Sprache", *BI Wissenschaftsverlag*, Mannheim, 1990.
- [126] Sarkar, S., Boyer, K. L.: "Integration, Inference and Management of Spatial Information using Bayesian networks: Perceptual Organization", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 15, S. 256-274, 1993.
- [127] Sarkar, S., Boyer, K. L.: "Computing Perceptual Organization in Computer Vision", *World Scientific Publishing Co.*, Singapore/New Jersey/London/Hong Kong, 232 S., 1994.
- [128] Schieche, B., Erasmi, S., Schrage, T., Hurlemann, P.: "Monitoring and Registering of Grassland and Fallow Fields with multitemporal ERS Data within a district of Lower Saxony, Germany", *Proceeding of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS'99)*, Hamburg, 1999.
- [129] Schiewe, J., "Suitability of MOMS-02 Data for hierachical map revision", *ISPRS Joint Workshop "Sensors and Mapping from Space"*, Working Groups I/1, I/3 und I/4, Hannover, S. 247-256, 1997.

- 
- [130] Schneider, V.: "Bayesnetze als Bewertungskalkül für Semantische Netze", *Diplomarbeit Universität Bielefeld*, Bielefeld, 1996.
- [131] Schröder, T.: "Nutzung von Bayes-Netzen für die Verarbeitung von Unsicherheit und Ungenauigkeit in einem wissensbasierten Bildinterpretationssystem", *Diplomarbeit Universität Hannover*, Hannover, 1999.
- [132] Schwarze, J.: "Simulation von Netzplänen mit stochastischen und abhängigen Vorgangsdauern", *Operations Research Proceedings 1980*, Berlin/Heidelberg/New York, S. 190-197, 1981.
- [133] Schwarze, J.: "Netzplantechnik", *Neue Wirtschafts-Briefe*, Herne/Berlin, 1990.
- [134] Schwidofsky, K., Ackermann, F.: "Photogrammetrie", *B. G. Teubner*, Stuttgart, 1976.
- [135] Shafer, G., Pearl, J.: "Readings in Uncertain Reasoning", *Morgan Kaufmann Publishers*, San Mateo, USA, 1990.
- [136] Shafer, G.: "A Mathematical Theory of Evidence", *Princeton University Press*, Princeton, USA, 1976.
- [137] Shapiro, S. C. (Hrsg.): "Encyclopedia of Artificial Intelligence", *John Wiley & Sons*, New York, 1992.
- [138] Siebe, E.: "3D-Urban Data for Planning Cellular Radio Networks", *XVIIIth ISPRS Congress, Intern. Archives of Photogrammetry and Remote Sensing*, Bd. 31/B4, S. 797-802, 1996.
- [139] Singh, A.: "Digital Change Detection Techniques Using Remotely-Sensed Data", *International Journal of Remote Sensing*, Vol. 10, S. 989-1003, 1989.
- [140] Smits, P. C., Annoni, A.: "Updating Land-Cover Maps by Texture Information from very high-resolution space-borne Imagery", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 37 (3), S. 1244-1254, 1999.
- [141] Solberg, A. H. S., Jain, A. K., Taxt, T.: "Multisource Classification of Remotely Sensed Data: Fusion of Landsat TM and SAR Images", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 32, S. 768-778, 1994.
- [142] Solberg, A. H. S., Taxt, T., Jain, A. K.: "A Markov Random Field Model for Classification of Multisource Satellite Imagery", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 34 (1), S. 100-113, 1996.

