

UNIVERSITÄT BREMEN  
FACHBEREICH 3  
STUDIENGANG INFORMATIK

Diplomarbeit

**Modellierung und Klassifikation von  
Raumkonzepten anhand der  
Dempster-Shafer-Theorie am Beispiel  
von Universitätsräumen**

Joana Hois

Dezember 2005

Erstgutachter: Prof. Dr. Kerstin Schill  
Zweitgutachter: Dr. Thomas Barkowsky



# Eidesstattliche Erklärung

Hiermit erkläre ich, dass ich die vorliegende Diplomarbeit selbständig verfasst und noch nicht anderweitig zu Prüfungszwecken vorgelegt habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt. Wörtliche oder sinngemäße Zitate sind als solche gekennzeichnet.

Bremen, Dezember 2005

Joana Hois



# Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b>	<b>9</b>
<b>2</b>	<b>Stand der Forschung</b>	<b>13</b>
2.1	Visuelle Wahrnehmung . . . . .	13
2.1.1	Wahrnehmung von Objekten . . . . .	15
2.1.2	Raumwahrnehmung . . . . .	16
2.2	Bildanalyse . . . . .	17
2.2.1	Konzepte der Bildverarbeitung . . . . .	18
2.2.2	Prozessablauf des Bildverstehens . . . . .	20
2.3	Wissensrepräsentation und Wissensverarbeitung . . . . .	23
2.3.1	Vorgehensweise bei der Modellierung wissensbasierter Systeme . . . . .	23
2.3.2	Ontologien . . . . .	24
2.4	Unsicheres Wissen . . . . .	27
2.4.1	Wahrscheinlichkeitsbasierte Theorie . . . . .	28
2.4.2	Fuzzy Logic . . . . .	28
2.4.3	Dempster-Shafer-Theorie . . . . .	29
2.5	Kombination von Ontologien und unsicherem Wissen . . . . .	32
2.5.1	Kombination von OWL und der Bayes-Theorie . . . . .	33
2.5.2	Kombination von SWRL und Fuzzy Logic . . . . .	34
2.6	Inferenzmechanismen und Problemlösen . . . . .	35
2.6.1	Suchstrategien . . . . .	35
2.6.2	Reasoning . . . . .	36
2.7	Lernen . . . . .	37
2.7.1	Begriff des Lernens . . . . .	37
2.7.2	Lernverfahren . . . . .	37
<b>3</b>	<b>Raumwahrnehmung mittels Blickpunktfolgen</b>	<b>41</b>
3.1	Sakkadische Augenbewegungen . . . . .	41
3.2	Blickpunktfolgen einzelner Objekte . . . . .	42
3.3	Szenenanalyse mit Okusys . . . . .	43

3.4	Blickpunktfolgen komplexer Szenen . . . . .	45
3.5	Erweiterung der Szenenanalyse . . . . .	48
<b>4</b>	<b>Ontologien und unsicheres Wissen</b>	<b>49</b>
4.1	Ontologie . . . . .	49
4.1.1	Differenzierte Ontologiedefinitionen . . . . .	49
4.1.2	Vereinigung der unterschiedlichen Ontologiesichtweisen . . . . .	51
4.1.3	OWL (Web Ontology Language) . . . . .	53
4.1.4	OWL DL . . . . .	54
4.1.5	Ontologie-Editor . . . . .	56
4.2	Unsicheres Wissen . . . . .	56
4.3	Kombination: Dempster-Shafer-Theorie und Ontologien . . . . .	57
4.3.1	Glaubensmaß-Modellierung in OWL . . . . .	58
4.3.2	Anfragen bezüglich konkreter Raumkonzepte . . . . .	58
<b>5</b>	<b>Beispielontologie: Raumkonzepte</b>	<b>61</b>
5.1	Niches: Raumkonzepte . . . . .	61
5.2	Physical Endurant: Raum . . . . .	63
5.3	Physical Object: einzelne Objekte . . . . .	64
5.4	Glaubensmaß für Raumkonzepte . . . . .	64
5.5	Taxonomie der Raumontologie . . . . .	68
5.5.1	Bildmaterial . . . . .	69
5.5.2	Graphische Bearbeitung des Bildmaterials . . . . .	69
5.5.3	Glaubensmaße . . . . .	71
5.5.4	Instanzen der Ontologie . . . . .	71
5.6	Reasoning . . . . .	74
5.6.1	Reasoner: Pellet . . . . .	75
5.6.2	Anfragesprache: RDQL . . . . .	75
5.6.3	Belief-Berechnung: Dempster's Rule of Combination . . . . .	78
5.7	Okusys-Erweiterung . . . . .	80
5.7.1	Lernen der Objekte . . . . .	80
5.7.2	Raumanalyse . . . . .	82
<b>6</b>	<b>Implementierung</b>	<b>91</b>
6.1	Klassenstruktur . . . . .	91
6.2	Unittests . . . . .	97
6.3	Okusys-spezifische Daten . . . . .	97

<b>7</b>	<b>Evaluation der Okusys-Erweiterung</b>	<b>99</b>
7.1	Ergebnisse aus Testfällen und deren Bewertung . . . . .	99
7.1.1	Testfall: „office05“ . . . . .	99
7.1.2	Testfall: „laboratory04“ (1) . . . . .	100
7.1.3	Testfall „laboratory04“ (2) . . . . .	102
7.1.4	Testfall „lecture_hall02“ . . . . .	104
7.1.5	Testfall „classroom03“ . . . . .	106
7.1.6	Testfall „kitchen01“ . . . . .	108
7.2	Bewertung der Ontologiemodellierung . . . . .	110
7.3	Bewertung des Bildmaterials . . . . .	110
7.4	Bewertung der Glaubensmaße . . . . .	111
<b>8</b>	<b>Zusammenfassung</b>	<b>113</b>
<b>A</b>	<b>Literaturverzeichnis</b>	<b>117</b>
<b>B</b>	<b>Abbildungsverzeichnis</b>	<b>123</b>
<b>C</b>	<b>Tabellenverzeichnis</b>	<b>125</b>
<b>D</b>	<b>CD-Datenträger</b>	<b>127</b>
<b>E</b>	<b>Testbilder der Raumkonzeptinstanzen</b>	<b>129</b>
<b>F</b>	<b>Glaubenswerte von Raumkonzepten der Rauminstanzen (Expertenbefragung)</b>	<b>135</b>
<b>G</b>	<b>Relationen und Taxonomie der Beispielontologie</b>	<b>139</b>





# 1 Einleitung

Der Ausgangspunkt dieser Diplomarbeit ist die Untersuchung, ob und wie geschlossene Räume in Gebäuden durch die in ihnen enthaltenen Objekte modelliert werden können. Es soll untersucht werden, welche Beziehungen zwischen den Objekten und den Räumen bestehen und inwieweit hierdurch Raummuster klassifiziert werden können. Dabei werden Grundlagen der menschlichen Wahrnehmung von Objekten und Räumen berücksichtigt und im System modelliert.

Ein Aspekt bei der Modellierung ist die Kombination von unsicherem und ontologischem Wissen. Ivan Herman vom World Wide Web Consortium<sup>1</sup> forderte jüngst in einem Interview: „Wie sich nicht klar in ja/nein-Strukturen abzulegendes Wissen [in Ontologien] abbilden [lässt], ist ebenfalls noch zu bearbeiten“ (Herman, 2005). Dieser Fragestellung soll hier nachgegangen werden. Im Detail sollen im Rahmen dieser Untersuchung folgende Aspekte behandelt werden:

Die Klassifikation der Räume soll auf Blickpunktfolgen basieren und sich so am biologischen Sehen orientieren. Ein bestehendes System (Schill u. a., 2001) klassifiziert dabei einzelne Objekte anhand von Sakkadenfolgen, die auf dem Objekt ausgeführt werden. Hierauf soll die vorliegende Arbeit aufbauen und komplexere Objekte (in diesem Fall Räume) klassifizieren. Das System wird damit um eine Abstraktionsstufe erweitert: Durch imitierte Augenbewegungen werden die einzelnen Objekte im Raum nach und nach fixiert. Dabei wird bei jedem fixierten Punkt jeweils das damit aktuell fixierte Objekt klassifiziert, indem an dieser lokalen Stelle ein Sakkadenzug auf dem Objekt ausgeführt wird. Durch die Kombination dieser klassifizierten Objekte wird ein komplexeres Objekt, das Raumkonzept, klassifiziert. Das zusätzliche Wissen, das zur Klassifikation des Raumkonzepts nötig ist, wird in einer Ontologie repräsentiert.

Für die Modellierung dieser Ontologie werden die Objekte eines Raumes und verschiedene Arten von Räumen hinsichtlich ihrer partonomischen und taxonomischen Beziehungen untersucht. In der Ontologie werden die Raummuster mit den enthaltenen Objekten und deren Eigenschaften dargestellt. Beispielhaft soll dies an einer Ontologie für Universitätsräume demonstriert werden. Um den Umfang auf ein überschaubares Maß zu reduzieren und gleichzeitig eine hinreichende Zahl an zu differenzierenden Raummustern untersuchen und die Ergebnisse prüfen zu können, werden in der Onto-

---

<sup>1</sup><http://www.w3.org/>

logie fünf Raumkonzepte modelliert: Hörsaal, Seminarraum, Praktikumsraum, Büro und Küche.

Als Ontologie-Sprache wird OWL<sup>2</sup> gewählt. Die Sprache soll aber um die Möglichkeit erweitert werden, unsicheres Wissen repräsentieren zu können. Wie bei dem bestehenden System, bei dem die Folge der Blickpunkte mit dem Glaubensmaß der Dempster-Shafer-Theorie (Shafer, 1976) belegt wird, sollen auch die Raumkonzepte mit einem Glaubensmaß versehen werden. Dies liegt darin begründet, dass einem Raum selten nur aufgrund seiner enthaltenen Objekte ein eindeutiges Raumkonzept zuzuweisen ist. Vielmehr können Kombinationen von Objekten auf bestimmte Raumkonzepte hindeuten. Auch ist es möglich, dass eine solche Kombination auf mehrere, unterschiedliche Raumkonzepte schließen lässt. So sollen die Wahrscheinlichkeiten für die jeweiligen Raumkonzepte durch ein Glaubensmaß repräsentiert werden.

Die Möglichkeit, ontologisches und epistemologisches Wissen zu verknüpfen, ist bisher kaum untersucht worden. In erster Linie wird in Ontologien nur sicheres und statisches Wissen dargestellt (Puppe u. a., 2003). Doch fordert nicht nur die hier beschriebene Anwendung eine solche Kombination. Gerade durch den verbreiteten Einsatz von Ontologien (nicht zuletzt durch die Entwicklung des Semantic Web<sup>3</sup>) steigt das Interesse an einer Modellierung unsicheren Wissens in Ontologien. Aktuelle Entwicklungen solcher Kombinationen werden in dieser Arbeit, die selbst einen ersten Ansatz für eine mögliche Kombination vorstellt, dargestellt.

Im Implementierungsteil soll das bestehende System um die hier modellierte Ontologie erweitert werden. Raummuster sollen anhand von Blickpunktfolgen zwischen und auf einzelnen Objekten klassifiziert werden können. In der anschließenden Testphase sollen die Güte der Ergebnisse und die Vor- und Nachteile der Kombination von Ontologien mit unsicherem Wissen evaluiert werden.

Zum Einsatz könnte ein solches System beispielsweise in einem autonomen Roboter kommen, der sich durch Räume bewegen kann und den Raum von seiner Position aus betrachtet. Aus diesem Blickwinkel kann mit dem System untersucht werden, um welches Raumkonzept es sich aktuell handeln könnte und auf Basis dieses Wissens können weitere Aktionen folgen.

Denkbar wäre aber auch dieses System in einen Agenten zu integrieren, der Räume beispielsweise per Kamera betrachten kann. Aufgrund der analysierten Raumkonzepte auch in Kombination mit den analysierten Objekten könnten Metainformation über einen Raum erschlossen werden. Bei späteren Anfragen bezüglich der Verwendung und Verwendbarkeit von Räumen könnte dieses System entsprechend reagieren. In so einem

---

<sup>2</sup><http://www.w3.org/TR/2004/REC-owl-features-20040210/>

<sup>3</sup><http://www.w3.org/2001/sw/>

---

„intelligenten Gebäude“ wären beispielsweise unbelegte Räume in Erfahrung zu bringen, die für eine bestimmte Veranstaltung mit entsprechenden Präsentationsmöglichkeiten geeignet sind.

In den einzelnen Kapitel werden die folgenden Themen behandelt:

- In Kapitel 2 soll in die allgemeine Thematik eingeführt und der aktuelle Stand der Forschung zusammengefasst werden. Die beteiligten Fachgebiete werden näher vorgestellt und grundlegende Begriffe geklärt.
- In den Kapiteln 3 und 4 werden die Themengebiete *sakkadische Augenbewegungen* und *unsicheres Wissen und Ontologien* detaillierter betrachtet. Ausgehend von diesen Themen wird speziell Bezug auf die vorliegende Arbeit genommen.
- Die vorangegangenen Überlegungen werden in Kapitel 5 durch das konkrete Beispiel der Universitätsräume umgesetzt. In Kapitel 6 folgt die Dokumentation der Implementierung. In Kapitel 7 folgt die Testphase und eine Bewertung der Ergebnisse und der Umsetzung.
- Abschließend werden in Kapitel 8 die Schlussfolgerungen, die aus der Untersuchung gezogen werden, zusammengefasst. Es folgt ein Ausblick auf weiter zu untersuchende Aspekte und offene Fragen.



## 2 Stand der Forschung

In diesem Kapitel soll ein Überblick über den derzeitigen Stand der Forschung gegeben werden. Dabei sind in Bezug auf den hier behandelten Gegenstand vor allem folgende Bereiche von Interesse: Es soll die Thematik des biologischen Sehens näher betrachtet werden, in welcher Weise der Mensch Objekte wahrnimmt und welche technischen Verfahren bei der Bilderkennung bisher eingesetzt werden. In diesem Zusammenhang geht es auch um die Objekterkennung im Raum. Denn auch in der vorliegenden Arbeit sollen Objekte im Raum erkannt werden, wobei das Verfahren des menschlichen Sehens als Ausgangspunkt dient. Anschließend wird das Forschungsgebiet der Wissensrepräsentation näher betrachtet. Hier sind vor allem Ontologien von Interesse. Die erkannten Objekte in den zugehörigen Räumen sollen in einer Wissensbasis modelliert werden. Als weiterer Punkt schließt sich das Gebiet des unsicheren Wissens an, welche Theorien es gibt und welche Vor- und Nachteile sie haben. Insbesondere die gewählte Domäne der Raumkonzepte verlangt durch ihre Vielfalt und Unsicherheit nach entsprechenden Methodiken. Weiter soll darauf eingegangen werden, ob und inwieweit unsicheres Wissen mit Ontologien bisher miteinander verknüpft wurden und wie sie bei dem hier beschriebenen Thema kombiniert werden können. Abschließend werden die Themen Inferenz und Lernen kurz vorgestellt, soweit sie für diese Arbeit relevant sind.

### 2.1 Visuelle Wahrnehmung

Computergestützte Systeme der Bildanalyse, bei denen Objekte segmentiert und klassifiziert werden, basieren üblicherweise auf den Abläufen, die beim menschlichen Sehen zu beobachten sind. Daher soll an dieser Stelle ein kurzer Überblick über die visuelle Wahrnehmung des Menschen gegeben werden. Systeme, die auf solchen Vorgängen basieren, werden anschließend näher beschrieben (siehe Kapitel 2.2).

Die Leistungen des menschlichen visuellen Systems bestehen darin, das Netzhautbild zu stabilisieren, es zu segmentieren, dreidimensionale Gegebenheiten zu extrahieren, Objekte zu klassifizieren und zu identifizieren, bedeutungsvolle Szenen zu konstruieren und Blick- und Kopfbewegungen zu steuern (Roth, 1997).

Abbildung 2.1 (nach Hubel (1989)) stellt eine schematische Darstellung der Sehbahnen des menschlichen Gehirns dar. Die visuelle Informationserzeugung findet auf der

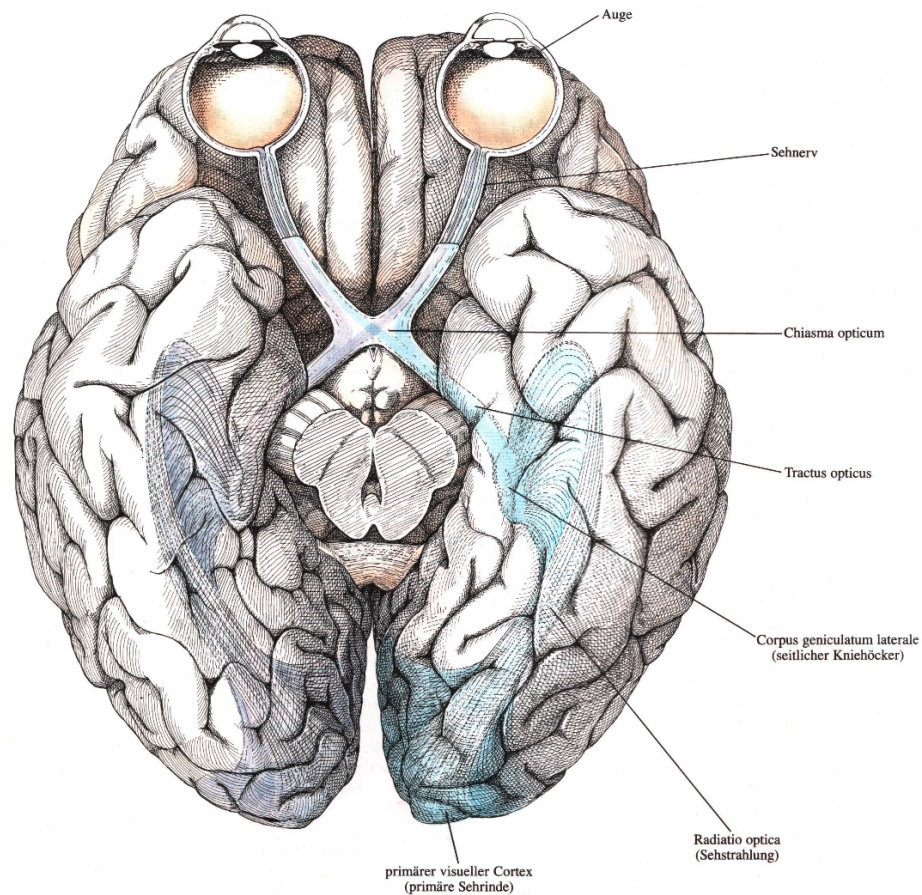


Abbildung 2.1: Visuelles Schema der Sehbahnen (nach Hubel (1989))

Netzhaut des Auges statt, auf der die Lichtreize eintreffen. Hier befinden sich drei Zellschichten: Photorezeptoren, Interneurone und Retinaganglienzellen, welche durch Faserschichten miteinander verknüpft sind. Auf diesen Schichten werden unterschiedliche Leistungen erbracht: Die Photorezeptoren bestehen aus lichtempfindlichen Stäbchen und farbempfindlichen Zapfen, die entsprechend Helligkeiten und Farben wahrnehmen können. Auf der Schicht der Interneurone befinden sich Horizontalzellen, Bipolarzellen und amakrine Zellen. Diese sind hauptsächlich für die Gestaltwahrnehmung zuständig: Größe und Richtung werden erkannt, wobei die visuelle Wahrnehmung begünstigt wird von starken Grauwertübergängen und ständigen Augenbewegungen (Sakkaden), mit denen Reize auf der Retina erzeugt werden (Abmayr, 1994).

Ein Teil der Retinaganglienzellen reagiert auf die Farbkontraste rot-grün bzw. blau-gelb. Andere reagieren auf geringe Helligkeitskontraste, räumliche Auflösung oder sich schnell bewegende Objekte. Auf der Schicht der Ganglienzellen vereinigen sich die Aktivitäten der Photorezeptoren und Interneurone, um zusammen mit den Aktivitäten der Ganglienzellen über Nervenbahnen an die Bereiche des Gehirns gesendet zu werden, in denen visuelle Informationen verarbeitet werden. Dabei kreuzen sich die Sehnerven im optischen Kreuz (Chiasma opticum), wo eine partielle paarweise Aufteilung stattfindet (Hubel, 1989).