- 
- [143] Stassopoulou, A., Petrou, M., Kittler, J.: "Application of a Bayesian network in a GIS based decision making system", *Int. Journal on Geographical Information Science*, Vol. 12, S. 23-45, 1998.
- [144] Steger, C.: "An Unbiased Detector of Curvilinear Structures" *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20 (2), S. 113-125, 1998.
- [145] Stewart, J. B., Finch, J. W.: "Application of Remote Sensing to Forest Hydrologie", *Journal of Hydrology*, Vol. 150 (2-4), S. 701-716, 1993.
- [146] Stilla, U., Michaelsen, E.: "Semantic modelling of man-made objects by production nets", In: Gruen, A., Baltsavias, E., Henricsson, O. P. (Hrsg.): *Automatic Extraction of Man-Made Objects from Aerial and Space Images (II)*, Birkhäuser, Basel, S. 43-52, 1997.
- [147] Stow, D. A., Tinney, L. R., Estes, J. E.: "Deriving Land Use/Cover Change Statistics From Landsat: a Study of Prime Agricultural Land", *Proceedings of the 14th International Symposium on Remote Sensing of Environment*, Ann Arbor, Michigan, S. 1227-1237, 1980.
- [148] Suter, M., Nüesch, D.: "Automated Generation of Visual Simulation Databases Using Remote Sensing and GIS", *Proceedings of IEEE Visualization'95*, Atlanta, S. 86-93, 1995.
- [149] Terres, J. M., Delince, J., van de Steene, M., Hawkins, A.: "The use of remote sensing and GIS capabilities to support the Reform of the Common Agricultural Policy of the European Community", *Remote Sensing Review*, Vol. 12, S. 53-60, 1995.
- [150] Todt, H.: "The effect of the distribution-type on the statistical calculation of networks", In: Lombaers, H. (Hrsg.): *Project planning by network analysis*, Amsterdam/London, 1969.
- [151] Tönjes, R., Growe, S., Bückner, J., Liedtke, C.-E.: "Knowledge Based Interpretation of Remote Sensing Images Using Semantic Nets", *Photogrammetric Engineering & Remote Sensing*, Vol. 65 (7), S. 811-821, 1999.
- [152] Tönjes, R.: "Wissensbasierte Interpretation und 3D-Rekonstruktion von Landschaftsszenen aus Luftbildern", Dissertation, Universität Hannover, *Fortschritt-Berichte VDI*, Reihe 10, Nr. 575, VDI-Verlag, Düsseldorf, 1999.

- 
- [153] Tönjes, R., Growe, S.: "Knowledge Based Road Extraction from Multisensor Imagery", *ISPRS Symposium "Object Recognition and Scene Classification from Multispectral and Multisensor Pixels"*, Columbus, Ohio, USA, IAPRS Vol. 32 Part 3/1, S. 387-393, 1998.
- [154] Tönjes, R.: "Control of Scene Reconstruction Using Explicit Knowledge", *Proc. of IEEE Workshop on Applications of Computer Vision*, Sarasota, USA, S. 15-20, 1996.
- [155] Tsang, E.: "TLP - a temporal planner", In: Haalam J., Mellish, C. (Hrsg.): *Advances in Artificial Intelligence*, John Wiley & Sons, S. 63-78, 1987.
- [156] Veterinary Informatics and Epidemiology (VIE), "CaDDiS - Cattle Disease Diagnostic System", <http://www.dis.strath.ac.uk/vie/CaDDiS/>
- [157] Vila, L.: "A Survey on Temporal Reasoning in Artificial Intelligence", *AI Communications*, Vol. 7, S. 4-28, 1994.
- [158] Vilain, M.: "A system for reasoning about time", *Proceedings of AAAI'82*, S. 197-201, 1982.
- [159] Vosselmann, G., de Knecht, J.: "Road Tracing by profile matching and Kalman filtering", In: Grün, A., Kübler, O., Agouris, P. (Hrsg.): *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Birkhäuser, Basel, S. 265-274, 1995.
- [160] Wang, F.: "A Knowledge-Based Vision System for Detecting Land Changes and Urban Fringes", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 31, S. 136-145, 1993.
- [161] Weydahl, D. J.: "Change Detection in SAR Images", *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS'91)*, S. 1421-1424, 1991.
- [162] Wiemker, R., Spitzer, H.: "Änderungsdetektion auf multispektralen Luftbildern - Perspektiven für den Open-Skies-Vertrag", in: Altmann, J., Neuneck, G. (Hrsg.): *Naturwissenschaftliche Beiträge zu Abrüstung und Verifikation, Verhandlungen der Fachsitzung auf der 60. Physikertagung der Deutschen Physikalischen Gesellschaft (DPG) in Jena 1996*, FONAS/DPG, Hamburg, S. 138-151, 1996.
- [163] Winston, P. H., "Artificial Intelligence - 3rd Edition", *Addison-Wesley*, Massachusetts, 1992.