Die Reize werden an verschiedene Bereiche des visuellen Cortex gesendet, welcher sich grob in drei Sehrinden zergliedern lässt. Die primäre Sehrinde beinhaltet eine komplette Abbildung der Retina, wobei die Fovea centralis, der Bereich des schärfsten Sehens, überrepräsentiert ist. Im Anschluss an den primären visuellen Cortex schließt sich die sekundäre Sehrinde an. Auch hier existiert eine Abbildung der Retina, die allerdings stark verzerrt ist. Den Abschluss bildet der tertiäre Cortex, auf dem ebenfalls in verzerrter Form ein Abbild des gesamten Gesichtsfeldes repräsentiert wird. Hier wird eher auf Orientierungs- als auf Farbreize reagiert. Weitere Areale im visuellen Cortex befassen sich mit der Farbwahrnehmung oder der Wahrnehmung von Richtung und Bewegung. Zusätzlich werden sensorische Informationen und motorisches Verhalten integriert. Insgesamt lässt sich allerdings nicht von einer strikten Hierarchie sprechen. Die circa 30 visuellen corticalen Areale sind miteinander verknüpft. In Untergruppen von ihnen finden eine Reihe verschiedener Verarbeitungsprozesse statt (Hubel, 1989).

### 2.1.1 Wahrnehmung von Objekten

Das bloße Erkennen von Licht und Farbe reicht für die Wahrnehmung von Objekten nicht aus. Weitere Verarbeitungsschritte sind notwendig, um zu erkennen, welche Regionen zusammengehören und ein Objekt bilden. Eine Kombination elementarer Merkmale muss zu einer sinnvollen Einheit zusammengefügt werden. Anhand welcher Kriterien sich eine solche Einheit zusammenfügt, wurde von Gestalt-Psychologen in den *Gestaltgesetzen der Wahrnehmungsorganisation* erstmals formuliert (Metzger, 1975). Demnach verläuft die Zergliederung einer visuellen Szene in einzelne Objekte im Gehirn nach diesen Prinzipien. Die Gestaltgesetze sind im einzelnen: das Gesetz der guten Gestalt, der Nähe, der Ähnlichkeit, der guten Fortsetzung, der Geschlossenheit und des gemeinsamen Schicksals (Metzger, 1975). Beim Erkennen von Objekten ist insbesondere die Form ausschlaggebend, während Farbeigenschaften eine eher geringe Rolle spielen. Daneben ist für die Separierung einzelner Elemente aber auch Erfahrung im Umgang mit der realen Welt entscheidend (Anderson, 2001). Durch Erfahrungen werden neue Informationen im Gedächtnis gespeichert. Diese sind dann für neue Wahrnehmung

verfügbar. Betrachtete Elemente können so mit ähnlichen bereits betrachteten Elementen assoziiert werden. Dabei reicht bereits eine kritische Menge an Daten, um Assoziationen herzustellen. Zudem gibt es unterschiedliche Grade der Klassenzugehörigkeit. Es gibt bestimmte Einschränkungen, die auf ein Element zutreffen müssen, damit eine Assoziation hergestellt werden kann. Gleichzeitig darf dieses Element aber auch gewisse Unterschiede aufweisen. Ein Mensch braucht nicht zu betrachten, ob ein Bär runde oder spitze Ohren hat, um zu erkennen, dass es sich um einen Bären und nicht um einen Hund handelt. Je öfter ein bestimmtes Objekt betrachtet wird, desto mehr automatisiert sich die Wiedererkennung. Das wahrgenommene Bild wird anhand von einfachster Eckdaten identifiziert. Ein solcher stark automatisierter Prozess läuft beispielsweise bei der Erkennung von bekannten Gesichtern ab (Goldenberg, 1998).

Über einen detaillierten Prozess der Speicherung von Bildern im Gehirn gibt es bisher in der Forschung keine einheitlichen Ergebnisse.

### 2.1.2 Raumwahrnehmung

„Wahrnehmung umfaßt zum einen die physikalischen, physiologischen und psychischen Prozesse, zum anderen das Resultat, den bewußten Erlebnisinhalt. Hier wiederum ist die einzelne Empfindung (...) von der komplexen Gesamtleistung der Wahrnehmung zu unterscheiden, zu der auch die Raumwahrnehmung gehört.“ (Ritter und Gründer, 1971ff)

Gerade bei der *Raumwahrnehmung* handelt es sich folglich um einen sehr komplexen Prozess, der sich aus einer Vielzahl von Kriterien zusammensetzt. Insbesondere findet er nicht nur im Gehirn statt sondern beruht auch auf sensomotorischer Wahrnehmung. Während der Schwerpunkt bei der visuellen Wahrnehmung noch auf dem Erkennen von Objekten liegt, erstreckt sich das Gebiet beim dreidimensionalen Raum von dem bloßen Erkennen einzelner Objekte und ihrer Lage im Raum bis hin zu der Navigation durch den Raum und der Manipulation von Gegenständen in diesem Raum. Des weiteren ließ sich durch eine Reihe von Experimenten feststellen, dass die räumliche Wahrnehmung nicht an die visuelle gebunden ist. Beide Bereiche haben unterschiedliche Repräsentationen im Gehirn (Knauff, 1997).

Wie räumliche Informationen im Gehirn dargestellt werden, wird häufig unter dem Titel *kognitive (Land-)Karte* (engl.: cognitive map) diskutiert und untersucht (Mallot, 2004). Der Leitgedanke bei diesen Untersuchungen besteht darin, dass die räumliche Repräsentation einer Landkarte ähnelt, auf der die räumlichen Verhältnisse der Umwelt dargestellt sind. Neben den Objekten und ihren Positionen werden demnach auch die räumlichen Relationen zwischen diesen Objekten gespeichert. Doch ob die Speicherung von Ortsinformationen im Gehirn einer Karte gleicht, ist bis heute nicht bewiesen.



Zwar sind Karten in unserem Kulturkreis eine natürliche und gleichzeitig sehr alte Art der Ortsrepräsentation, doch würde eine Karte viel Speicherplatz benötigen und Menschen fällt es oft schwer, Karten zu lesen (Mallot, 2004).

Zudem können Objekte im Raum die Funktion sogenannter *landmarks* (Orientierungspunkte) einnehmen und dadurch zusätzliche Eigenschaften bekommen. Diese Punkte bieten zusätzliches Wissen zum Raum- und Distanzwissen. Sie dienen der Orientierung und als Referenzpunkte. Zudem stellt die räumliche Organisation von Objekten ein wesentliches Element des Schemawissens, also Konzepten, von Szenen dar (Engelkamp, 1990).

„Die Art der Repräsentation räumlicher Sachverhalte ist keineswegs nur eine Funktion des empirischen Sachverhalts und des Wahrnehmungsinhalts, sondern wird maßgeblich durch Vorerfahrung, Erwartung sowie konzeptuellem Wissen beeinflusst“ (Knauff, 1997). Dabei wird das konzeptuelle Wissen aus unterschiedlich vielen Erfahrungen einheitlich zusammengefasst. In ihm bilden sich hierarchische Strukturen. Diese gilt auch für Raumkonzepte. So können diese beispielsweise anhand ihrer Funktionalität (Küche, Bad, Wohnraum etc.) zusammengefasst werden. Wird den Probanden ein typisches Raumkonzept dargeboten, so ist die Erinnerung an einzelne Gegenstände größer als bei untypischen Raumkonzepten. Beispielsweise erinnern sich Versuchspersonen eher an ein Sofa in einem Wohnzimmer als an eine Dusche. So hat das Gehirn eine Erwartungshaltung an die aktuelle Szene. Bestimmten Raumkonzepten werden spezifische Objekte zugeordnet, die in diesem Raum erwartet werden (Engelkamp, 1990).

In der vorliegenden Untersuchung soll das Erkennen von Objekten anhand von Sakkaden erfolgen. Wie oben bereits erwähnt, spielen diese kontinuierlichen Augenbewegungen bei der visuellen Wahrnehmung eine entscheidende Rolle. Durch Sakkaden werden auf der Netzhaut Reize erzeugt und damit Nervenzellen für die Verarbeitung der Signale im Gehirn angeregt. Die Idee, ausgehend von neurobiologischen Erkenntnissen ein Modell zu generieren, das mittels Sakkaden Objekte erkennen kann, ist bereits erfolgreich umgesetzt (siehe Kapitel 3.3). Hierauf aufbauend sollen komplexere Objekte - in diesem Fall Raumkonzepte - modelliert und identifiziert werden. Dabei bedient sich diese Modellierung der Tatsache, dass einem Raumkonzept spezielle Objekte zugeordnet werden können. Hier stellt sich die Frage, ob sich Erkenntnisse aus der Raumwahrnehmung mit der Analyse von 2D-Eingabebildern verknüpfen lassen. Wie genau dies realisiert werden kann, wird in Kapitel 3 beschrieben.

## 2.2 Bildanalyse

Das Forschungsgebiet des *Bildverstehens*, auch genannt wissensbasierte Bildanalyse, befasst sich mit der Analyse und Interpretation von visuellen Informationen. Es beschäftigt sich damit, wie biologische Sehsysteme maschinell nachgebildet und damit technische Probleme gelöst werden können. Auf dem Gebiet des Bildverstehens gibt es heute verschiedene konzeptuelle Sichten, die sich aus der historischen Entwicklung dieses Fachgebiets entwickelt haben (Neumann, 2003).

### 2.2.1 Konzepte der Bildverarbeitung

Das Gebiet der *Bildverarbeitung* beschäftigt sich mit Methoden und Techniken, die zur Transformation, Modifikation oder Codierung von natürlichen Bildern eingesetzt werden. Ein Teil dieser Techniken dient dazu, aus vorhandenem Bildmaterial neue Bilder zu generieren, die zusätzliche Informationen für die visuelle Auswertung liefern können. Das *Bildverstehen* hingegen wird genutzt, um Bilder automatisch zu erkennen. In Sehsystemen findet man häufig eine Kombination aus Bildverarbeitung und Bildverstehen (Abmayr, 1994).

Anfangs wurde im Bereich der *Mustererkennung* (engl. pattern recognition) vor allem die Klassifikation von Buchstaben untersucht. Der klassische Ablauf eines Bildverarbeitungssystems gliedert sich in folgende Abschnitte: Ein Eingangssignal (das Bild) wird erfasst und vorverarbeitet. Hierbei werden beispielsweise Rauschpunkte entfernt. Daran schließt die Segmentierung an. Basierend auf dieser Segmentierung werden signifikante Merkmale extrahiert. Anhand dieser Merkmale, die üblicherweise durch einen Merkmalsvektor beschrieben werden, kann die Bildstruktur in Klassen eingeordnet werden. Anhand der Lage im Merkmalsraum erfolgt dann die Klassifikation (Abmayr, 1994).

Diese Methode wird auch heute noch erfolgreich eingesetzt in Anwendungsbereichen, in denen mit einer überschaubaren Anzahl von Klassen gearbeitet wird und Eingabedaten aus 2D-Objekten bestehen. Doch für die Verarbeitung von Eingabebildern, die eine dreidimensionale Szene einer vielfältigen Domäne darstellen, haben sich merkmalsbasierte Systeme als unbrauchbar erwiesen (Neumann, 2003).

Eine andere Idee wurde mit der Analyse von Polyederszenen verfolgt (Waltz, 1975). Dabei werden Blockswelten mittels eines Markierungsalgorithmus anhand ihrer geometrischen Merkmale beschrieben. Doch Realweltszenarien können mit Polyedermodellen nicht hinreichend genug beschrieben werden, auch Beleuchtungseffekte stellen Probleme dar.

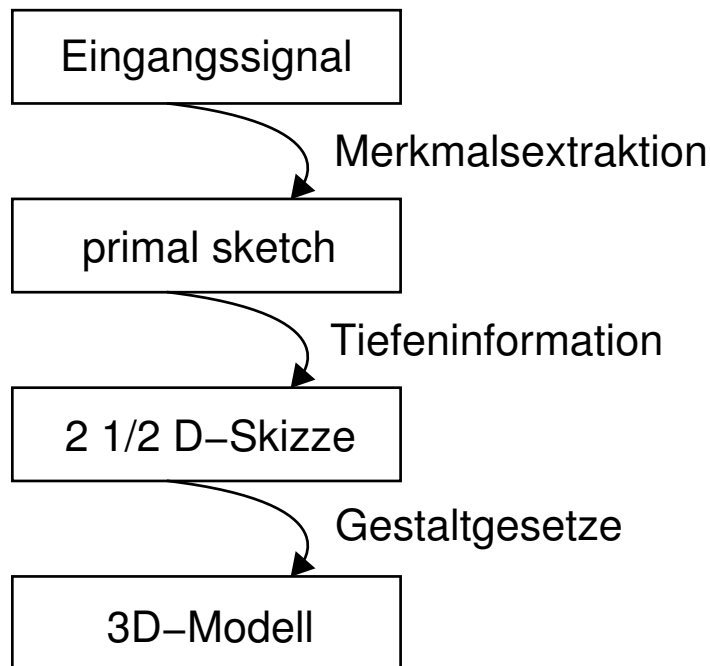


Abbildung 2.2: Ablauf visueller Objektwahrnehmung (nach Marr (1982))

Ein weiterer Ansatz basiert darauf, dass komplexe Objekte aus einfachen Komponenten kombiniert werden können. Bekannte Objekte können beispielsweise aus einer Kombination von einfachen röhrenförmigen Komponenten bestehen (Marr, 1982). Marrs System orientiert sich dabei an biologischen Sehsystemen, die aus einer hierarchischen Verarbeitungsstruktur bestehen. Nach Marr lässt sich der Ablauf bis zur visuellen Objektwahrnehmung in drei Schritten skizzieren, die in Abbildung 2.2 dargestellt sind: Ein Eingangssignal, der Lichtreiz, trifft auf die Netzhaut und wird dort in neuronale Signale umgewandelt. Hier werden Informationen über geometrische Gegebenheiten und Helligkeitsunterschiede extrahiert, um die *Primärskizze* (primal sketch) zu erhalten. In Kombination mit der Tiefeninformation, die Aufschluss über die Lagebeziehungen der Elemente zueinander und zum Betrachter gibt, ergibt sich laut Marr eine 2 1/2D-Skizze. Die einzelnen Objekte dieser Skizze werden mittels der Gesetze zur Wahrnehmungsorganisation differenziert, hierarchisch angeordnet und ein 3D-Modell wird generiert. Verknüpft mit Kontextinformationen können die Objekterkennung und weitere komplexe kognitive Prozesse folgen (Marr, 1982).

Auch wenn die bei Marr beschriebene Architektur zum Teil nicht mit biologischen Fakten vereinbar ist, so führte sein Ansatz dazu, dass sich auf dem Gebiet des Bildver-

stehens verstärkt an kognitiven Erkenntnissen orientiert wurde (Neumann, 2003).

Neuere Methoden lassen sich unter dem Begriff „*Active Vision*“ (Bajcsy, 1988) zusammenfassen. Auch diese Methode basiert auf biologisch und kognitiv orientierten Erkenntnissen. Visuelles Sehen ist nicht nur eine passive Wahrnehmung der Umwelt. Vielmehr bewegen wir unsere Augen und unseren Kopf, fokussieren Objekte und betrachten sie aus unterschiedlichen Blickwinkeln.

Die Theorie basiert auf lokalen und globalen Komponenten. Dabei arbeiten die lokalen Komponenten mit internen Parametern, beispielsweise mit Werten für die Breite eines Bandpassfilters bei der Kantendetektion. Diese Parameter bestimmen den Bereich für mögliche Messwerte und deren Ungenauigkeiten. Globale Komponenten werten die Kombination der lokalen Komponenten aus und entscheiden das weitere Verfahren. Der Grundgedanke hierbei besteht darin, dass Feedback an die globalen Komponenten geliefert wird und verschiedene Daten gesammelt werden, die zu einer Optimierung des Gesamtablaufs führen.

Diese Methode dient nicht dazu, die komplette Szene ähnlich Marrs Ansatz zu beschreiben, sondern zielgerichtetes Handeln zu unterstützen. Auch die vorliegende Arbeit orientiert sich an dieser Idee. Zielgerichtete Steuerung der Augenbewegungen soll Aufschluss über die zu untersuchende Szene geben.

Schließlich sind noch die konnektionistischen Modellbildungen zu erwähnen, die eine weitere Alternative zu dem Sehsystem von Marr bilden. Der Vorteil liegt hier vor allem in der massiven Parallelverarbeitung (Neumann, 2003).

Der Anwendungsbereich der Bildverarbeitung ist vielfältig. Doch für alle Anwendungen gilt, dass sie spezifisch auf ihren Anwendungsbereich zugeschnitten sind. Eine universelle Umsetzung des Sehsystems konnte bisher nicht umgesetzt werden (Hermes, 2005).

Das in der vorliegenden Arbeit entwickelte System hat als Anwendungsgebiet Raummuster. Auch dieses Verfahren beruht auf kognitiven, biologischen und sensomotorischen Erkenntnissen. Die Idee besteht darin, durch Sakkadenfolgen Rückschlüsse auf das betrachtete Objekt zu ziehen, so dass man auch hier von aktiven Prozessen sprechen kann. Diese Möglichkeit wurde bisher für einzelne Objekte erfolgreich realisiert (Schill u. a., 2001). In diesem System, werden Bottom-up- und Top-down-Prozesse der Wahrnehmung miteinander verknüpft. Während der Bottom-up-Phase wird das zu untersuchende Objekt analysiert, um mögliche Kandidaten für Fixationspunkte zu identifizieren. Im Laufe dieser Vorverarbeitung werden sowohl lineare als auch nichtlineare Filter eingesetzt. In der Phase des Top-down-Prozesses werden die Augenbewegungen analysiert, mit einer erlernten Wissensstruktur verglichen und anhand eines Evidenzwerts die Klassifikation generiert. Nähere Beschreibungen zu diesem System folgen in Kapitel 3.

### 2.2.2 Prozessablauf des Bildverstehens

Wie oben bereits erwähnt, geht es beim Bildverstehen um die Deutung einer Szene anhand von Bildern. Notwendig hierfür sind also Bilder als Eingabe, auf denen Objekte der Szene abgebildet sind. Die Ausgabe besteht aus der computerinternen Beschreibung der Szene. Szenen setzen sich aus zeitlich-räumlichen Ausschnitten der Realwelt zusammen. Dabei sind die Eingabebilder in der Regel zweidimensionale Abbildungen einer Szene. Die Beschreibung der Szene gibt Auskunft darüber, welche Objekte sich wann und wo im Bild befinden (Neumann, 2003).

Beim Bildverstehen handelt es sich häufig um eine zielgerichtete Aktivität. „Zur Steuerung der einzelnen Prozesse gehört ein Handlungskontext, insbesondere Aufmerksamkeitskontrolle, Fokussierungsmechanismen, Bewegungsverfolgung, sowie das Zusammenwirken von Grob- und Feinanalyse“ (Neumann, 2003). Dabei stellt sich der Prozess des Bildverstehens als ein komplexes System zahlreicher Teilkomponenten dar. Diese Teilkomponenten können sehr vielfältig und unterschiedlich abstrahiert sein. Abbildung 2.3 stellt einen konzeptuellen Rahmen für ein Sehsystem dar (nach Niemann (1985)).

Im wesentlichen enthalten diese Systeme vier Stufen. In einer Vorverarbeitungsphase werden Rohbilder der Szene, die als Eingangsdaten dienen, untersucht. Auf dieser ersten Stufe erfolgen somit Berechnungen unterschiedlicher Filterprozesse. Diese Bottom-up-Ergebnisse bilden die Ausgangslage für die zweite Stufe, Bildelemente, die als Objekte in der realen Szene auftreten, zu extrahieren und als Szenenelemente zu interpretieren. Dabei werden primitive Objekte separiert, denen Eigenschaften zugewiesen werden können. In der dritten Stufe findet die Objekterkennung durch die Bilddaten, die erkannten Szenenelemente und durch das Vorwissen über mögliche Objektmodelle statt. Abhängig von der Wissensbasis können Objekte aus unterschiedlichen Blickwinkeln, Größen und Eigenschaften ermittelt werden. Hier wird das Bild aus der Top-down-Sicht behandelt. Weitergehende Verarbeitungsschritte gehören zur vierten Stufe. Hier können komplexe Zusammenhänge der Szene erkannt werden. Ergibt die Analyse der Szene beispielsweise die Objekte Tafel, Stühle, Tische, Projektor, kann mit Hilfe einer vorhandenen Wissensbasis auf einen Vorlesungssaal geschlossen werden. Von der vierten Stufe aus wird auch der Gesamtprozess überwacht. Abhängig von dem Wissen werden Verarbeitungsschritte bestimmt.

Allgemein werden beim Bildverstehen folglich Bottom-up- und Top-down-Prozesse kombiniert. Zusätzlich bedient man sich einer Wissensbasis für weitere Informationen, die nicht aus der Szene direkt ableitbar sind. Möglichkeiten der Wissensmodellierung werden im nächsten Abschnitt beschrieben.

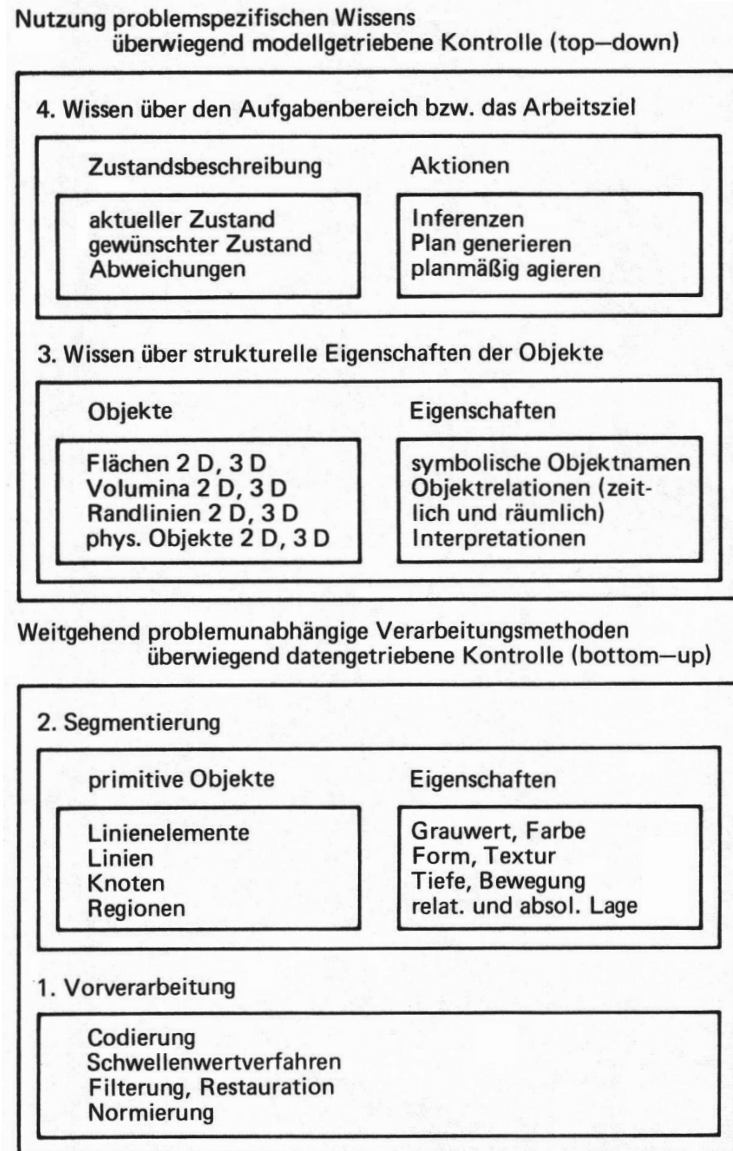


Abbildung 2.3: Prozessablauf des Analysesystems (nach Niemann (1985))

## 2.3 Wissensrepräsentation und Wissensverarbeitung

Im Bereich der künstlichen Intelligenz versteht man unter dem Begriff der *Wissensrepräsentation und -verarbeitung* (engl. knowledge engineering) die Erfassung, Verwaltung, Verwendung und Transformation von Aussagen und Kenntnissen über einen relevanten Ausschnitt der Welt (Puppe u. a., 2003).

Der relevante Ausschnitt hängt dabei von der Domäne ab, die modelliert werden soll. Die Modellierung kann üblicherweise auf weitere Bereiche beliebig erweitert werden. Doch abhängig von der Funktion, die das System ausführen soll, liegt es nahe, den Bereich zur Komplexitätsvermeidung anwendungsbezogen zu lassen. Auch die Tiefe der Struktur kann der Aufgabe angepasst werden. Ein weiterer Punkt, der bei der Modellierung zu betrachten ist, ist die inferenzielle Kompetenz (Owsnicki-Klewe u. a., 2003): Die Modellierung sollte Schlussfolgerungen oder Problemlösen durch ihre Struktur unterstützen.

Wissensrepräsentationen als Basis zum Lösen von Problemen entstanden ursprünglich aus der Idee heraus, dass diese Vorgehensweise ähnlich der des Menschen sei, da Menschen ihr Wissen nutzen, um Probleme zu lösen. Aus heutiger Sicht ist diese Übertragung nicht mehr uneingeschränkt haltbar, da das menschliche Gehirn nicht wie ein reiner Interpretierer von Produktionsregeln funktioniert. Vielmehr wird hier Wissen in neue Strukturen zergliedert, die für das Problem geeigneter scheinen, durch Sprechen und Vorstellen werden neue Erkenntnisse gewonnen. So dienen Wissensrepräsentationen heute vor allem dazu das Wissen von Experten zu modellieren und damit Konstruktionen für Problemlösungen vorzugeben. So steht Knowledge Engineering üblicherweise in Zusammenhang mit maschineller Wissensverarbeitung. Diese Maschine soll neben der reinen Wissensrepräsentation dieses Wissen nicht nur erwerben, sondern mit Hilfe dieses Wissens weitere Erkenntnisse erschließen oder Probleme lösen können. Dabei spricht man von *wissensbasierten Systemen* oder *Expertensystemen* (Puppe u. a., 2003).

### 2.3.1 Vorgehensweise bei der Modellierung wissensbasierter Systeme

Dem Modellierungsprozess liegen syntaktische und semantische Normierungen zugrunde. So wird das repräsentierte Wissen in einer formalen Sprache verfasst, der eine bestimmte Semantik zugewiesen wird. Für den Vorgang der Wissensmodellierung gibt es in der Literatur unterschiedliche Vorschläge. Zusammengefasst lassen sie sich in folgende Phasen gliedern (nach Russell und Norvig (2003)):

1. *Problemidentifikation*: Das Aufgabengebiet wird analysiert. Die Aufgaben, die

eine Wissensbasis erfüllen soll, werden hier definiert.

2. *Wissensakquisition*: Sie dient dem Zusammentragen des vorhandenen Wissens über die Domäne. Dabei gibt es verschiedene Vorgehensweisen, angefangen bei der Expertenbefragung mittels unterschiedlicher Interview-Methoden oder Protokollierung von Tests mit Versuchspersonen über Literaturrecherche und -analyse bis hin zu Wissenserwerb aus Daten bzw. Data-Mining-Methoden.
3. *Modellwahl*: Da das Wissen bisher nur informell vorhanden ist, wird in dieser Phase ein Modell für die Repräsentation ausgewählt, das die Art der Wissensabbildung bestimmt. Die Wahl hängt meist vom Einsatzgebiet des Systems ab.
4. *Abbildung des Wissens*: Das erworbene, konzeptionelle Wissen wird nun in das im vorherigen Schritt festgelegte Format gebracht.
5. *Abbildung des Zustands*: Hier werden Instanzen von den modellierten Konzepten in die Wissensbasis eingepflegt.
6. *Testen und Wartung*: Die Testphase bildet den Prozessabschluss. Im laufenden Betrieb kann dann das System gewartet und gegebenenfalls aktualisiert und verändert werden.

An diesen Phasen orientiert sich auch die vorliegende Arbeit. Hier liegt die Aufgabe darin, Räume anhand der erkannten Strukturen (der enthaltenen Objekte) zu klassifizieren. Die Aufarbeitung des Wissens über die Raumkonzepte geschieht dabei anhand der Datensätze der Räume, die zum Teil empirisch zusammengetragen werden. Modelliert wird dieses Wissen in einer *Ontologie*. Die Integration von Instanzen und die Testphase schließen sich daran an.

### 2.3.2 Ontologien

Im Bereich der künstlichen Intelligenz wird in der Literatur am häufigsten folgende Definition für den Begriff einer Ontologie herangezogen: „An ontology is an explicit specification of a conceptualization“ (Gruber, 1993). Später wurde diese Definition noch präzisiert durch: „An ontology is a formal, explicit specification of a shared conceptualisation.“ (Studer u. a., 1998). Eine Ontologie definiert demnach Konzepte und deren Beziehungen, die in einer Domäne auftreten können. Zudem ist sie in einer formalen Sprache explizit definiert und bildet die Basis für eine geteilte Sicht auf die Domäne, so dass über diese kommuniziert werden kann (Uchold und Grüninger, 1996).



Doch ist die Verwendung des Begriffs Ontologie problematisch. Angelehnt ist er an den philosophischen Begriff der Ontologie, womit das Gebiet bezeichnet wird, das sich mit der Frage, was existiert, bzw. mit dem *Seienden* beschäftigt, und auf Aristoteles zurückgeht (Ferber, 1999). Andere Fachdisziplinen haben wiederum ihren eigenen Begriff einer Ontologie geprägt und unterschieden sich voneinander, insbesondere hinsichtlich dessen, was modelliert werden kann. Diese verschiedenen Sichtweisen werden in Kapitel 4.1.1 näher untersucht und spielen bei der praktischen Umsetzung der Ontologie-Modellierung in Kapitel 5 nochmals eine Rolle.

Im Hinblick allerdings auf die Verwendung von Ontologien im Bereich der künstlichen Intelligenz soll im Folgenden ein Überblick über nähere Eigenschaften und Kategorisierungsmöglichkeiten von Ontologien gegeben werden. Eine standardisierte Kategorisierung gibt es dabei nicht.

Ontologien bieten die Möglichkeit, Konzepte zu definieren, die sich durch bestimmte Eigenschaften auszeichnen, die ebenfalls modelliert werden können. Diese Konzepte repräsentieren den modellierten Ausschnitt der Welt. Sie können abstrakt oder konkret, elementar oder zusammengesetzt, real oder fiktiv sein. „In short, a concept can be anything about which something is said, and, therefore, could also be the description of a task, function, action, strategy, reasoning process, etc.“ (Corcho und Gómez-Pérez, 2000). Des Weiteren können Beziehungen zwischen Konzepten existieren, welche zum Beispiel reflexiv, symmetrisch und transitiv sein können. Durch diese Beziehungen, die eine Klasse haben kann, können unter anderem partonomischen Aspekte modelliert werden. Durch hierarchische Beziehungen können Vererbungsstrukturen modelliert werden. Diese Vererbungshierarchie stellt den taxonomischen Aspekt der Ontologie dar.

Wohlgeformtheit der Taxonomie einer Ontologie wird durch drei Eigenschaften bedingt (nach Smith (2005)):

1. Die Taxonomie sollte in eine azyklische Baumstruktur gegliedert sein. Ein Beispiel hierfür ist ein Ausschnitt nach Gómez-Pérez u. a. (2004) aus UMLS<sup>1</sup> (Unified Medical Language System), einer domänenspezifischen Ontologie medizinischer Fachbegriffe (siehe Abb. 2.4). Die Verbindungen stellen dabei die Unter-/Oberklassenrelationen dar. Unterklassen stellen eine Spezialisierung ihrer Oberklasse dar und stehen in einer sogenannten „is-a“-Relation.
2. Die Endknoten sollten die niedrigste Kategorisierungsstufe darstellen, das heißt diese Knoten können nicht in weitere Unterklassen unterteilt werden und stellen die atomaren Konzepte der Ontologie dar.

---

<sup>1</sup><http://www.nlm.nih.gov/research/umls/>

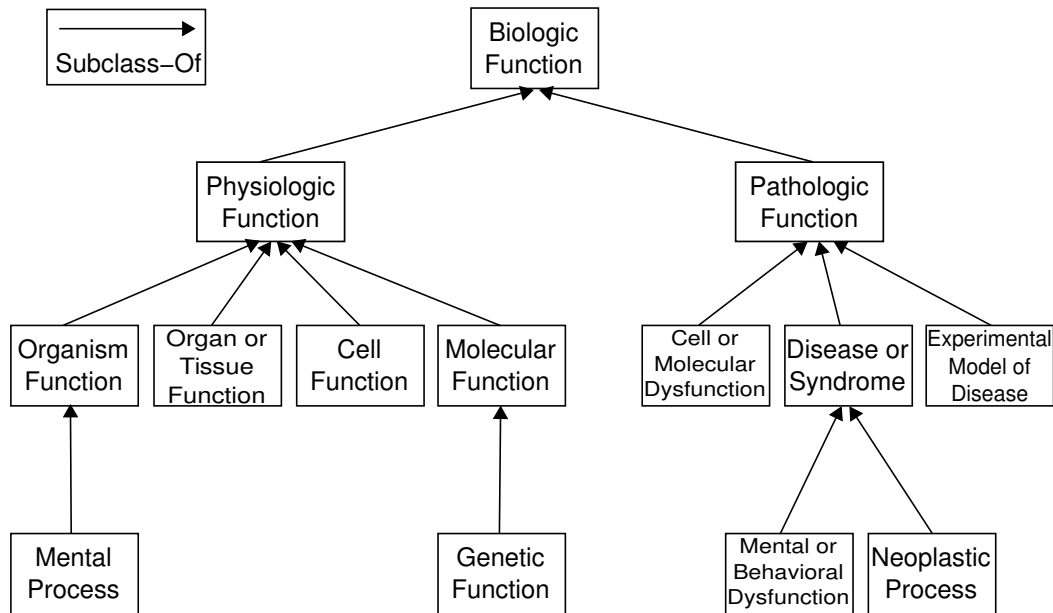


Abbildung 2.4: Ausschnitt aus der Taxonomie von UMLS (nach Gómez-Pérez u. a. (2004))

3. Die Taxonomie sollte *vereinigt* (engl. unified) sein, das heißt es existiert ein Konzept, welches Oberklasse aller anderen Konzepte ist.

Entgegen der zweiten Forderungen gibt es jedoch auch Taxonomien, die wie ein Verband modelliert sind und neben dem Konzept, das Oberklasse aller anderen ist, auch ein Konzept haben, das Unterklasse aller anderen ist. So eine Struktur findet sich beispielsweise in der Top-Level-Ontologie (siehe dazu Kapitel 4.1.1) nach Sowa (2000).

Wie formal das Vokabular einer Ontologie ausfällt, hängt von der verwendeten Sprache ab, in der die Ontologie verfasst wird. Diese kann von informell (beispielsweise natürliche Sprache) über semi-informell (Thesauri) und semi-formal (XML) bis formal (Prädikatenlogik) variieren (Uschold und Grüninger, 1996). Hiervon hängt auch die Ausdrucksstärke einer Ontologie und der Unterschied zwischen *heavyweight*- von *lightweight*-Ontologien ab (Gómez-Pérez u. a., 2004).

Werden Ontologien formal in Beschreibungslogiken verfasst, so beinhalten sie terminologisches Wissen (*TBox*), womit die Begriffsdefinitionen der Konzepte bezeichnet werden, und assertorisches Wissen (*ABox*), womit die Instanzen der Konzepte definiert werden (Puppe u. a., 2003). Auch die Ontologiemodellierungssprache *OWL DL* (siehe Kapitel 4.1.3) ist beschreibungslogikbasiert.

Zusätzlich spricht man bei der Ontologiemodellierung von *Ontological Commitment*, also der Verpflichtung, dass die verwendeten Konzepte und Begriffe innerhalb der Ontologie konsistent und kohärent verwendet werden, womit allerdings nichts über die Vollständigkeit ausgesagt wird (Gómez-Pérez u. a., 2004).

In der vorliegenden Arbeit wurde sich bei der Modellierung der Wissensbasis für eine Ontologie entschieden, da diese besonders gut die gewünschten Modellierungen von Partonomie und Taxonomie unterstützt. Ein weiterer interessanter Aspekt ist dabei die Verknüpfung mit unsicherem Wissen (siehe auch Kapitel 2.5).

## 2.4 Unsicheres Wissen

Im vorherigen Abschnitt wurde dargestellt, wie Wissen modelliert werden kann. Bisher wurde allerdings nur präzises und sicheres Wissen betrachtet. Doch trifft man gerade in der Realität häufig auf ungenaues oder vages Wissen. Dennoch sind Menschen in der Lage, Erkenntnisse hieraus zu ziehen und Probleme zu lösen. So hat sich im Bereich der künstlichen Intelligenz das Forschungsgebiet, das sich mit *unsicherem Wissen* (engl. *uncertain knowledge*, *uncertainty*) befasst, entwickelt. Auch Softwaresysteme sollen mit unsicherem Wissen umgehen und trotz dieser Unsicherheit Lösungsvorschläge liefern können.

Bei unsicherem Wissen unterscheidet man drei verschiedene Wissensenseigenschaften: Impräzision, Vagheit und Unsicherheit (Borgelt u. a., 2003). Impräzises Wissen ordnet einer Aussage keinen Wert sondern eine Menge von Wertemöglichkeiten zu. Eine Aussage hierfür ist beispielsweise „Das Leuchtmittel kann eine LED oder Neonröhre oder Halogenbirne sein“. Solche Aussagen können mit Mitteln der symbolischen Logik behandelt werden. Bei vagen Aussagen zeichnen sich die Aussagen durch ihre Unschärfe aus. Sie geben keine Wertemöglichkeiten an sondern einen nicht klar definierten Bereich. Eine Aussage für vages Wissen wäre zum Beispiel „Das Leuchtmittel ist hell“. Zusätzlich existieren auch Aussagen, deren Wertebereich nicht definitiv bekannt ist. In diesem Fall kann nicht wie im obigen Beispiel ausgeschlossen werden, dass das Leuchtmittel keine Energiesparbirne ist.

Es wurde bereits erwähnt, dass wissensbasierte Systeme ihr Wissen nutzen, um daraus neue Erkenntnisse abzuleiten. Das Wissen kann nun in als Wahrheit, Wahrscheinlichkeit, Möglichkeit oder einem Grad der Möglichkeit auftreten (Borgelt u. a., 2003). Wie erwähnt, wird aus diesem Wissen neues Wissen erschlossen. Bei unsicherem Wissen funktionieren die Mittel der deduktiven Logik jedoch häufig nicht mehr. Um mit unsicherem Wissen umgehen zu können, werden andere Lösungsstrategien benötigt. Die Art der Strategie hängt dabei von der Art der Unbestimmtheit des Wissens ab.

Einige solcher Strategien sollen im Folgenden kurz vorgestellt werden.

### 2.4.1 Wahrscheinlichkeitsbasierte Theorie

Bei *wahrscheinlichkeitsbasierten Theorien* (engl. Probability Theory) wird versucht, Schlussfolgerungen anhand von Wahrscheinlichkeiten zu ziehen (engl. Probabilistic Reasoning). Probability wird dabei als *Grad der Zusicherung* (engl. confirmation) definiert (Bhatnagar und Kanal, 1986). Es ist bekannt, welche Ereignisse (Aussagen) eintreten können. Welches der Ereignisse schließlich eintritt, ist aber unsicher. Es existiert lediglich eine Wahrscheinlichkeit des Eintretens. Jeder Aussage wird ein numerischer Wert zugewiesen. Dieser liegt üblicherweise zwischen 0 und 1 und legt damit fest, wie wahrscheinlich (nahe dem Wert 1) oder unwahrscheinlich (nahe dem Wert 0) eine Aussage ist. Auch den Folgerungen aus den Aussagen können so Wahrscheinlichkeiten zugewiesen werden. Die ermittelte Lösungsmöglichkeit hängt dann von der Höhe des Wertes ab (Aliev u. a., 2000).

**Bayes-Theorie** Die Bayes-Theorie ist ein spezielles Verfahren des Probabilistic Reasonings. Hier wird die Unsicherheit durch bedingte und unbedingte Wahrscheinlichkeiten repräsentiert. Die Idee liegt darin, zu ermitteln, wie hoch die Wahrscheinlichkeit für eine Aussage ist unter der Bedingung der Wahrscheinlichkeit einer anderen Aussage. Dabei wird die sogenannte *Bayes' Rule* angewendet:

$$\mathbf{P}(Y|X) = \frac{\mathbf{P}(X|Y)\mathbf{P}(Y)}{\mathbf{P}(X)} \quad (2.1)$$

$\mathbf{P}(A)$  ist die nichtbedingte Wahrscheinlichkeit von A,  $\mathbf{P}(A|B)$  die bedingte Wahrscheinlichkeit von A unter der Bedingung, dass B mit einer bestimmten Wahrscheinlichkeit gilt. Diese Regel ist für beliebig viele Aussagen erweiterbar, wodurch sich eine *Wahrscheinlichkeitsverteilung (probability distribution)* erstellen lässt.