- 
- [164] Xiang, Y., Pant, B., Eisen, A., Beddoes, M., Poole, D.: "Multiply sectioned Bayesian networks for neuromuscular diagnosis", *Artificial Intelligence in Medicine*, Vol. 5 (4), S. 293-314, 1993.
- [165] Yip, K.: "Tense, aspect and the cognitive representation of time", *Proceedings IJCAI'85*, S. 806-814, 1985.
- [166] Zadeh, L. A.: "Fuzzy Sets." *Information and Control*, Vol. 8, S. 338-353, 1965.
- [167] Zadeh, L. A.: "The Concept of a Linguistic Variable and its Applications to Approximate Reasoning", Teil I: *Inform. Science*, Vol. 8, S. 199-249, Teil II: *Inform. Science*, Vol. 8, S. 301-357, Teil III: *Inform. Science*, Vol. 9, S. 43-80, 1975.
- [168] Zhu, P., Chirlian, P. M.: "On Critical Point Detection of Digital Shapes", *Ieee Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, Vol. 17 (8), S. 737-748, 1995.

# Schlagwortverzeichnis

## Zahlen

2D-Bildbereich, 19  
 3D-Rekonstruktion, 11  
 3D-Szenenbereich, 19

## A

A\*-Algorithmus, 28, 64, 89, 90  
 ACRONYM, 4, 14  
 AIDA, 5, 8  
   Steuerung, 22–33  
   Systemarchitektur, 11–13  
   Wissensrepräsentation, 14–22  
 Analyse  
   bottom-up, 13  
   datengetrieben, 13  
   modellgetrieben, 13  
   top down, 13  
 Änderungsdetektion, 2, 106–112  
 Anfangszeitpunkt, frühester, 49, 57  
 ATKIS, 12, 13  
 Attributberechnungsmethode, 17  
 Attribut, 16–17  
   Abbildung auf Bayes-Knoten, 85–87  
 Attributgewicht, 17, 32, 86  
 attributierte Relation, 19, 51, 83  
 Attributwert, 16, 85

## B

Bayes-Kante, 69  
 Bayes-Knoten, 69  
 Bayes-Netz, 30, 65  
   Abbildung eines semantischen Netzes, 80–87  
   als Wissensbasis, 71–73  
   Behandlung von Schleifen, 78–79  
   Definition, 69  
   Propagierung von Evidenz, 73–79  
   Theorie, 67–79  
   zur Bewertung des semantischen Netzes,  
   80–89

Bayes-Regel, 70–71  
 bedingte Glaubwürdigkeit, 78  
 bedingte Unabhängigkeit, 68–69  
 Belief, 67  
 Belief-Update, 73–79  
 Bestandteil, 15, 17  
 Bewertung, 17, 29  
 Bewertungsmethode, 16  
 Bewertungssystem  
   possibilistisches, 29–33  
   probabilistisches, 64–92  
   Vergleich, 90–92, 104–106, 112–114  
 Bild-zu-Bild-Verfahren, 3  
 Bildinterpretation, 2, 9  
   multitemporale, 34–63  
   Steuerung, 22  
   Strategie, 23, 58–63  
   Suchprozeß, 27  
   wissensbasierte, 8–33  
 Bildinterpretationssystem, 11  
 Bildprimitiv, 9, 12, 119  
 Bildverarbeitung, 12, 96–97, 122–123  
 bottom-up, 13, 23, 65, 76, 92, 114  
 BPI, 4, 14  
 Brachland, 36, 54

## C

Change Vector Analysis, 3  
 Clustering-Verfahren, 78  
 Conditioning-Verfahren, 78–80, 88, 92  
 CPM, 46–48

## D

DAG, 69  
 datengetrieben, 23, 33, 56  
 Dempster-Shafer-Theorie, 30  
 DHM, 9, 96  
 diagnostischer Support, 75, 80, 83, 88, 90, 104,  
 114

digitales Höhenmodell, 9, 96  
 Dummy-Knoten , 86

## E

endlicher Automat, 41  
 Endzeitpunkt, spätestes, 49, 57  
 Ereignisknotennetz, 47–49  
 ERNEST, 5, 14, 15  
 Evidenz, 16, 30, 67, 69–70  
 Evidenztheorie, 30

## F

fehlende Instanz, 16, 86  
 Folgezustand, 39  
   Prädiktion von, 56–58, 60  
 Formmerkmal, 97, 98  
 Fuzzy-Set, 30, 85