Die Bayes-Theorie wurde zu Bayes'schen Netzen erweitert. Diese Netze bieten die Möglichkeit, das Wissen graphisch zu veranschaulichen. Durch einen gerichteten, azyklischen Graphen werden die obigen abhängigen Wahrscheinlichkeiten modelliert. Jeder Knoten entspricht einer Aussage und besitzt eine eigene Wahrscheinlichkeitsverteilung, die sich aus den Vaterknoten ergeben. Unidirektionale Kanten entsprechen der Aussage, dass der Startknoten direkten Einfluss auf den Zielknoten hat (Russell und Norvig, 2003).

Die Bayes-Theorie spielt in Kapitel 2.5.1 nochmals eine Rolle.

### 2.4.2 Fuzzy Logic

Mit der sogenannten *fuzzy set theory* besteht die Möglichkeit, vages Wissen zu modellieren. Anstelle von wahren oder falschen Aussagen spricht man hier von einem Grad der Richtigkeit (Zadeh und Kacprzyk, 1992). Das vorhandene Wissen ist *unscharf*. Hier wird beispielsweise nicht entschieden, ob eine Landschaft bewaldet ist oder nicht. Sie ist stattdessen zu einem gewissen Grad bewaldet. Dieser Grad wird durch einen Wert zwischen 0 und 1 angegeben.

Es ist allerdings zu beachten, dass es bei der Fuzzy Logic weniger um die Unsicherheit in der realen Welt geht. Vielmehr geht es darum, dass sprachliche Begriffe keine klaren Grenzen haben. Klassifikationen wie warm, kalt, groß oder klein gehören in diese Kategorie. Ist ein Objekt zu einem gewissen Grad groß, gehört es damit zu der Menge großer Objekte, die zusammen das *fuzzy set* ergeben (Russell und Norvig, 2003).

Der heutige Anwendungsbereich der Fuzzy Logic liegt vor allem in kommerziellen Industrieprodukten, beispielsweise Mikrowellen, Videokameras, Waschmaschinen etc. Der erfolgreiche Einsatz in diesem Bereich wird vor allem dem überschaubaren Regelsatz und den fuzzy logic-Operatoren, deren Funktionen für nahtlose Übergänge sorgen, zugeschrieben (Russell und Norvig, 2003).

In Kapitel 2.5.2 wird nochmals auf die Fuzzy Logic zurückgekommen.

### 2.4.3 Dempster-Shafer-Theorie

Die *Dempster-Shafer-Theorie* ist eine Weiterentwicklung der Probabilistic Theory. Die Grundlage entwickelte Arthur Dempster (Dempster, 1968), erweitert wurde sie durch Glenn Shafer (Shafer, 1976), so dass sich die heutige Bezeichnung Dempster-Shafer-Theorie etabliert hat. An manchen Stellen wird aber auch der Begriff *Belief-Theorie* verwendet.

Anstelle von Wahrscheinlichkeiten werden den Aussagen *Glaubenmaße* (engl. degrees of belief) zugewiesen. Das unsichere Wissen wird durch eine *Belief Function* repräsentiert, während beispielsweise bei der Bayes-Theorie eine Wahrscheinlichkeitsverteilung genutzt wird. Zudem können Beliefs über ganze Mengen von Aussagen getroffen werden und nicht nur über eine einzelne Aussage. Dies ist beispielsweise der Fall, wenn neue Informationen auf unterschiedliche Aussagen hindeuten können. Die Summe muss dabei nicht zwingend 1 betragen. Dem liegt zugrunde, dass einer Hypothese ein Teil des Beliefs zukommt, dass damit aber nicht automatisch der Negation der Hypothese der übrige Belief zugeordnet werden muss (Shafer, 1976).

Die Menge aller Aussagen einer Domäne ist in dem sogenannten *Frame of Discernment*  $\Theta$  dargestellt. Den Aussagen können Glaubensmaße zugewiesen werden. Diese werden

in der *Basic Probability Assignment*  $m : 2^\Theta \rightarrow [0, 1]$  zugewiesen. Den Elementen der Potenzmenge von  $\Theta$  werden also Werte zwischen 0 und 1 zugewiesen. Dabei gilt:

1.  $m(\emptyset) = 0$
2.  $\sum_{A \subset \Theta} m(A) = 1$

$m(A)$  wird dabei als „basic probability number“ (Shafer, 1976) bezeichnet und gibt das Glaubensmaß an, das man exakt  $A$  zuweist. Um mit der Basic Probability Assignment den gesamten Belief für  $A$  zu berechnen, der sich aus den Glaubensmaßen der einzelnen Teilmengen von  $A$  zusammensetzt, nutzt man die *Belief Function*

$$\text{Bel}(A) = \sum_{B \subset A} m(B) \quad (2.2)$$

über die Basic Probability Assignment. Besteht  $A$  aus einer Menge von Aussagen, so werden durch die Belief Function die Summe aller Teilmengen von  $A$  berechnet, die das gesamte Glaubensmaß für  $A$  angeben. Ist  $A$  eine atomare Aussage, so gilt  $m(A) = \text{Bel}(A)$ .

Belief Functions, die sich auf ein Frame of Discernment beziehen aber von unterschiedlichen Basic Probability Assignments stammen, lassen sich mit *Dempster's Rule of Combination* kombinieren. Mit dieser Regel wird eine neue Belief Function berechnet, in der das Wissen aus den unterschiedlichen Quellen zusammengetragen wird. Dabei kann eine Kombination  $\text{Bel}_1, \dots, \text{Bel}_n$  durch paarweise Verknüpfung  $((\dots(\text{Bel}_1 \oplus \text{Bel}_2) \dots) \oplus \text{Bel}_{n-1}) \dots \text{Bel}_n$  berechnet werden. Für die genaue Darstellung der Kombination zweier Belief Functions sei auf (Shafer, 1976), Kapitel 3, §1 verwiesen. An dieser Stelle soll stattdessen ein Beispiel (nach Shafer (1992)) für Belief Functions und deren Kombination folgen.

**Beispiel** Betty sagt aus, ein Ast sei auf Glenns Auto gefallen. Das Glaubensmaß von Glenn, dass Betty glaubwürdig ist, liegt bei 0.9, und dass sie nicht glaubwürdig ist, bei 0.1. Daraus ergibt sich Glenns Glaubensmaß, dass der Ast wirklich auf sein Auto gefallen ist, von 0.9, während das Glaubensmaß, dass dies nicht geschehen ist, 0 ist. Der Glaube, dass es nicht geschehen ist, liegt bei 0 und nicht bei 0.1, da sich Glenn nicht sicher sein kann, dass kein Ast auf sein Auto gefallen ist, auch wenn Betty nicht glaubwürdig ist. Es existiert insofern kein anderes Glaubensmaß, dass darauf hinweisen würde, dass kein Auto auf das Dach gefallen ist.

Nun kommen neue Aussagen durch eine weitere Person Sally hinzu. Diese gibt unabhängig von Betty ebenfalls an, dass ein Ast auf dem Auto gelandet ist, und auch ihre Glaubwürdigkeit liegt bei 0.9.

Trägt man nun die Glaubwürdigkeiten der unterschiedlichen Quellen zusammen, ergeben sich die Ergebnisse, die in Tabelle 2.1 dargestellt sind. Die Glaubensmaße werden dabei multipliziert. Es zeigt sich, dass der Belief dafür, dass beide Quellen glaubwürdig sind, bei 0.81 liegt. Der Belief, dass mindestens eine Quelle glaubwürdig ist, liegt demnach bei 0.99 ( $0.81 + 0.09 + 0.09$ ). Somit liegt der Belief, dass ein Ast auf das Auto gefallen ist, bei 0.99, da dies beide Quellen unabhängig voneinander ausgesagt haben und mindestens eine von ihnen glaubwürdig ist.

Tabelle 2.1: Kombination der Beliefwerte: Glaubwürdigkeit der Quellen

	Betty: glaubwürdig	Betty: unglaubwürdig
Sally: glaubwürdig	0.81	0.09
Sally: unglaubwürdig	0.09	0.01

Insgesamt wird der Belief für die Frage, ob ein Ast auf Glens Auto gefallen ist, aus dem Glaubensmaß bezüglich einer anderen Frage, nämlich wie glaubwürdig eine Quelle ist, generiert. In diesem Fall unterstützen (*confirm*) beide Quellen die Aussage. Würde Sally hingegen aussagen, es wäre kein Ast auf Glens Auto gefallen, so würde ihre Aussage die von Betty nicht unterstützen und es ergäben sich entsprechend geringere Beliefs für die jeweiligen Aussagen. Da diese nicht-unterstützenden (*non-confirming*) Beliefs in dieser Arbeit aber nicht benötigt werden, sollen sie hier auch nicht weiter behandelt werden.

**Dempster-Shafer-Theorie: Vor- und Nachteile** Es wird argumentiert, dass ein großer Vorteil dieser Theorie darin liege, dass explizit Unwissenheit repräsentiert werden kann. Auch ermögliche Dempster's Rule of Combination die Möglichkeit, neues Wissen mit altem zu verknüpfen bzw. die verschiedene Glaubensmaße zu vereinen (Bhatnagar und Kanal, 1986). Dieses *Belief Updating* wird zum Teil als Darstellung kognitiver Leistung angesehen, einer „Revision von Glaubensgraden“ (Plach, 1998).

Zudem wird im Gegensatz zu klassischen Wahrscheinlichkeitstheorien bei der Dempster-Shafer-Theorie zwischen positiven und negativen Hinweisen auf die Unterstützung einer Hypothese unterschieden. Wissen über entsprechende Informationen können hier berücksichtigt und verarbeitet werden (Plach, 1998).

Auch dass sich die Glaubensmaße der Hypothesen nicht zu 1 addieren müssen, ist ein großer Vorteil der Dempster-Shafer-Theorie. So werden nicht zwangsläufig der Negation

einer Hypothese Werte zugewiesen, obwohl man über diese noch keine Informationen besitzt.

Ein weiterer Vorteil liegt darin, dass nicht nur den atomaren Aussagen Glaubensmaße zugewiesen werden können. Insbesondere wenn nicht alle Glaubensmaße bekannt sind: Existieren beispielsweise keine Werte für atomare Aussagen sondern nur für eine Menge von atomaren Aussagen, so kann problemlos mit diesen Werten für weitere Beliefs gerechnet werden. Der Belief muss dabei nicht unter den atomaren Aussagen aufgeteilt werden.

Nachteilig sei aber zu nennen, dass nur wenig über Schlussfolgerungsverfahren in Kombination mit Wissen, das durch Belief Functions dargestellt wird, untersucht wurde (Bhatnagar und Kanal, 1986). Allerdings ermöglicht Dempster's Rule of Combination die Möglichkeit der Kombination mehrerer Belief Functions. Dies ist bei der vorliegenden Anwendung sehr nützlich: Im Laufe der Analyse eines Bildes ergeben sich neue Informationen bezüglich des möglichen Raumkonzepts der Szene. Dieser Belief lässt sich entsprechend mit der Kombinationsregel berechnen.

Problematisch hingegen ist die Rechenkomplexität von Dempster's Rule of Combination, die exponentiell ist. Doch es existieren Vereinfachungen der Berechnung, durch die sich die Rechenkomplexität verringern läßt (Barnett, 1981).

Auch soll ein System erweitert werden, das im Rahmen der Objekterkennung bereits mit dieser Theorie arbeitet (siehe Kapitel 3.3). All diese Gründe sprechen dafür, sich bei der vorliegenden Arbeit für die Anwendung der Dempster-Shafer-Theorie zu entscheiden.

In den Kapiteln 4.3, 5.5.3 und 5.6.3 wird die praktische Umsetzung dargestellt.

## 2.5 Kombination von Ontologien und unsicherem Wissen

Es wurde bereits festgehalten, dass in Ontologien in erster Linie präzises Wissen modelliert wird. Auf Ungenauigkeit oder Unsicherheit wird weitestgehend verzichtet. Abhängig von der Domäne, die die Ontologie modelliert, kann diese Methode durchaus problemlos funktionieren.

In der vorliegenden Arbeit geht es allerdings um die Klassifikation von Räumen anhand ihrer enthaltenen Objekte. Welche Objekte in Räumen auftreten, kann nicht im Vorfeld präzise herausgearbeitet werden. Theoretisch sind hier beliebig viele Objekte vorstellbar und möglich. Daher soll sich auf diejenigen Objekte beschränkt werden, die zu der Klassifikation beitragen oder hauptsächlich in diesen Räumen auftreten



können. Die Objekte wiederum sollen anhand der für sie typischen Sakkadenzüge klassifiziert werden. Gerade hier kann aber der Fall eintreten, dass ein Objekt fokussiert wird, welches aber nicht klassifiziert werden kann, da es in der Wissensbasis nicht vorhanden ist. Genauso besteht aber auch die Möglichkeit, dass anhand der erkannten Objekte gar keine eindeutige Raumklassifikation vorgenommen werden kann. Gerade weil die verschiedenen Raumkonzepte häufig nicht logisch voneinander getrennt werden können, kann es sinnvoll sein, eine Klassifikation anhand des größten Glaubensmaßes vorzunehmen anstatt keine Klassifikation durchführen zu können. Eine Kombination von Ontologien und unsicherem Wissen kann eine Lösung für dieses Problem sein.

In der Literatur jedoch ist die Idee, Ontologien mit unsicherem Wissen zu verknüpfen, bisher nur wenig umgesetzt worden. Hier herrscht überwiegend eine strikte Trennung zwischen *ontologischem* und *epistemologischem* Wissen. Dabei beschreibt das epistemologische Wissen jenes Wissen, welches sich aus Meinungen oder Glauben ergibt. Die zentralen Fragen der Epistemologie sind „Was sollen wir glauben“ und „Was können wir wissen?“. In diesem Zusammenhang werden auch die Quellen untersucht, die zu diesem Glauben führen (Brendel und Koppelberg, 1996). Zwei realisierte Umsetzungen, in denen beide Wissensformen kombiniert werden, werden im folgenden Abschnitt näher behandelt. Die Kombination von Ontologien und unsicherem Wissen ist eine zentrale Idee dieser Arbeit. Es soll untersucht werden, in welchem Rahmen eine Kombination möglich ist. In dieser Arbeit soll, wie bereits oben erklärt, die Dempster-Shafer-Theorie für die Arbeit mit unsicherem Wissen dienen. Es soll herausgearbeitet werden, ob sich diese Theorie in eine Ontologie integrieren lässt, welche möglichen Probleme dabei auftreten und welche Vor- und Nachteile sich daraus ergeben. Detailliert werden diese Fragen in Kapitel 4 untersucht.

### 2.5.1 Kombination von OWL und der Bayes-Theorie

Bisher ist bereits der Versuch unternommen worden, die Themengebiete Ontologien und unsicheres Wissen durch eine Kombination von OWL (siehe Kapitel 4.1.3) und der Bayes-Theorie (siehe Kapitel 2.4.1) zu vereinen. Im Folgenden soll ein kurzer Überblick dieser Umsetzung (Ding u. a., 2004; Ding und Peng, 2004) gegeben werden:

Die (bedingten) Wahrscheinlichkeiten der Bayes-Theorie werden durch die OWL-Klassen „PriorProbObj“, „CondProbObjT“ und „CondProbObjF“ realisiert. Die Klasse „PriorProbObj“ verfügt über zwei Eigenschaften: „hasVariable“ und „hasProbValue“. Dabei legt die zweite Eigenschaft einen Wert zwischen 0 und 1 fest, der die Wahrscheinlichkeit für die Klasse der ersten Variablen definiert. Eine Instanz dieser Klasse setzt die Wahrscheinlichkeit  $\mathbf{P}(X)$  für eine Variable  $A$  fest. Die Klasse „CondProbObjT“ beschreibt die bedingte Wahrscheinlichkeit der Bayes-Theorie  $\mathbf{P}(X|Y)$ , „CondProbObj-

jF“ beschreibt  $P(X|\bar{Y})$ . Beide Klassen besitzen folgende Eigenschaften: „hasVariable“, „hasProbValue“ und „hasCondition“. Die ersten beiden Eigenschaften entsprechen denen der Klasse „PriorProbObj“. Die Eigenschaft „hasCondition“ gibt die Bedingung (Y) der aktuellen Variablen an.

In diesem Ansatz wird des Weiteren eine Möglichkeit beschrieben, wie diese Wissensstruktur in ein Bayes'sches Netz, also einen gerichteten, azyklischen Graphen, transformiert werden kann. Auch können bedingte Wahrscheinlichkeitsverteilungen für jeden Knoten (jede Klasse) berechnet werden. Ebenso können Reasoning-Strategien realisiert werden, doch befindet sich dieser Ansatz noch in der Entwicklungsphase. Insbesondere soll hier noch untersucht werden, ob Inferenzalgorithmen für Bayes'sche Netze übernommen werden können oder neue entwickelt werden müssen.

In dieser Umsetzung der Verknüpfung von Bayes-Theorie und Ontologien wird bisher allerdings nicht die Möglichkeit behandelt, dass unvollständiges Wissen vorhanden ist oder einzelne Parameter erst noch bestimmt werden müssen. Hier ist die Voraussetzung, dass alle Parameter, auch ob eine Variable wahr oder falsch ist, bereits bekannt sind. Dadurch wird die gesamte Wahrscheinlichkeitsverteilung errechnet und in die Ontologie integriert. In der vorliegenden Arbeit ist gerade diese Voraussetzung aber nicht gegeben. Die einzelnen Parameter werden erst nach und nach ermittelt und nach Möglichkeit soll schon nach wenigen Schritten, ohne alle Parameter zu überprüfen, ein Ergebnis vorliegen. Auch bietet dieses Verfahren keine hierarchischen Strukturen, die für die vorliegende Arbeit von Interesse sind. Zudem ist in dem vorhandenen System (Schill u. a., 2001) das unsichere Wissen mit der Dempster-Shafer-Theorie umgesetzt worden. Um konsistent in der Darstellung des unsicheren Wissens innerhalb des Systems zu bleiben, liegt es nahe, auch in dieser Arbeit diese Theorie zu verwenden.

### 2.5.2 Kombination von SWRL und Fuzzy Logic

Seit kurzem wird an einer Kombination von Fuzzy Logic und der *Semantic Web Rule Language* (SWRL)<sup>2</sup>, einer Erweiterung von OWL DL durch eine Teilsprache der Rule Markup Language (RuleML)<sup>3</sup>, gearbeitet (Pan u. a., 2005). Durch die Erweiterung mit RuleML sind in OWL DL unäre und binäre Datalog-Prädikate darstellbar. Diese Prädikate ähneln in ihrer Form Hornklauseln. Diese Darstellung soll es beispielsweise ermöglichen, Regeln aus bestehenden Systemen in die OWL DL-Syntax zu übernehmen.

Die Kombination von SWRL durch Fuzzy Logic besteht nun darin, dass man Wahrheitswerte zwischen 0 und 1 modellieren kann, die bezüglich OWL DL angeben, mit

---

<sup>2</sup><http://www.w3.org/Submission/2004/SUBM-SWRL-20040521/>

<sup>3</sup><http://www.ruleml.org/>

welchem Grad der Wahrscheinlichkeit eine Instanz zu einem Konzept gehört (siehe dazu Kapitel 4.1.4). Prinzipiell kann man damit ausdrücken, in welchem Maße eine Klasse, die durch Eigenschaften bestimmt ist, auf eine Instanz zutreffen. Beispielsweise kann man in SWRL folgende Aussage (nach Pan u. a. (2005)) treffen:

$\text{parent}(?x, ?p) \wedge \text{Happy}(?p) \rightarrow \text{Happy}(?x) * 0.8$

Diese Aussage bedeutet, dass, wenn jemand einen Elternteil hat, der glücklich ist, dann selbst zu einem gewissen Grad (von 0.8) glücklich ist. Zu diesem Grad gehört die Instanz „x“ der Klasse „Happy“ an.

Wie im vorherigen Abschnitt wird zwar auch hier nicht mit der Dempster-Shafer-Theorie gearbeitet, allerdings ließe sich die Modellierung der Wissensgrade auf die der Glaubensmaße übertragen. Dennoch ist auch diese Lösung nicht geeignet für die vorliegende Arbeit, denn bei dieser Kombination lassen sich nur Klassenzugehörigkeiten mit einer gewissen Unsicherheit beschreiben. In der vorliegenden Arbeit sind es aber die Beziehungen, die zu einem gewissen Maß an Unsicherheit modelliert werden sollen. Diese in Kapitel 4.1.2 vorgestellten Beziehungen lassen sich insbesondere nicht durch Klassenzugehörigkeiten modellieren, da eine Instanz nicht zu beiden der in Beziehung stehenden Konzepte gehören kann.

## 2.6 Inferenzmechanismen und Problemlösen

In Abschnitt 2.3 wurde bereits erwähnt, dass die Struktur der Wissensrepräsentation Inferenzmechanismen und damit das Problemlösen begünstigen sollte. An dieser Stelle soll ein kurzer Überblick über Problemlösungsstrategien folgen und der Reasoning-Prozess von Ontologien erläutert werden. Probabilistic Reasoning wurde im vorigen Abschnitt bereits behandelt.

### 2.6.1 Suchstrategien

Häufig werden *Suchstrategien* auf eine Wissensbasis angewendet, die graphisch in ihrer Struktur einem Baum entspricht. Jeder Knoten im Baum repräsentiert einen möglichen Zustand, der Wurzelknoten ist der initiale Punkt der Suche (Schill, 1990). Von hier aus gibt es verschiedene Möglichkeiten, den gesuchten Zielknoten, die Problemlösung, sofern einer vorhanden, zu finden.

**Breiten- und Tiefensuche** Beide Verfahren nutzen kein zusätzliches Vorwissen, die Suche orientiert sich strikt an der Baumstruktur. Dabei werden bei der Breitensuche stets die Knoten einer Ebene betrachtet, bevor die nächste Ebene untersucht werden

kann. Bei der Tiefensuche hingegen werden nacheinander die Pfade im Baum bis zu den Blattknoten verfolgt. Dabei kann man iterativ vorgehen, indem bei jeder Iteration die Tiefe, also die Anzahl der Ebenen, erhöht wird und auf der tiefsten Ebene mit dem Untersuchen des Pfades abgebrochen wird. Beide Methoden haben einen hohen Rechenaufwand (Schill, 1990).

**Hill-Climbing und A\*** Diese Verfahren nutzen zusätzliches Vorwissen, um die Suche zielgerichteter ausführen zu können. Beim Hill-Climbing-Verfahren wird kein Suchbaum genutzt. Stattdessen wird versucht, das globale Maximum einer Funktion zu finden, indem an der aktuellen, lokalen Stelle als nächstes der Schritt ausgewählt wird, der einen höheren Funktionswert liefert als der aktuelle. Bei dieser schrittweisen Steigerung („Hill-Climbing“) besteht allerdings das Problem, auf einem Plateau keinen Nachfolger zu finden oder bei einem lokalen Maximum zu enden (Russell und Norvig, 2003).

A\* nutzt eine heuristische Funktion, um den besten Pfad im Baum zu ermitteln. Diese setzt sich aus der Summe zweier Werte zusammen. Der eine Wert gibt die Kosten zum nächsten Knoten an, der andere Wert gibt, bestimmt durch die heuristische Funktion, die Kosten des nächsten Knotens zum Ziel an. Dabei darf die Heuristik nicht Kosten überschätzen, um die optimale Lösung zu finden. Insofern hängt die Suche durch A\* von der Güte der Heuristik ab (Schill, 1990).

**Forward/Backward Chaining und Backtracking** Je nach verwendeter Suchstrategie können zusätzliche Verfahren verwendet werden, um die Suche zu optimieren. So ist der Ausgangspunkt der Suche beim Forward Chaining der aktuelle Zustand, von dem aus mögliche Lösungsstrategien zum Ziel entwickelt werden. Beim Backward Chaining wird vom Zielknoten aus ein möglicher Lösungsweg zum aktuellen Zustand gesucht. Je nach Anwendungsfall kann sich für eines dieser Verfahren entschieden werden (Schill, 1990).

Gerät man bei der Suche in einen Zustand, von dem aus das Ziel nicht mehr zu erreichen ist, so kann man mittels Backtracking vor den Zustand zurückspringen, von dem der Pfad nicht mehr zur Lösung führen kann (Russell und Norvig, 2003).

Insgesamt gibt es noch eine Reihe weiterer Suchstrategien und zusätzliche Verfahrensweisen bei der Planung der Suche (beispielsweise bei Russell und Norvig (2003)), die aber im Rahmen dieser Arbeit nicht weiter relevant sind.

## 2.6.2 Reasoning

Beim *Reasoning* werden aus Prämissen anhand bestimmter Kriterien Schlussfolgerungen gezogen. Das Wissen liegt hierbei üblicherweise in einer formalen Notation vor, beispielsweise in einer aussagenlogischen Form. Oben wurde bereits Probabilistic Reasoning, also Schlussfolgern bei Angabe von Wahrscheinlichkeiten, vorgestellt (siehe Kapitel 2.4).

Bei Ontologien können *Reasoner* zusätzliches Wissen durch Inferenz erschließen. Dabei wird meist deduktives oder induktives Reasoning verwendet. Es wird also entweder vom Allgemeinen auf das Spezielle geschlossen oder umgekehrt. Abhängig von der Modellierungssprache und ihrer Ausdrucksmächtigkeit existieren entsprechende Reasoning-Mechanismen (Gómez-Pérez u. a., 2004). Dabei besteht die Hauptaufgabe der Reasoner bei Sprachen, die auf Beschreibungslogiken basieren, darin, die Erfüllbarkeit und Subsumption von Konzepten und somit die Konsistenz der Taxonomie und zusätzliche taxonomische Beziehungen zu ermitteln.

Bei der vorliegenden Arbeit wird für die praktische Umsetzung der Reasoner *Pellet* (siehe Kapitel 5.6.1) verwendet, dessen Inferenzalgorithmus auf dem Tableau-Kalkül basiert und der somit deduktiv terminologische Inferenzen generiert.

## 2.7 Lernen

Da das Gebiet *Lernen* nicht von sehr hoher Relevanz in der vorliegenden Arbeit ist, soll es nur kurz dargestellt werden. Es findet dadurch Verwendung, dass es im Rahmen des vorhandenen Systems genutzt wird. Es soll in diese Richtung aber nicht erweitert werden. Die Verwendung im System bezieht sich auf das Lernen der Objekte anhand von Sakkaden (Schill u. a., 2001).

### 2.7.1 Begriff des Lernens

Generell bezeichnet man in der Psychologie Lernen „als eine *Veränderung* von *Verhaltensdispositionen* durch *Erfahrung*. Man spricht von *Verhalten* und nicht von Wissen, weil nur Verhalten beobachtet werden kann (...). Man spricht von *Verhaltensdisposition*, also der Fähigkeit oder Möglichkeit, ein bestimmtes Verhalten zu zeigen, weil ein Lernvorgang auch stattfinden kann, ohne dass das entsprechende Verhalten direkt gezeigt wird.“ (Ferber, 2003)

Entsprechend besitzen Menschen die Fähigkeit, ihr Verhalten zu reflektieren und auch zu ändern. In Situationen wie der Weiterverarbeitung von Wahrnehmung oder dem Lösen von Problemen führt dies dazu, dass in zukünftigen, ähnlichen Situationen besser

oder optimal reagiert werden kann. In einem Teilgebiet der künstlichen Intelligenz, dem *maschinellen Lernen*, wird versucht, diesen Vorgang in einem Modell abzubilden oder ihn für den Lösungsvorgang zu nutzen. So können Lernvorgänge in Expertensystemen umgesetzt werden, anstatt diese lediglich durch einen immer größer werdenden Satz an Regeln zu beeinflussen. Letzteres wäre zu aufwendig und könnte zu Inkonsistenzen führen. Expertensysteme sollen vielmehr in die Lage versetzt werden, selbst zu lernen bzw. menschliche Lernprozesse abzubilden und ihr Wissen zu modifizieren (Reiss u. a., 1992).

### 2.7.2 Lernverfahren

Im Bereich der künstlichen Intelligenz gibt es verschiedene Methoden, Lernverfahren als Prozess zu modellieren. Eine Möglichkeit, Lernprozesse zu definieren, liegt in ihrer verwendeten Strategie, der Form der Wissensrepräsentation, ihrem Anwendungsbereich und dem Zweck des Lernprozesses. Der Zweck kann darin bestehen, neues Wissen zu erzeugen (*synthetisches Verfahren*), oder vorhandenes Wissen neu zu strukturieren (*analytisches Verfahren*), um gewünschte Ziele umsetzen zu können (Kodratoff und Michalski, 1990). Zudem unterscheidet man *überwachtes*, *unüberwachtes* und *Reinforcement-Lernverfahren*. Bei überwachtem Lernen werden den Eingabewerten ein bestimmter Ausgabewert zugewiesen, an dem sich das System orientieren kann. Dies ist bei unüberwachtem Lernen nicht der Fall. Hier werden beispielsweise die Eingabedaten kategorisiert, um charakteristische Muster zu extrahieren und unterscheiden zu können. Das Reinforcement-Lernen unterstützt den Prozess durch „Belohnung und Bestrafung“ und gibt so eine Güte des Ergebnisses an (Russell und Norvig, 2003). Eine strikte und eindeutige Differenzierung von Lernverfahren gibt es aber nicht.

**Induktive Lernverfahren** Hierbei handelt es sich um ein synthetisches Lernverfahren, in dem aus speziellen Informationen auf allgemeinere Sachverhalte geschlossen wird. Am Anfang steht üblicherweise ein Satz an Trainingsdaten zur Verfügung, der Eingaben eine bestimmte Ausgabe bzw. Kategorie zuordnet. Es wird folglich eine Funktion anhand der Ein- und Ausgabedaten gelernt. Sind diese Trainingsdaten gelernt worden, können auf neue Eingangsdaten entsprechende Ausgabedaten generiert werden.

Ein bekanntes Beispiel hierfür sind *Entscheidungsbaumverfahren*. Hier besteht die Eingabe aus einem Satz an Attributen, denen Werte oder Wertebereiche zugewiesen werden können. Der Entscheidungsbaum stellt die Entscheidungsregeln, nach denen eine Eingabe kategorisiert wird, graphisch dar und besteht aus aufeinanderfolgenden hierarchischen Entscheidungen (Russell und Norvig, 2003).

Dabei spricht man auch von *induktivem Lernen aus Beispielen*. Eine Alternative hierzu ist das *induktive Lernen von Regeln*. Hier wird versucht, aus den gegebenen Beispielen einen Satz an Regeln abzuleiten. Eine entsprechende Repräsentation dieser Regeln kann beispielsweise in prädikatenlogischer Form geschehen, so dass durch entsprechende logische Operationen zusätzliches Hintergrundwissen genutzt werden kann (Emde, 1996).

**Erklärungsbasierte Lernverfahren** Zu analytischen Methoden zählen verschiedene *erklärungsbasierte Lernverfahren*. Solche Verfahren deduzieren aus einem Beispiel einer Problemlösung eine Regel, mit der ähnliche Probleme gelöst werden sollen. Hierfür wird insbesondere untersucht, warum die Problemlösung korrekt ist. Die problemrelevanten Elemente werden extrahiert und mittels Generalisierung wird hieraus eine Regel konstruiert. Die Regel gibt an, welche Bedingungen erfüllt sein müssen, damit die Problemlösung korrekt funktionieren kann (Emde, 1996).

**Statistische Lernverfahren** Grundlage bei diesem Ansatz ist die statistische Datenmodellierung. Auf der Basis solcher Datenmodelle können Verfahren aus dem Bereich von Entscheidungstheorien oder Wahrscheinlichkeitsverfahren verwendet werden. Hier werden Hypothesen aufgrund der Eingangsdaten ermittelt, die in der Regel mit einer bestimmten Wahrscheinlichkeit versehen sind. Dabei handelt es sich wiederum um ein synthetisches Verfahren (Russell und Norvig, 2003).

*Neuronale Netze* gehören zur Gruppe der statistischen Lernmethoden und dienen generell der Erforschung biologischer Informationsverarbeitung (Mallot u. a., 1992). Sie modellieren Aspekte der biologischen Neuronen. Ein Neuron setzt sich dabei aus einer Anzahl von gewichteten Eingabewerten zusammen, die entweder direkt aus der Eingabeschicht oder von anderen Neuronen stammen. Eine Neuronfunktion erstellt ein Ergebnis aus diesen Eingabewerten und abhängig von dem Resultat „feuert“ das Neuron ab einem bestimmten Schwellenwert. Wenngleich neuronale Netze häufig für den Benutzer schwer nachzuvollziehen sind, so können sie gerade mit verrauschten Daten sehr gut umgehen.

Auch in dem zugrunde liegenden System (Schill u. a., 2001) kommt ein neuronales Netz bei der Bildklassifikation zum Einsatz. Abhängig von den bewerteten Testdaten lernt das System diejenigen Sakkadenzüge, die den höchsten Informationsgewinn versprechen. Näheres hierzu ist in Kapitel 3.3 zu finden.





## 3 Raumwahrnehmung mittels Blickpunktfolgen

„Jede perzeptive und imaginative Raumkognition ist blickpunktbezogen. Soweit man Gegenstände oder Konstellationen von Gegenständen im Raum wahrnimmt oder sie sich vorstellt, geschieht dies notwendigerweise zu jedem Zeitpunkt aus einem bestimmten Blickpunkt.“ (Herrmann, 1996). Auf diese Blickpunktbezogenheit beim Betrachten komplexer Szenen stützt sich die vorliegende Arbeit. Daher soll in diesem Abschnitt das Gebiet der Augenbewegungen, sich dadurch ergebende Wahrnehmung und Raumwahrnehmung und welche experimentellen Ergebnisse sich in diesem Zusammenhang in der Forschung ergeben haben, dargestellt werden.

### 3.1 Sakkadische Augenbewegungen

Das menschliche Auge führt unentwegt sakkadische Augenbewegungen aus. Dabei können diese Bewegungen Geschwindigkeiten von bis zu 600 Grad pro Sekunde erreichen (Mallot, 2000). Dies ermöglicht es, den Bereich des schärfsten Sehens, die Fovea centralis, stets an den aktuellen Ort des Interesses zu lenken. So wechseln sich Augenbewegungen und nachfolgende Fixationsmomente ab, in denen der visuelle Reiz analysiert wird. Die Sakkaden werden dabei überwiegend unwillkürlich ausgeführt, sie können aber auch gezielt beeinflusst werden. Dies hängt zu einem großen Teil von kognitiven Prozessen und zu einem geringeren Teil von sensomotorischen Faktoren ab (Deubel, 1994).

So unterscheidet man reflexive und freiwillige (*voluntary*) Sakkaden (Richards, 2003). Reflexive Sakkaden entstehen ungewollt aufgrund von optischen Reizen, während freiwillige Sakkaden absichtlicher Kontrolle unterliegen und durch Aufmerksamkeit bestimmt sind.

Informationen darüber, was für Eigenschaften Augenbewegungen beim Erkennen von Objekten aufweisen, sind vor allem durch Experimente mit Versuchspersonen erfasst worden. Bei diesen Experimenten, bei denen die Augenbewegungen aufgezeichnet werden, lassen sich viele Rückschlüsse auf das visuelle Wahrnehmen schließen. Diese Ergebnisse werden im Folgenden vorgestellt, zuerst bezüglich des Betrachtens und Erkennens einzelner Objekte, danach bezüglich komplexer Szenen.

Dabei ist allerdings der Begriff „Objekt“ nicht klar definiert. Bei der Bildbetrachtung wird meist ein spezieller Bildteil betrachtet, ohne dass gleich ein ganzes Objekt wahrgenommen werden könnte. Durch die Betrachtung weiterer Teile können dann Rückschlüsse auf das Objekt gezogen werden (Kosslyn, 1994). Doch ob es sich dabei nun um ein Objekt oder nur um einen Teil eines Objekts handelt ist häufig nicht eindeutig festzulegen: Ein Gesicht kann je nach Kontext als Teil eines Körpers, als ein einzelnes Objekt oder als die Summe seiner Linien und Punkte sowie deren Lage und Abstände betrachtet werden. Aus Experimenten ist zudem zu erkennen, dass Menschen beim Abzeichnen üblicherweise die Objekte in Umriss, Grundstruktur und Details zerlegen, um so das Objekt zu rekonstruieren (Goldenberg, 1998). Es lassen sich jedoch Hierarchien von Objekten erkennen. Die betrachteten Objekte werden üblicherweise in die Kategorie sortiert, in der die Objekteigenschaften am speziellsten repräsentiert sind. Wird beispielsweise ein Straußenvogel betrachtet, so wird dieser auch als solcher klassifiziert, während ein Sperling als Vogel klassifiziert wird (Kosslyn, 1994).

In dieser Arbeit wird folgende Unterscheidung verwendet: Ein Raumkonzept ist eine komplexe Szene, die sich aus einzelnen Objekten zusammensetzt. Die einzelnen Objekte sind Gegenstände, die sich im Raum befinden können. Diese Gegenstände sind in einer Hierarchie angeordnet.

## 3.2 Blickpunktfolgen einzelner Objekte

In Experimenten hat sich gezeigt, dass wiederkehrende, gleiche Sakkadenzüge bei der Erkennung von gleichen Objekten zu beobachten sind. Unterschiedliche Objekte werden auch mit unterschiedlichen Sakkadenzügen betrachtet. Bei verschiedenen Probanden jedoch sind bei gleichen Objekten auch gleiche Sakkadenzüge zu beobachten. Insbesondere die ersten Sakkaden des Sakkadenzugs haben hohe Ähnlichkeit (Noton und Stark, 1971a).

Die Fixationspunkte der Sakkadenzüge ergeben sich dabei aus den signifikanten Merkmalen, den *Features*, im Bild. Die gewählten Features besitzen dabei üblicherweise einen hohen Informationsgewinn, der zum Erkennen des Objekts beiträgt. Ihre Auswahl erfolgt anhand von Kantenenden oder -ecken und unvorhersehbaren oder komplizierten Bildpunkten (Noton und Stark, 1971b).

Die Reihenfolge der fixierten Punkte hängt dabei stark von der Erwartungshaltung des Betrachters ab. „We do not just see, we look.“ (Bajcsy, 1988). Die Betrachtung eines Punktes führt dazu, dass Hypothesen über das Betrachtete entwickelt werden und die darauf folgenden Blickpunkte durch diese Hypothesen beeinflusst werden (Kosslyn, 1994).

Hier lassen sich zwei Richtungen unterscheiden, wie sie beim Bildverstehen entsprechend modelliert werden (siehe Kapitel 2.2.2): Zum einen den Bottom-up-Prozess, bei dem Features im Bild aufgrund von Bildmerkmalen gewonnen werden, die Einfluss auf den Sakkadenzug haben. Zum anderen den Top-down-Prozess, bei dem Wissen aus Erfahrung den Sakkadenzug beeinflusst.

Daraus resultiert die Idee, einzelne Objekte anhand von Sakkadenzügen zu klassifizieren. Ein System, das diesen Ansatz erfolgreich umgesetzt hat, wird im Folgenden vorgestellt.

### 3.3 Szenenanalyse mit Okusys

Das *Okusys-System* ist ein hybrides System, das wissensbasierte Schlussfolgerungen auf Basis der Dempster-Shafer-Theorie mit einem neuronalen Netz für die Bildvorverarbeitung kombiniert, um das Dargestellte eines 2D-Eingabebilds anhand sakkadischer Augenbewegungen zu lernen und zu identifizieren (Schill u. a., 2001). Dabei kann es mit unvollständigem, mehrdeutigem oder inkonsistentem Wissen umgehen.

Die Umsetzung erfolgt durch eine Kombination von Bottom-up- und Top-down-Verfahren: Das Bottom-up-Verfahren ist für die vorverarbeitenden Schritte verantwortlich und extrahiert signifikante Merkmale im Bild. Das Top-down-Verfahren verfügt über eine Kontrollinstanz, in der die kognitiven Aspekte des Wahrnehmungsprozesses umgesetzt werden. Die einzelnen Schritte der Verfahren werden im Folgenden erklärt.

Die Features, die bei der Vorverarbeitung extrahiert werden, sollen sowohl für mögliche Fixationspunkte in Frage kommen als auch Informationen liefern, mit denen darauffolgende Fixationspunkte ermittelt werden können. Die Kriterien für die Auswahl der Features erfolgen anhand von linearen und nicht-linearen Filtern. Die nicht-linearen Filter extrahieren die Features, bei denen ein hoher Orientierungswechsel im Bild vorliegt, die durch Statistiken höherer Ordnung ermittelt werden können. Dieser Verarbeitungsschritt wird mit zusätzlichen Filteralgorithmen einmalig am Anfang einer Bildanalyse ausgeführt und gespeichert, so dass auf diesem vorverarbeiteten Bild während der Analyse gearbeitet werden kann. Wird während der Fixationsphase ein Feature fixiert, berechnen lineare Filter einen Featurevektor für den entsprechenden Fixationspunkt. Informationen aus diesem Vektor werden für den nächsten Schritt verarbeitet.