## G

Gauß-Krüger-Koordinaten, 93  
 generisches Szenenmodell, 10, 13  
 Geo-Daten, 1, 118  
 Geoinformationssystem, 1, 12  
 Geokodierung, 93  
 Geokoordinatensystem, 12  
 Gesamtpuffer, 49  
 Gibbs-Zufallsfelder, 108  
 GIS, 1, 12, 40, 93  
 GIS-Daten, 12, 21–23, 106, 108, 112, 114  
 GIS-Ebene, 21  
 Glaubwürdigkeit, 30, 66, 67, 70  
 Graphische Benutzerschnittstelle, 124–125

## H

hierarchische Kante, 17–18  
 Hyperspektral-Sensor, 9  
 Hypothese, 16, 17

## I

Image Differencing, 3  
 Image Ratioing, 3  
 Image Regression, 3  
 Inferenz, temporale, 58–60  
 Inferenzmaschine, 14, 23–27  
 Inferenzregel, 22  
   multitemporale, 61–63  
 Instanz, 16  
   fehlende, 16, 86  
   komplette, 16, 24, 61  
   partielle, 16, 24  
 Instanzierung, 16, 24  
   komplette, 24  
   multitemporale, 62  
   partielle, 24  
 IR, 20, 22

## K

Kante, 17–19  
   Abbildung auf Bayes-Kanten, 81–85  
   appearance-of, 15  
   attributierte, 19, 51, 83  
   concrete-of, 18, 82  
   data-of, 18, 19, 83  
   hierarchische, 17–18  
   instance-of, 17, 82  
   is-a, 15, 18, 19, 82  
   kontextabhängige, 18  
   part-of, 15, 17, 19, 82  
   temporale, 51, 83  
   topologische, 18–19  
 Kantenmatrix, 73, 76, 84, 91  
 kausaler Support, 75, 80, 88, 90, 104, 114  
 Klärwerk, 20  
 Klassifikationsvergleich, 3, 111  
 Knoten, 16  
   Abbildung auf Bayes-Knoten, 80–81  
 Knotengewicht, 16, 32, 88  
 komplette Instanz, 16  
 Konkretisierung, 15, 18  
 Konzept, 16  
 konzeptionelle Ebene, 18, 19, 51

kritischer Punkt, 119  
 künstliche Intelligenz, 11, 40, 43

## L

$\lambda$ -Nachricht, 74, 75  
 Landschaftsszene, 8  
 landwirtschaftliche Nutzfläche, 36, 53  
 Laserscanner, 9  
 Likelihood, 70, 73  
 Loop, in Bayes-Netzen, 78–79, 88  
 Luftbild, 1, 9
 

- multisensoriell, 9
- multitemporal, 2, 9

 Luftbildauswertung, 1

## M

Markov-Kette, 42–43  
 medizinische Diagnostik, 67  
 Messengelände, 54–56, 58–60, 93  
 MESSIE, 4, 14  
 Mixing Weight, 79  
 modellgetrieben, 23, 33, 56  
 Möglichkeit, 31, 90  
 MOSES, 5, 14  
 MPM, 46–48  
 multisensoriell, Luftbild, 9  
 Multispektral-Sensor, 9  
 multitemporal
 

- Luftbild, 2, 9
- Luftbildauswertung, 35–39

## N

Netzplantechnik, 41, 46–50  
 Notwendigkeit, 31, 90

## O

Orthophoto, 93

## P

partielle Instanz, 16  
 PERT, 46–48  
 Pfadgrammatik, 17  
 physikalische Ebene, 21  
 $\pi$ -Nachricht, 75  
 Plausibilität, 30  
 Polygonapproximation, 119–120  
 Possibility-Theorie, 30  
 Post Classification Comparison, 3  
 posteriori-Wahrscheinlichkeit, 70  
 Prädikatenlogik, 14  
 Prädiktion, 39, 40, 56–58  
 Principal Component Analysis, 3  
 probabilistisches Netzwerk, 65  
 Produktion, 14  
 Propagierung von Evidenz, 73–79

## R

Regel, 14, 23
 

- Hypothesengenerierung, 24
- Instanzierung, 24
- Mehrfachbindung, 25
- multitemporale, 61–63
- Spezialisierung, 25
- temporale Inferenz, 60