Die Augenbewegungen, die das System auf einem Objekt ausführt, sind allerdings nicht in strikte Klassen getrennt, das heißt die Klassifikation erfolgt nicht anhand festgelegter Sakkadenzüge. Vielmehr unterstützen Sakkadenzüge Hypothesen bestimmter Klassen. Diese Hypothesenhinweise werden durch Beliefs der Dempster-Shafer-Theorie

angegeben. Die Sakkadenzüge mit den zugehörigen Beliefs für Hypothesen können vom System gelernt werden.

Dabei verfügt das System über ein Lernverfahren, das sich an den Augenbewegungen orientiert, die den höchsten Informationsgewinn besitzen. Der Informationsgewinn ergibt sich aus den bereits untersuchten Features des Analysebilds. Diese Informationen können mit dem Wissen über Hypothesen aus vorher getätigten Analysen verglichen werden und der Belief für die aktuelle Hypothese kann daraus generiert werden. Dabei lernt das System die Hypothesen mit Hilfe von überwachtem Lernen: Dem System wird nach jeder Analyse ein klassifizierendes Label, das die Hypothese repräsentiert, als Feedback mitgeteilt. In der Wissensbasis werden die Glaubenswerte von Augenbewegungen bezüglich Hypothesen gespeichert.

Diese Glaubenswerte über die Hypothesen entsprechen dabei denen aus der Basic Probability Assignment der Dempster-Shafer-Theorie (siehe Kapitel 2.4.3). Diese Glaubenswerte können entsprechend Dempster's Rule of Combination miteinander verknüpft werden, wenn analysierte Augenbewegungen des Systems auf verschiedene Hypothesen hindeuten.

Zusätzlich nutzt das System eine Lernstrategie, bei der jeweils diejenige Augenbewegung gewählt wird, die zu einem Fixationspunkt führt, der im nächsten Schritt den höchsten Informationsgewinn verspricht. Die Lernstrategie *Inference by Information Gain*, IBIG (Schill, 1997), berechnet den Informationsgewinn für die Fixationspunkte durch die Differenz von den Beliefs der Hypothesen bezüglich der aktuellen Fixationsfolge, die aufgrund der bisherigen Analyse gewonnen werden konnten, und den möglichen Beliefs, die aufgrund der bisher nicht analysierten Eigenschaften noch erlangt werden können. Wie die in Kapitel 2.6.1 vorgestellten Suchstrategien, stellt auch IBIG die Hypothesen in einer Baumstruktur dar. Bei der Auswahl der nächsten Aktion, die zum nächsten Zustand führt, wird aber wie beschrieben kein heuristisches Verfahren sondern die Berechnung der Beliefs verwendet. Außerdem wird nicht nur ein Pfad zur Zeit betrachtet, sondern sämtliche Hypothesen, die durch die Basic Probability Assignment als Lösung in Frage kommen. So ist kein zeitaufwendiges Backtracking notwendig, um zu einer Hypothese eines anderen Pfades zu gelangen, wenn diese den nächsten, höchsten Informationszuwachs verspricht.

Durch die Verwendung einer baumartigen Datenstruktur kann zusätzlich die Rechenkomplexität der Kombinationsregel durch Approximationsverfahren verringert werden. So werden nur die unterstützenden Evidenzen berechnet, um die Rechenzeit von exponentiell auf linear zu reduzieren (Barnett, 1981). Hierdurch lassen sich zusätzlich summierte Beliefs einer Teilhierarchie im Baum approximieren (Gordon und Shortliffe, 1985).

Das System durchläuft während einer Analyse folgende Schritte: Durch die Vorverar-

beitung werden die Features im Bild extrahiert, die im folgenden als Fixationspunkte in Frage kommen. Initial wird ein zufälliger Fixationspunkt im Bild gewählt. Der Featurevektor für die aktuelle Position wird erstellt und durch die IBIG-Lernstrategie werden die neuen Beliefwerte für die Hypothesen mit Hilfe des vorher gelernten Wissens berechnet. Das Feature, das aufgrund dieser Berechnung den höchsten Informationsgewinn im nächsten Schritt verspricht, wird als nächster Fixationspunkt ausgewählt und eine entsprechende sakkadische Augenbewegung zu diesem Punkt ausgeführt. Wie hoch schließlich der tatsächliche Informationsgewinn an diesem Punkt ausfällt, wird in der Wissensbasis gespeichert, so dass diese Informationen in zukünftige Analysen einfließen können.

Durch die Kombination sensorischer und motorischer Aktionen, also zum einen dem Wahrnehmen der Informationen des aktuell fixierten Punktes und der aktiven Bewegung zum nächsten Fixationspunkt, kann man bei diesem System von einer Umsetzung von „Active Vision“ (siehe Kapitel 2.2.1) sprechen. Eine Erweiterung dieses Systems soll in dieser Arbeit vorgestellt werden (siehe unten), bei der die einzelnen Gegenstände im Raum mit der hier beschriebenen Methode gelernt werden (Beispielanwendung in Kapitel 5.7.1).

### 3.4 Blickpunktfolgen komplexer Szenen

Bei komplexen Szenen, die sich aus einer Vielzahl von Objekten zusammensetzen, verlaufen die Augenbewegungen häufig von Objekt zu Objekt (siehe Abbildung 3.1 und 3.2). Die einzelnen Bildobjekte werden erkannt und daraus Rückschlüsse auf die ganze Szene gezogen. Dabei ist es sehr auffällig, dass manche Objekte sehr häufig und über einen längeren Zeitraum fixiert werden, während andere Objekte kaum oder keine Beachtung erhalten. Die Aufmerksamkeit des Betrachters richtet sich dabei vor allem auf solche Objekte, die entweder einen hohen Informationsgehalt für die Szene haben, deren Anwesenheit in der dargestellten Szene eher unbekannt, irritierend oder unverständlich ist, oder die durch ihr bloßes Erscheinungsbild interessant erscheinen. Dabei beeinflussen weder der Detaillierungsgrad eines Objekts im Bild noch dessen Helligkeit oder Farbe die Augenbewegungen beim Betrachten. Objektumrisse hingegen haben einen stärkeren Einfluss (Yarbus, 1967).

Experimente haben auch gezeigt, dass Betrachter komplexer Szenen weitergehende Schlüsse über den Kontext der Szene ziehen, als nur das Dargestellte. So machen sich Betrachter beispielsweise Gedanken über mögliche Beziehungen zwischen im Bild auftretenden Personen oder welche Atmosphäre der Situation zugrunde liegen könnte. Die Schlussfolgerungen gehen also weit über das bloße Raumkonzept hinaus (Yarbus,

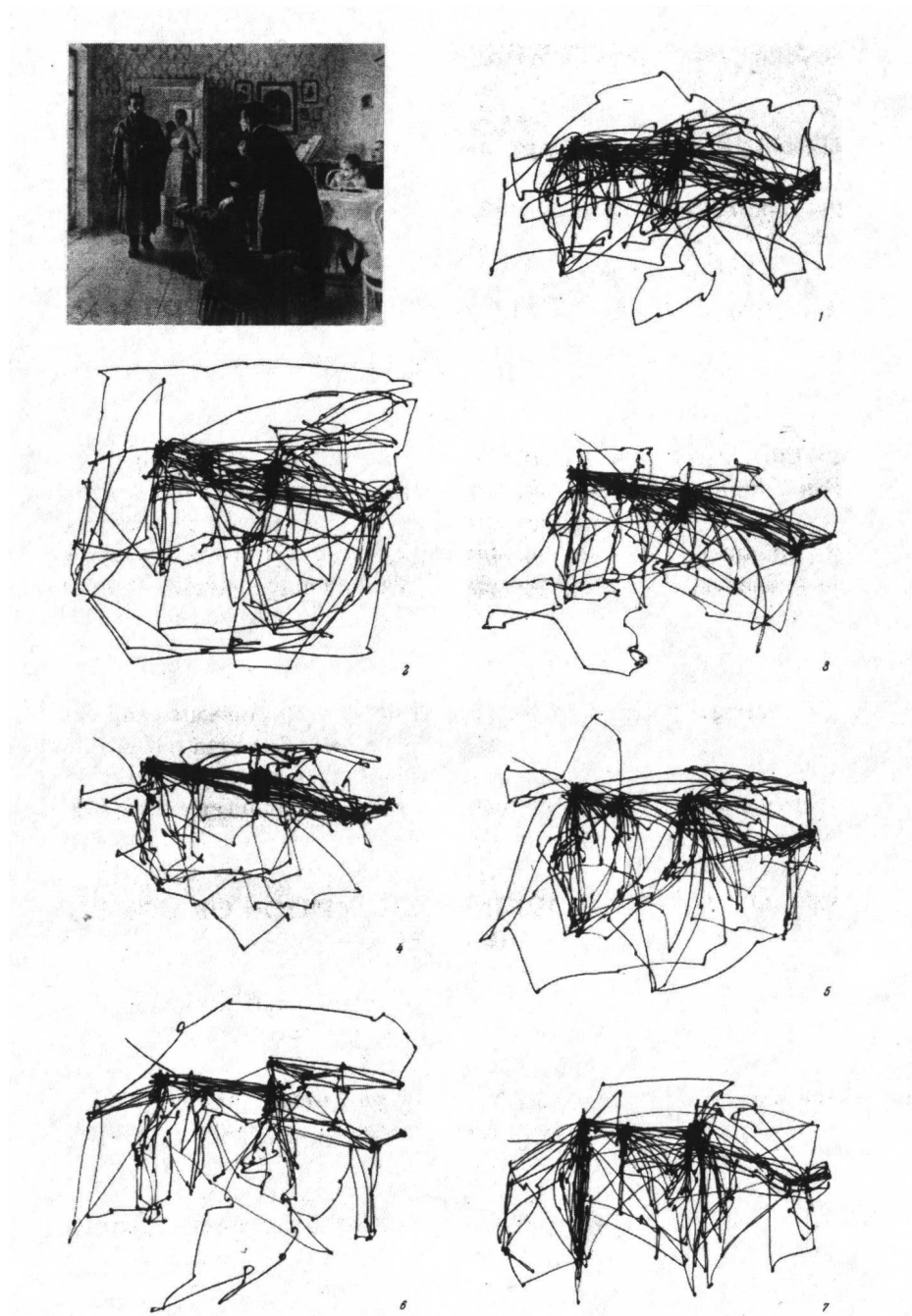


Fig. 107. Reproduction from I.E. Repin's picture "An Unexpected Visitor" and records of the eye movements of seven different subjects. Each subject examined the picture freely (without instruction) with both eyes for 3 minutes.

Abbildung 3.1: Sakkadenzüge unterschiedlicher Probanden, aus Yarbus (1967)

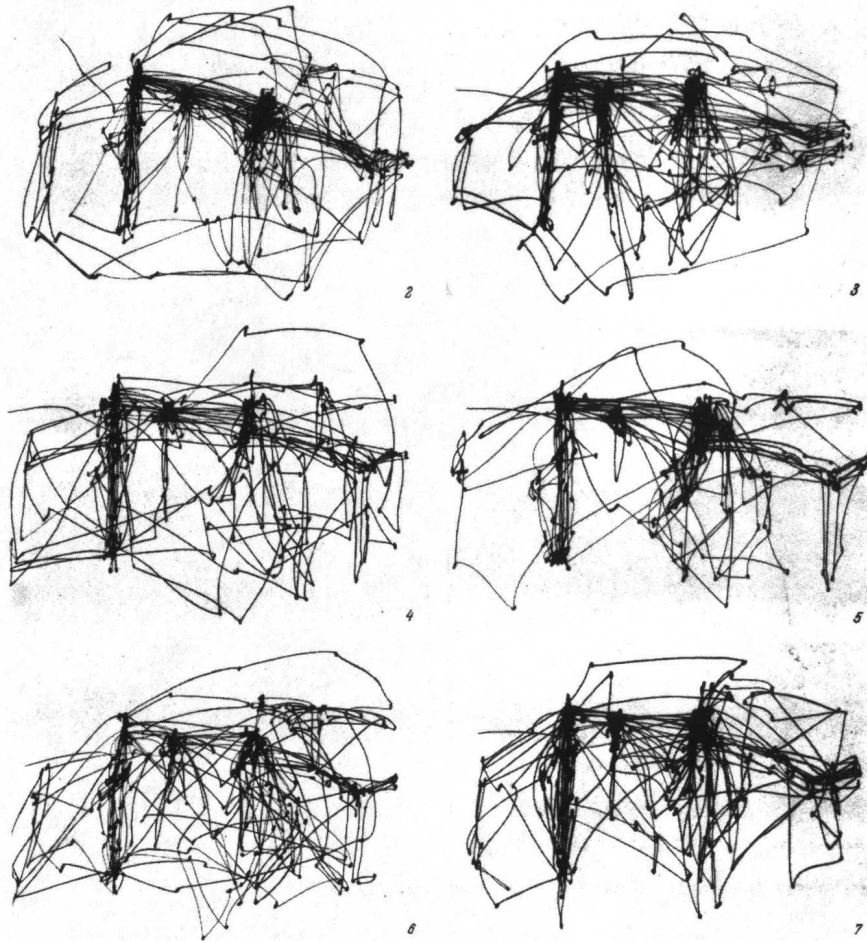


Fig. 108. Seven records of eye movements by the same subject, examining Repin's picture freely with both eyes. The records, arranged in chronological order, lasted 3 minutes. The interval between records was 1 or 2 days.

Abbildung 3.2: Sakkadenzüge desselben Probanden, aus Yarbus (1967)

1967). Gleichzeitig ist aber auch festzuhalten, dass Schlussfolgerungen davon abhängen, ob und welche Aufgabe einem Probanden beim Betrachten einer Szene gestellt wird.

Zusammenfassend lässt sich feststellen, dass einzelne Teile im Bild erkannt werden und mit zusätzlicher Hilfe von Hintergrundwissen können Erkenntnisse über die komplexe Szene gewonnen werden (Yarbus, 1967).

### 3.5 Erweiterung der Szenenanalyse

Die hier beschriebenen Ergebnisse sollen zeigen, dass beim Betrachten einer komplexen Szene einzelne Objekte erkannt werden und über die Kombination dieser erkannten Objekte auf die komplexe Szene und sogar darüberhinaus gehende Sachverhalte Rückschlüsse gezogen werden. Diese experimentellen Befunde macht sich die vorliegende Arbeit zunutze und der hier beschriebene Sachverhalt soll in einem System realisiert und geprüft werden.

Die einzelnen Objekte (Gegenstände) sollen mit dem oben vorgestellten System klassifiziert werden. Als Erweiterung soll durch die einzeln erkannten Objekte die komplexe Szene (das Raumkonzept) klassifiziert werden. Im folgenden Kapitel werden hierfür notwendige Überlegungen zur Wissensrepräsentation angestellt, in Kapitel 5 wird eine beispielhafte Umsetzung für Universitätsräume vorgestellt.



## 4 Ontologien und unsicheres Wissen

In den Kapiteln 2.3.2 und 2.4 finden sich grundlegende Informationen zu den Themen Ontologie und unsicheres Wissen. An dieser Stelle soll detaillierter auf beide Themenbereiche eingegangen werden, um anschließend ihre Kombinationsmöglichkeit vorzustellen.

### 4.1 Ontologie

In Kapitel 2.3.2 wurde darauf hingewiesen, dass es keine eindeutige Definition des Begriffs Ontologie gibt. Die verschiedenen Sichtweisen darüber, was unter einer Ontologie zu verstehen ist, begründen sich insbesondere durch die verschiedenen Fachdisziplinen, in denen Ontologien genutzt werden, und ihre dortigen Verwendungszwecke. Die Herkunft aus der Philosophie wurde eingangs bereits erwähnt.

#### 4.1.1 Differenzierte Ontologiedefinitionen

In der *Linguistik* werden mit Ontologien semantische Konstrukte beschrieben (Gómez-Pérez u. a., 2004). Sie sind in erster Linie an die Semantik grammatikalischer Einheiten gebunden. Ihr Anwendungsgebiet findet sich häufig im Zusammenhang mit *Sprachverarbeitungsprozessen* (engl.: natural language processing). Ein bekanntes Beispiel für eine linguistische Ontologie ist WordNet<sup>1</sup>, hauptsächlich eine lexikalische Datenbank der englischen Sprache, welche auf psycholinguistischen Ergebnissen beruht.

Gruppen unterschiedlicher Disziplinen vor allem aus der Philosophie beschäftigen sich wiederum mit sogenannten *Upper-Level-* oder *Top-Level-Ontologien* (Gómez-Pérez u. a., 2004). In solchen Ontologietypen werden ausschließlich generelle Konzepte modelliert. Das gesamte Weltwissen soll hier repräsentiert und in Konzepte strukturiert werden. Auf eine bestimmte Domäne sind sie nicht beschränkt. Das Ziel dieser Modellierung liegt darin, ein Framework zu erstellen. Existierende Ontologien sollen in die dort beschriebenen Konzepte integriert werden. Ontologien, die neu erstellt werden, sollen auf Basis dieser Konzepte modelliert werden. Bisherige Versuche von Top-Level-Ontologien

---

<sup>1</sup><http://wordnet.princeton.edu/>

greifen nur auf taxonomische Modellierungskonzepte zurück und variieren stark in ihren Klassifikationskriterien. Die IEEE Standard Upper Ontology (SUO) Working Group versucht, mit SUMO<sup>2</sup> eine Top-Level-Ontologie zu entwickeln, die ein allgemeiner Standard werden soll. Als weiteres Beispiel einer Top-Level-Ontologie sei Cyc<sup>3</sup> erwähnt. Diese Ontologie versucht sämtliches Weltwissen zu erfassen, umfasst derzeit etwa 47000 Konzepte und ist in einer Open-Source-Version<sup>4</sup> zugänglich.

Dennoch mag die Entwicklung verschiedener Top-Level-Ontologien kritisch betrachtet werden: Selbst eine sehr ausgereifte, allumfassende Ontologie kann nur dann von Nutzen sein, wenn auch ihre konkrete Verwendung möglich ist. Sie kann zwar als Vorlage einer konkreten Ontologieimplementierung dienen, doch bedeutet dies nicht zwangsläufig, dass damit eine Integration, Kommunikation oder ein Austausch konkreter Ontologien gewährleistet wird. Denn die Ontologien, die erstellt werden, können gegenseitig konkurrieren. Sollen Ontologien in die Strukturen integriert werden, muss zudem die Semantik der konkret verwendeten Begrifflichkeiten im Einzelfall überprüft und Konflikte müssen aufgelöst werden.

Zudem sind in bisherigen Top-Level-Ontologien ausschließlich taxonomische Beziehungen modelliert. Mögliche Anfragen, die an eine solche Ontologie gestellt werden können, können demnach nur bezüglich Unter-/Oberklassen-Beziehungen oder Teilmen-gen gestellt werden. Für die hier beschriebene Analyse von Bildern reicht das aber nicht aus. Hier werden Anfragen bezüglich partonomischer Beziehungen gestellt, beispielsweise um welches Raumkonzept es sich handelt, wenn bestimmte Einrichtungsgegenstände bekannt sind.

Im Gegensatz zu einer Top-Level-Ontologie wird in *Domain Ontologies* nur ein sehr spezieller Themenbereich modelliert. Hier werden nicht nur die partonomischen Beziehungen und Relationen zwischen Konzepten sowie deren Eigenschaften definiert, auch allgemeine Prinzipien und Theorien der Domäne werden modelliert (Gómez-Pérez u. a., 2004). Beispielsweise kann in einer Ontologie über Internet-Webseiten modelliert werden, wie die Syntax einer URL definiert ist. Domain Ontologies haben einen hohen Praxisbezug und werden gegenwärtig bereits in unterschiedlichen Themenfeldern benutzt, beispielsweise auf Gebieten der Medizin, des E-Commerce oder der Ingenieurwissenschaften. Als konkrete Umsetzung sei E-cl@ss<sup>5</sup> genannt, in der Materialklassifikation und Warengruppen definiert sind und die der Kommunikation zwischen Lieferanten und Kunden dient. Im Einsatz ist E-cl@ss unter anderem bei SAP, BASF und Siemens.

---

<sup>2</sup><http://www.ontologyportal.org/>

<sup>3</sup><http://www.cyc.com/>

<sup>4</sup><http://www.opencyc.org/>

<sup>5</sup><http://www.eclass.de/>

Problematisch bei Domain Ontologies ist allerdings, dass sie häufig zu speziell auf ihren Themenbereich zugeschnitten sind. Das erschwert insbesondere die Wiederverwendbarkeit. Auch die Kommunikation unter mehreren Domain Ontologies führt oft zu Problemen, da beispielsweise unterschiedliche Namen für denselben Begriff oder gleiche Begriffsbezeichnungen mit unterschiedlicher Bedeutung verwendet werden.

Im Fall der hier beschriebenen Ontologianwendung würde zwar eine domänenspezifische Ontologie erstellbar und nutzbar sein, doch sollen gerade die eben erwähnten Probleme von Domain Ontologies vermieden werden. Vielmehr soll sich die Struktur der Ontologie an einer allgemeinen, domänenunspezifischeren Ontologie orientieren. Zwar wird sie hier isoliert betrachtet und vorgestellt, doch ließe sie sich ohne Probleme in die übergeordnete Ontologie als Teil integrieren. Diese übergeordnete Ontologie wird in dem Teilprojekt „I1-[OntoSpace]“ des „SFB/TR8 Spatial Cognition“<sup>6</sup> entwickelt, mit dem eine Kooperation besteht. In diesem Projekt geht es um die Erstellung einer linguistischen, räumlichen Ontologie auf Basis von *GUM*<sup>7</sup> und *DOLCE*<sup>8</sup>.

Bei *GUM* (*Generalized Upper Model*) (Bateman u. a., 1995) handelt es sich um eine linguistische Ontologie. Hier werden im Gegensatz zu dem erwähnten WordNet die semantischen Ausdrucksmöglichkeiten von größeren grammatikalischen Einheiten definiert (Gómez-Pérez u. a., 2004). *DOLCE* (*Descriptive Ontology for Linguistic and Cognitive Engineering*) ist eine der Ontologiebibliotheken von *WonderWeb* (Masolo u. a., 2003), die als Anleitung für neu zu erstellende Ontologien gedacht sind und in die existierende Ontologien integriert werden können. Die Bibliotheken sollen als Standard beliebiger Themengebiete für das Semantic Web dienen. Ihr Ziel kommt somit einer Top-Level-Ontologie sehr nahe. Insofern ist es erforderlich, die unterschiedlichen Ontologiedefinitionen zu kombinieren und die zu erstellende Ontologie zur Raumanalyse in den Gesamtkontext einzugliedern.

### 4.1.2 Vereinigung der unterschiedlichen Ontologiesichtweisen

Im vorigen Abschnitt wurde aufgeführt, warum verschiedene Ontologiedefinitionen zu berücksichtigen und zu vereinen sind. Kombiniert werden die hier zu erstellende domänenspezifische Ontologie von Raumkonzepten mit der Ontologie aus „I1-[OntoSpace]“.

Die Kategorisierung der hier zu modellierenden Raumkonzepte hängt von der funktionalen Verwendung des Raumes ab, die sich durch enthaltene Einrichtungsgegenstände auszeichnet. Beispielsweise trifft dann das Konzept eines Büros auf einen Raum zu,

---

<sup>6</sup><http://www.sfbtr8.uni-bremen.de/i1.html>

<sup>7</sup><http://www.fb10.uni-bremen.de/anglistik/langpro/webpace/jb/gum/index.htm>

<sup>8</sup><http://www.loa-cnr.it/DOLCE.html>

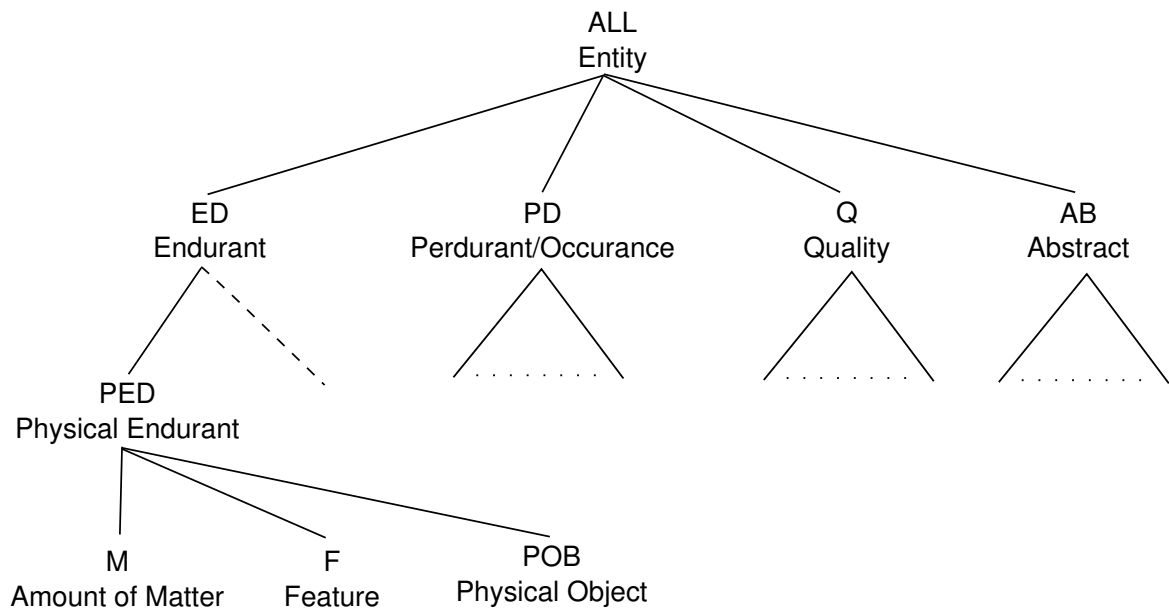


Abbildung 4.1: Auszug aus der Taxonomie der Basiskonzepte von DOLCE

wenn dieser bestimmtes Mobiliar oder elektronische Gegenstände beinhaltet. Eine solche Kategorisierung sieht DOLCE allerdings nicht vor: Hier werden Räume in physikalische Einheiten sortiert, die sich durch räumliche Maße definieren (siehe Abbildung 4.1).

Die hier zu definierenden Räume können zwar in einer Unterklassenhierarchie unter dem Konzept *physical endurant* (*PED*) modelliert werden, doch werden sie damit nicht hinlänglich beschrieben und können so auch nicht anhand ihres Mobiliars klassifiziert werden. Außerdem sind in DOLCE generell keine Eigenschaften von Konzepten beschrieben. Daher müssen Eigenschaften für die Raumkonzepte ergänzt werden, so dass ein Raum bestimmte Einrichtungsgegenstände enthalten kann (siehe konkrete Modellierung in Kapitel 5).

Mit dieser Erweiterung sind die Raumkonzepte aber weiterhin noch nach ihren physikalischen Gegenbenheiten strukturiert. Um dennoch eine funktionale Gliederung integrieren zu können, wird das Konzept sogenannter *Niches* verwendet (Smith und Varzi, 1999). Niches besitzen nach der Definition von Smith eine räumliche Ausdehnung und existieren nur für einen bestimmten Zeitraum. Sie definieren, welche Aufgaben bzw. welches Dasein eine (physikalische) Einheit ausmacht. Damit können unterschiedliche Sichtweisen auf Objekte, die sich nicht auf die rein physikalischen Maße beschränken, modelliert werden. Im hier beschriebenen Fall bieten Niches die Möglichkeit, Funktionen eines Raums zu modellieren.

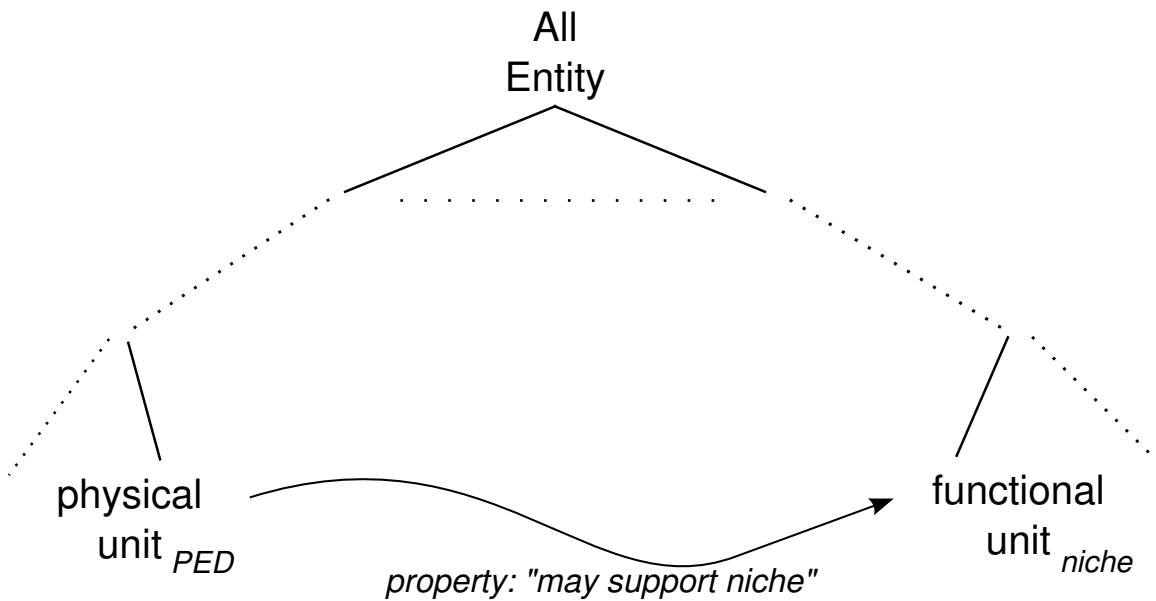


Abbildung 4.2: Niches als Eigenschaft der Konzepte

Bei der vorliegenden Umsetzung werden Raumkonzepte als Niches modelliert (Bateman, 2005): Sie stehen für eine Eigenschaft, die ein Raum als physikalische Einheit (physical unit) besitzen kann. Dieser Raum kann eine Niche *unterstützen* (may support), das heißt er kann für einen gewissen Zeitraum eine Funktion (functional unit), die durch Niches dargestellt wird, einnehmen. Die Niche selbst ist ebenfalls ein Konzept aus der Ontologie. Dieser Zusammenhang wird in Abbildung 4.2 verdeutlicht. In Kapitel 5.1 werden diese Zusammenhänge konkret in der Ontologie umgesetzt und anhand der Raumkonzepte, die dabei gerade die Funktion von Niches übernehmen, beschrieben.

Diese Umsetzung stützt sich zusätzlich auf Smith, dessen Ontologie nicht aus einem einzigen Taxonomiebaum besteht, der *wohlgeformt* ist und in dem Konzepte eindeutig modelliert sind. Stattdessen werden verschiedene Taxonomieebenen zu einem Baum vereint, die spezielle *Views (Facets or Factors)* der Domäne repräsentieren (Smith, 2005), das heißt es sind unterschiedliche Perspektiven der Ontologie modelliert.

### 4.1.3 OWL (Web Ontology Language)

Obwohl die Ontologiemodellierungssprache *OWL* erst seit eineinhalb Jahren existiert, ist sie auf dem Gebiet der formalen Ontologierepräsentation bereits weit verbreitet. Dies liegt unter anderem an der Entwicklung des *Semantic Web*, wovon die Sprache ein Teil

ist, durch das World Wide Web Consortium (W3C)<sup>9</sup>. OWL ist eine Erweiterung des *Resource Description Framework (RDF)*<sup>10</sup> mit zum Teil zusätzlichen Einschränkungen und einem größeren Vokabular. Entstanden ist sie als Nachfolger von *DAML* und *OIL*<sup>11</sup>. OWL ist dafür vorgesehen, Informationen durch Anwendungen bearbeiten zu lassen und sie nicht nur für Benutzer zur Betrachtung zur Verfügung zu stellen. Mit OWL lassen sich Begriffe, deren Eigenschaften und Beziehungen repräsentieren. Dabei hält sich die Sprache syntaktisch an das XML-Format.

OWL ist in drei Subsprachen gegliedert: *OWL Lite*, *OWL DL* und *OWL Full*. OWL Lite besitzt die geringste Komplexität und stellt neben der Klassifikationshierarchie nur einfache Bedingungen zur Verfügung. Der Name OWL DL ergibt sich aus der Verbindung zur *Beschreibungslogik* (engl. description logic). OWL DL bietet eine höhere Ausdrucksmächtigkeit als OWL Lite, bleibt aber dennoch berechen- und entscheidbar. Mit OWL DL können prinzipiell alle Sprachkonstrukte umgesetzt werden, doch bestehen bei manchen der Konstrukte Einschränkungen. OWL Full hingegen bietet die maximale Ausdrucksstärke und syntaktische Freiheiten von RDF allerdings ohne Berechenbarkeitsgarantien.

In der vorliegenden Arbeit soll die Berechenbarkeit für Anfragen garantiert sein, um Antworten auf Anfragen in einer angemessenen Zeit zu erhalten und erwarten zu können, so dass OWL Full nicht in Frage kommt. Die Ausdrucksstärke von OWL Lite genügt den Komplexitätsanforderungen wiederum nicht, beispielsweise existiert das Klassenaxiom *disjointWith* in OWL Lite nicht. Daher wird in der vorliegenden Arbeit OWL DL eingesetzt.

### 4.1.4 OWL DL

Es soll ein kurzer Überblick gegeben werden über den Aufbau einer OWL-Datei und welche Konstrukte die Sprache bietet. Gerade ersteres bezieht sich auf alle Untersprachen von OWL, doch soll sich dieser Überblick vornehmlich auf die Möglichkeiten der Sprache OWL DL beziehen.

Den Anfang einer OWL-Datei bilden Angaben über den XML-Namespace, die im weiteren Verlauf der Datei verwendeten Begriffe befinden sich demnach in dem angegebenen Namespace. So ist eine höhere Lesbarkeit der Begriffe und eine Auflösung möglicher Namenskonflikte bei der Kombination verschiedener OWL-Ontologien gewährleistet. Auch Abkürzungen können am Anfang einer OWL-Datei definiert werden. Hieran

---

<sup>9</sup><http://www.w3.org/Consortium/>

<sup>10</sup><http://www.w3.org/RDF/>

<sup>11</sup><http://www.w3.org/TR/2001/NOTE-daml+oil-reference-20011218>

schließen sich allgemeine Angaben der Ontologie an, beispielsweise die Versionsnummer, Kommentare, importierte Ontologien und ähnliches.

Die elementaren Sprachkomponenten beziehen sich vor allem auf die Beschreibung von Klassen, Instanzen, Eigenschaften von und Relationen zwischen Klassen. Jede Klasse ist dabei Unterklasse von *owl:Thing*. Auch die leere Klasse *owl:Nothing* wird bereitgestellt. Klassen werden mittels

```
<owl:Class rdf:ID="Klasse1" />
```

erstmalig benannt und können schon als Unterklasse einer anderen Klasse deklariert werden:

```
<owl:Class rdf:ID="Klasse11">  
  <rdfs:subClassOf rdf:ID="Klasse1" />  
</owl:Class>
```

Mit diesen Konstrukten kann die Taxonomie beschrieben werden. Individuen (Instanzen) werden ähnlich instanziiert: *<Klasse1 rdf:ID="Klasse1-Instanz">*.

Um Eigenschaften von Klassen zu beschreiben stellt OWL *Properties* bereit, die den Klassen zugewiesen werden können. Dabei unterscheidet man in OWL *ObjectProperty* und *DatatypeProperty*. *ObjectProperties* weisen einer Klasse eine weitere Klasse zu, so dass eine Relation zwischen diesen Klassen hergestellt wird. *DatatypeProperties* hingegen weisen einer Klasse einen Datentypen zu. Dieser Datentyp kann von der Art String, Integer, Boolean u. a. sein. Des weiteren werden bei der Definition von *Properties* ihre *Domain* und *Range* angegeben. Die *Domain* legt die Klassen fest, welche die *Properties* besitzen können, mit *Range* werden die Klassen bezeichnet, die mit den *Properties* angegeben werden können. So bezeichnet bei

```
<owl:ObjectProperty rdf:ID="hatTitel">  
  <rdfs:domain rdf:ID="Buch" />  
  <rdfs:range rdf:ID="Buchtitel" />  
</owl:ObjectProperty>
```

die Klasse *Buch* die *Domain* der Property *hatTitel*, *Buchtitel* bezeichnet die *Range*. Die konkreten Instanzen können dann entsprechend den *Properties* ihrer Klassen, den sie angehören, in Verbindung zu anderen Instanzen oder Werten von Datentypen stehen.

*Properties* können mittels *rdfs:subPropertyOf* wie Klassen hierarchisch strukturiert werden. Sie können transitiv, symmetrisch, funktional und invers definiert werden. Zudem besteht die Möglichkeit, weitere Einschränkungen für die *Properties* zu definieren. Solche *Restrictions* können festlegen, dass eine Eigenschaft auf alle Instanzen

einer Klasse zutreffen (`someValuesFrom`), dass Eigenschaften einer Klasse aus einem bestimmten Range stammen (`allValuesFrom`) und dass Eigenschaften Kardinalitäten oder einen bestimmten Wert besitzen müssen.

Des Weiteren können Äquivalenzen zwischen Klassen beschrieben werden, was vor allem bei der Kombination mehrerer Ontologien nützlich sein kann. Es können weiterhin komplexere Klassenstrukturen gebildet werden, beispielsweise Schnittmenge, Vereinigung und Komplement. Enumerierte und disjunkte Klassen können ebenfalls definiert werden (Antoniou und van Harmelen, 2004).

### 4.1.5 Ontologie-Editor

In der vorliegenden Arbeit wird der Ontologie-Editor *protégé*, Version 3.1.1, verwendet, mit dem die Ontologie modelliert wird. Der Editor wird an der *Stanford University School of Medicine* entwickelt und ist frei erhältlich<sup>12</sup>. Er unterstützt Modellierungsmöglichkeiten mit Frames, XML Schema, RDF(S) und OWL und bietet eine graphisch übersichtliche Gliederung der Ontologie. Auch ein Reasoner lässt sich einbinden.

## 4.2 Unsicheres Wissen

Es sei noch einmal festgehalten, dass unsicheres Wissen gerade deshalb in der Ontologie modelliert werden soll, da die Frage, um welches Raumkonzept es sich bei einer Bildanalyse handelt, nicht mit absoluter Gewissheit berechnet werden kann. Durch das Auftreten einer Reihe einzelner Objekte im Bild kann nur selten ein Konzept eindeutig zugewiesen werden. Zum einen können weitere Objekte extrahiert und analysiert werden, die gegen eine aktuelle Hypothese aufgrund der bis dato extrahierten Objekte spricht. Zum anderen können nicht genügend Informationen durch die vorkommenden Objekte gewonnen werden, so dass keine eindeutige Klassifikation möglich ist. Auch ist denkbar, dass der Bildausschnitt einen Raum zeigt, der unterschiedlichen Funktionen dient und der somit unterschiedliche Raumkonzepte repräsentieren kann. Es lassen sich nicht alle Raumkonzepte logisch voneinander abgrenzen, wie es in einer Ontologie eigentlich üblich wäre.

Insofern kann nur mit einem gewissen Glaubensmaß angegeben werden, wie wahrscheinlich die aktuell zu untersuchende Bildszene einem bestimmten Konzept entspricht. Dass dies keine Unzulänglichkeit ist, sondern an der Anfrage und Domäne liegt, zeigen die Begutachtungen von Raumausschnitten durch menschliche Betrachter (siehe

---

<sup>12</sup><http://protege.stanford.edu/>



Kapitel 5.5.3): Zum Teil gibt es eine hohe Unentschlossenheit der Gutachter, welches Raumkonzept auf den jeweiligen Raumausschnitt zutrifft.

Zwei Möglichkeiten, wie unsicheres Wissen in Ontologien einfließen kann, wurde bereits in Kapitel 2.5 dargestellt. Wie Glaubensmaße für Raumkonzepte in der Ontologie modelliert werden können, wird im Folgenden vorgestellt.

## 4.3 Kombination: Dempster-Shafer-Theorie und Ontologien

In Kapitel 2.4.3 wurde die Dempster-Shafer-Theorie bereits ausführlich erklärt. An dieser Stelle soll eine mögliche Umsetzung in Ontologien durch die Nutzung der Beliefwerte folgen.

Als erstes müssen die Hypothesen des Anwendungsfalls herausgearbeitet werden, denen überhaupt Glaubenswerte zugewiesen werden sollen. Im vorliegenden Fall sind es die Raumkonzepte, die zu einem gewissen Glaubensmaß unterstützt werden, soweit es Informationen über die Objekte im Raum gibt. Die Hinweise, die man über die einzelnen Objekte im Raum besitzt, beeinflussen das Glaubensmaß zu der Frage, um was für ein Raumkonzept es sich bei der aktuellen Szene handelt.