 Reihenmeßkamera, 8, 9, 19  
 Rekonstruktion, 11  
 Relation, 17  
 Restkosten, 28  
 RMK, 9, 12, 20

## S

SAR, 9, 12, 20  
 Satellitenbild, 1, 9  
 Segmentierung, 9, 13, 96–97, 119  
 Segmentierungsmethode, 16  
 semantische Ebene, 19  
 semantisches Netz, 14–19, 20, 28, 54–56, 95
 

- Abbildung auf ein Bayes-Netz, 80–87

Sensordaten, 13  
Sensorebene, 21  
Sensoren, 8  
SIGMA, 4, 14  
Simulation, 11  
SPAM, 4, 14  
Spezialisierung, 15, 18  
Stereokorrespondenzanalyse, 96  
stochastische Simulation, 78, 79  
Strategie, 23, 62  
strukturellen Mustererkennung, 11  
strukturelles Wissen, 10  
Suchbaum, 22, 28  
Suchbaumknoten, 22, 28, 64, 100, 103, 105  
    possibilistische Bewertung, 29–33  
    probabilistische Bewertung, 87–92  
symbolische Szenenbeschreibung, 9, 12  
Synthetic Aperture Radar, 9  
Szenenanalyse, 12

## T

temporal constraint network, 44  
temporal reasoning, 40  
temporale Relation, 51, 83  
temporales Wissen, 11, 34–63  
Texturanalyse, 108  
top-down, 13, 23, 65, 76, 114  
Topologie, 21  
topologische Kante, 18–19  
topologisches Wissen, 10

## U

Unglaubwürdigkeit, 30  
Unwissenheit, 30, 90

## V

Vegetation Index Differencing, 3  
VISIONS, 4, 14  
Visualisierung, 11  
Vorgangsknotennetz, 47–49

Vorgangspfeilnetz, 47–49

## W

Wertebereich, 17, 85  
Wissen  
    deklaratives, 14, 15  
    prozedurales, 14  
    strukturelles, 10  
    temporales, 11, 34–63  
    topologisches, 10  
wissensbasiertes System, 11  
Wissensbasis, 19–23, 94–96, 106–108  
Wissensnutzung, 38–39  
Wissensrepräsentation, 14, 36–38

## Z

Zeitintervall, 43–45  
zeitliches Schlußfolgern, 43–46  
Zeitstempel, 44, 50, 52, 59  
Zentralprojektion, 19  
Zielkonzept, 27, 99  
Zufallsvariable, 68  
Zustand, 36, 41  
Zustandsdauer, 37, 52, 103, 107  
Zustandsübergang, 37, 41  
Zustandsübergangsdauer, 52, 103, 107  
Zustandsübergangsdiagramm, 41–42, 51–56,  
    94, 107  
Zustandsübergangswahrscheinlichkeit, 37, 52,  
    105, 107, 112  
Zustandswahrscheinlichkeit, 37, 105

# Lebenslauf

Stefan Growe

## **Persönliches:**

24.9.1967 Geboren in Bad Oeynhausen  
Familienstand: Verheiratet

## **Ausbildung:**

1973 - 1977 Grundschule Nord, Bad Oeynhausen  
1977 - 1986 Immanuel-Kant-Gymnasium Bad Oeynhausen  
Abschluß: Allgemeine Hochschulreife

## **Wehrdienst:**

1986 - 1988 Soldat auf Zeit, Reserveoffizier

## **Studium:**

1988 - 1994 Universität Hannover  
Studium der Elektrotechnik, Schwerpunkt Nachrichtenverarbeitung  
Abschluß: Diplom (Dipl.-Ing.)  
1996 - 2000 Fernuniversität Hagen  
Aufbaustudium zum Wirtschaftsingenieur  
Abschluß: Diplom (Dipl.-Wirt. Ing.)

## **Beruf:**

1994 - 2000 Wissenschaftlicher Mitarbeiter am Institut für  
Theoretische Nachrichtentechnik und Informationsverarbeitung,  
Universität Hannover  
seit 11/2000 Consultant bei Ericsson Consulting GmbH, Düsseldorf  
im Bereich Mobile Internet Solutions