Bei der Beantwortung der Frage, welches Raumkonzept zutrifft, kommen lediglich atomare Aussagen in Form der einzelnen Raumkonzepte in Frage. Das bedeutet, dass die Raumkonzepte durchaus in einer hierarchischen Struktur auftreten, doch werden lediglich die Konzepte auf der untersten Stufe betrachtet. Den Raumkonzepten, die sich noch in weitere Unterkonzepte zergliedern, werden keine Glaubensmaße zugewiesen, da bei der hier vorgestellten Anwendung nur die untersten Raumkonzepte der Hierarchie identifiziert werden sollen. Das hat den Vorteil, dass die Glaubensmaße und die Beliefs bezüglich der Raumkonzepte gleich behandelt werden können. Der Frame of Discernment besteht somit nur aus disjunkten Konzepten, welche in (Gordon und Shortliffe, 1985) auch *Singleton Hypotheses* genannt werden.

In Bezug auf die ontologische Wissensmodellierung von Raumkonzepten, wie sie oben vorgestellt wurde (siehe Kapitel 4.1.2), ergibt sich somit folgendes Verhältnis: Der Raum als physikalische Einheit, der sich durch verschiedene Objekte auszeichnet, kann ein Glaubensmaß in Bezug auf ein Raumkonzept besitzen. Dies ist anschaulich in Abbildung 4.3 dargestellt. Auf diese Weise wird die ontologische und epistemologische Wissensmodellierung kombiniert.

Auf der Seite des sicheren (ontologischen) Wissens wird der physikalische Raum modelliert, der verschiedene Objekte beinhalten (*contains*) kann. Ebenso können die

Niches modelliert werden, die in diesem Fall die funktionalen Aspekte von Räumen modellieren. Wie oben bereits erläutert, existieren physikalische Räume als Konzepte, die die Eigenschaft besitzen, dass sie Niches, also Raumkonzepte, unterstützen können (*may support*). Insofern stellt sich die obige Modellierung durch Niches auch als sehr geeignet dar, um um Beliefwerte erweitert zu werden. Denn die Beziehung „may support“ wird mit dem Glaubensmaß aus der Dempster-Shafer-Theorie belegt und die oben beschriebene Relation *may support niche* wird durch die Eigenschaft eines Beliefs erweitert. Hier wird also das unsichere (epistemologische) Wissen modelliert.

### 4.3.1 Glaubensmaß-Modellierung in OWL

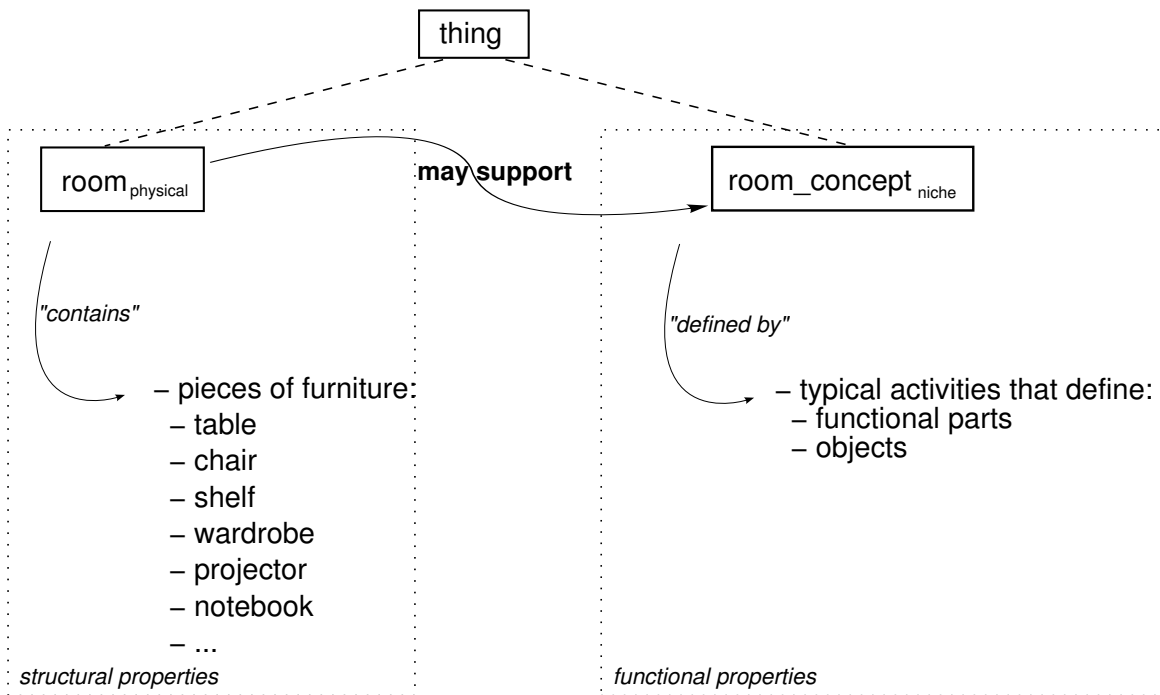
Um die Glaubensmaße aus der Dempster-Shafer-Theorie zu modellieren, muss die Beziehung zwischen dem physikalischen Raum und dem Raumkonzept (der Niche) mit einem Belief belegt werden. Da in OWL DL nur binäre Relationen darstellbar sind, die keine zusätzlichen Informationen beinhalten können, werden die Glaubensmaße in der Ontologie als Konzept modelliert. Dieses Konzept *niche\_evidence* hat drei Eigenschaften, deren Kardinalität jeweils 1 beträgt:

1. *refersToPhysical\_room* (owl:ObjectProperty): Der Range dieser Beziehung betrifft das Konzept *physical\_room*, das die physikalischen Räume modelliert.
2. *refersToNiche* (owl:ObjectProperty): Der Range dieser Relation bezieht sich auf das Konzept *Niche*. Unterklasse dieses Konzepts sind die konkreten Raumkonzepte.
3. *has\_belief\_value* (owl:DatatypeProperty): Hiermit wird das Glaubensmaß gemäß der Dempster-Shafer-Theorie modelliert. Der Range dieser Eigenschaft kann dabei eine Dezimalzahl zwischen 0 und 1 sein.

Dass bei der Instanziierung der *niche\_evidence*-Objekte mehrfach einem physikalischen Raum dasselbe Raumkonzept zugewiesen werden kann, das heißt es wird für exakt einen physikalischen Raum dasselbe Raumkonzept zwei Mal instanziiert mit jeweils gleichen oder unterschiedlichen Glaubensmaßen, ist unproblematisch. Denn so können die Glaubensmaße unterschiedlicher Quellen modelliert werden.

### 4.3.2 Anfragen bezüglich konkreter Raumkonzepte

Die Modellierung des unsicheren Wissens durch Glaubensmaße der Dempster-Shafer-Theorie reicht allerdings noch nicht aus: Die Aufgabe liegt weiterhin darin, aufgrund



ontological  
epistemological

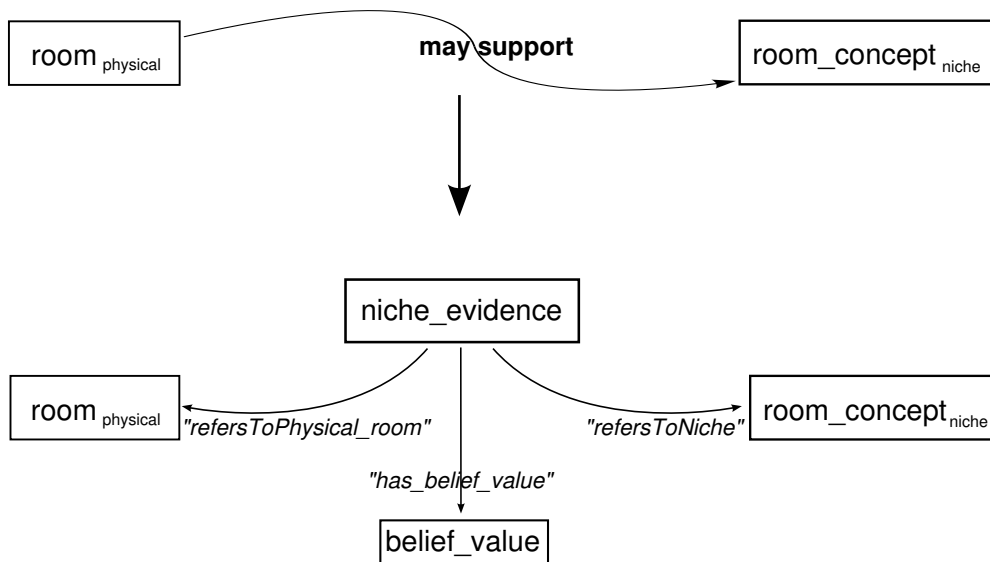


Abbildung 4.3: ontologische (oben) und epistemologische (unten) Sichtweise

einer Anzahl an bekannten Einrichtungsgegenständen, die den physikalischen Raum auszeichnen, herauszufinden, um welches Raumkonzept es sich handelt. Dafür müssen Anfragen an die Ontologie gestellt werden, die ein Reasoner auswertet.

Oben wurde bereits erwähnt, dass eine Ontologie in OWL DL aus einem terminologischen Teil (der im vorigen Kapitel beschrieben wurde) und einem assertorischen Teil besteht. Der assertorische Teil, die ABox, gibt Informationen über die Instanzen, welche ein fester Bestandteil der Ontologie sind. Kommen entsprechende Instanzen des Konzepts *niche\_evidence* in der Ontologie vor, können bezüglich des Wissens über diese Instanzen Anfragen gestellt werden.

Wird nun aktuell ein physikalischer Raum untersucht und sind bereits erste Objekte, die in diesem Raum enthalten sind, bekannt, so kann eine Ontologieanfrage gestartet werden. Diese trägt zuerst alle Instanzen von physikalischen Räumen in der Ontologie zusammen, in denen die gleichen Objekte (und eventuell mehr) enthalten sind, wie in dem aktuellen Raum. Im darauf folgenden Schritt muss untersucht werden, welche Instanzen von *niche\_evidence* mit diesen Räumen in Beziehung stehen und welche Raumkonzepte zu welchem Glaubensmaß dadurch unterstützt werden. Die Beliefs der so ermittelten Raumkonzepte können mittels Dempster's Rule of Combination berechnet und als Antwort ausgegeben werden. Wie bei der Berechnung für die Hypothesen einzelner Objekte (siehe Kapitel 3.3), werden auch hier nur die unterstützenden Beliefs berücksichtigt.

Die Modellierung der Glaubensmaße und ihre Verwendung bei Anfragen werden in Kapitel 5 am Beispiel von Universitätsräumen konkret umgesetzt und untersucht.

## 5 Beispielontologie: Raumkonzepte

Die vorangegangenen Überlegungen bezüglich der Berücksichtigung unterschiedlicher Ontologiesichtweisen und die Kombination von Ontologien und unsicherem Wissen sollen an einem konkreten Beispiel verdeutlicht werden. Das Ziel soll darin bestehen, anhand eines Bildausschnitts eines Raumes das Raumkonzept zu klassifizieren, um das es sich bei dem Bildausschnitt handeln könnte. Dabei werden für die Analyse einzelne Objekte (Einrichtungsgegenstände) im Raum nacheinander extrahiert und analysiert. Informationen über die Raumkonzepte und die enthaltenen Gegenstände sind in einer Ontologie gespeichert. Diese Ontologie nimmt die oben beschriebene Struktur ein. Anhand der analysierten Objekte können Anfragen gestellt werden, welches Raumkonzept vorliegen kann. Das ermittelte Konzept wird mit einem Belief nach Maßgabe der Dempster-Shafer-Theorie angegeben.

Die Analyse der einzelnen Objekte soll anhand des bestehenden Okusys-Systems (siehe Kapitel 3.3) erfolgen. Hier werden die einzelnen Objekte gelernt. Das System wird erweitert, um Objekte extrahieren und Anfragen an die Ontologie starten und auswerten zu können. Damit wird das System semantisch um eine Abstraktionsstufe erweitert: Bisher wurden einzelne Objekte analysiert und gelernt. Nun können auch mehrere Objekte im Bild analysiert und auf der Basis dieser Analyse die Objektkonfiguration, also das Raumkonzept, analysiert werden. In Anlehnung an den menschlichen Sehvorgang wird so eine Unterscheidung getroffen zwischen den Sakkaden, die auf einem Objekt ausgeführt werden und hier als *Intrasakkaden* bezeichnet werden, und den Augenbewegungen, die von einem zum nächsten Objekt führen und als *Intersakkaden* bezeichnet werden.

### 5.1 Niches: Raumkonzepte

Die zu identifizierenden Konzepte der Räume sind entstanden, indem empirisch Material erfasst und ausgewertet wurde (siehe Kapitel 5.5.1). Die einzelnen Raumkonzepte sind: Seminarraum (classroom), Hörsaal (lecture hall), Küche (kitchen), Praktikumsraum (laboratory) und Büro (office). Die Kategorisierung erfolgt anhand der Funktion, die ein Raum ausfüllt. Diese fünf Raumkonzepte sind hierarchisch strukturiert: Seminarraum und Hörsaal sind Unterklasse des Konzepts Veranstaltungsraum, Büro und

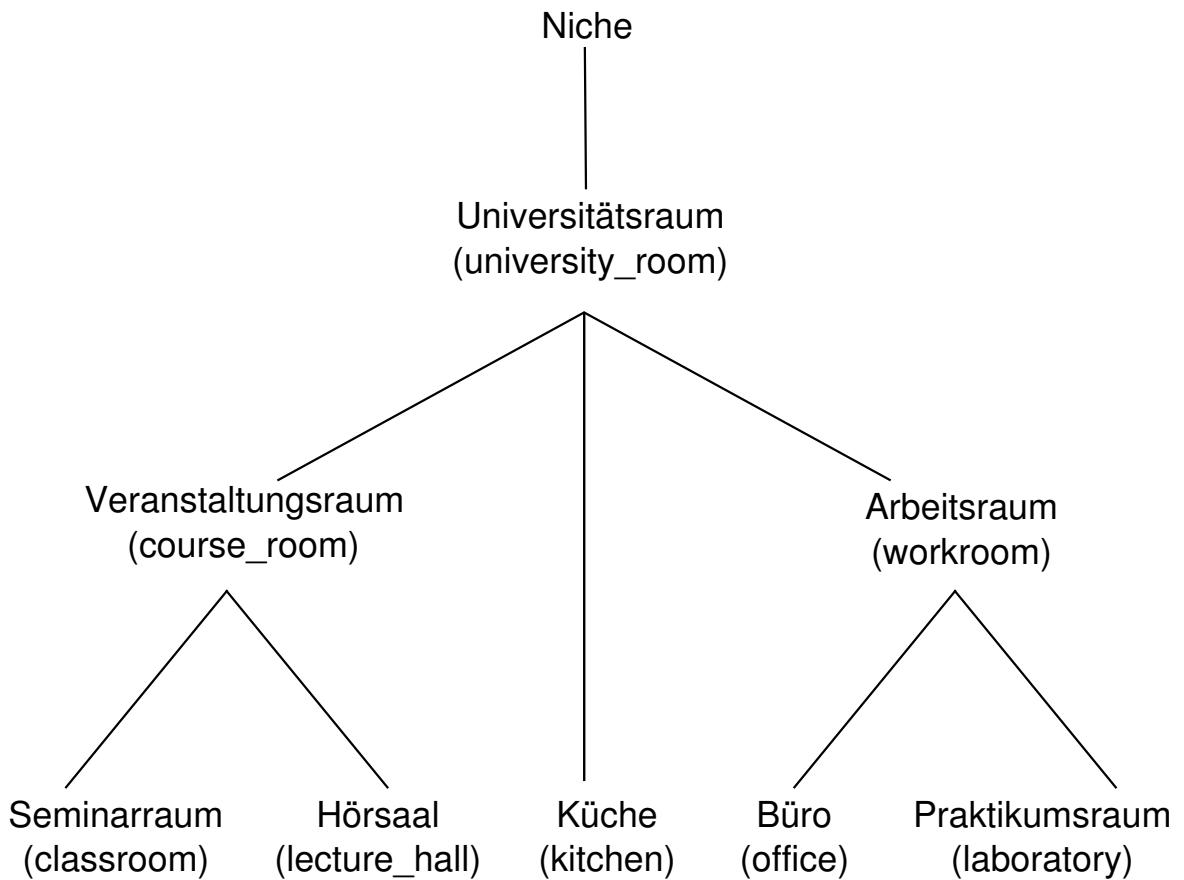


Abbildung 5.1: Universitätsraumkonzepte

Praktikumsraum sind Unterklasse von Arbeitsraum. Die Konzepte Veranstaltungsraum, Arbeitsraum und Küche sind Unterklassen des Konzepts Universitätsraum. Da diese Konzepte mit Ausnahme ihrer hierarchischen Struktur keine weiteren Eigenschaften besitzen, soll der OWL-Quellcode, in dem sie modelliert sind, hier nicht aufgeführt werden. Stattdessen ist die Struktur in Abbildung 5.1 dargestellt.

Wie bereits erwähnt, sind die Räume selbst physikalische Einheiten. Sie können aber die Funktion von einem oder mehreren der genannten Raumkonzepte besitzen. Daher entsprechen die Raumkonzepte dem genannten Konzept der Niche (siehe Kapitel 4.1.2). So ist in diesem Fall das Konzept Universitätsraum Unterklasse von Niche.

## 5.2 Physical Endurant: Raum

Der Raum bzw. Raumausschnitt besteht unabhängig von seiner Funktion aus geometrischen Ausmaßen und enthaltenen Objekten. Die Ausmaße, wie Länge, Breite, Höhe etc., werden hier nicht betrachtet, da bei den hier beschriebenen Raumkonzepten nur geringe Rückschlüsse auf eine Raumfunktion anhand der Ausmaße gezogen werden kann und das Bildmaterial diese Daten nicht hergibt. Dafür werden die Objekte betrachtet, die in einem Raum vorhanden sein können.

In der Ontologie wird ein Raumausschnitt durch das Konzept „physical\_room“ modelliert. Dieses Konzept besitzt die Eigenschaft „contains“, die eine Instanz des Konzepts mit einem „physical object“ verbindet. Diese Relation modelliert die Verknüpfung zwischen einem Raum und seinen enthaltenen Gegenständen. Die Klasse „physical\_room“ und ihre Relation ist in OWL wie folgt deklariert:

```
<owl:Class rdf:ID="physical_room">
  <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
    a physical room defined by its geographic dimension
    - can contain pieces of furniture etc.
    - can have a certain role, i.e. niche, with a certain evidence
  </rdfs:comment>
  <owl:disjointWith rdf:resource="#niche" />
  <owl:disjointWith rdf:resource="#interior" />
  <owl:disjointWith rdf:resource="#niche_evidence" />
</owl:Class>

<owl:ObjectProperty rdf:about="#contains">
  <rdfs:domain rdf:resource="#physical_room" />
  <rdfs:range rdf:resource="#physical_object" />
</owl:ObjectProperty>
```

Dabei ist zu bemerken, dass das Konzept „physical\_room“ disjunkt ist mit den Konzepten „niche“, „physical\_object“ und „niche\_evidence“. Diese vier Konzepte befinden sich in der Taxonomie auf einer Hierarchieebene (siehe Abbildung 5.3) und keine Instanz kann in mehr als einem dieser Konzepte enthalten sein. Die komplette Taxonomie ist in Anhang G abgebildet.

Bezüglich der OWL-Notation sei noch gesagt, dass die Rauten vor den Konzeptnamen (ebenso bei Instanz- oder Relationenbezeichnungen) auf den für die gesamte OWL-Datei angegebenen XML-Namespaces verweisen.

Das Property „contains“ verbindet damit den physikalischen Raum mit möglichen Einrichtungsgegenständen („interior“), also den einzelnen Objekten im Raum, und stellt somit die partonomischen Beziehungen der beiden Konzepte dar: „interior“ ist Teil von „physical\_room“.

### 5.3 Physical Object: einzelne Objekte

Alle Einrichtungsgegenstände, die aufgrund des gesichteten Bildmaterials für eine Kategorisierung in Frage kommen, sind ebenfalls hierarchisch in der Ontologie strukturiert. Die Gegenstände selbst haben keine weiteren Eigenschaften, außer dass alle untereinander disjunkt sind. Daher ist auch hier auf den OWL-Quellcode verzichtet worden. Eine Darstellung aller in der Ontologie verwendeten Einrichtungsgegenstände ist dafür in Abbildung 5.2 abgebildet.

In der Ontologie werden die einzelnen Objekte nicht weiter spezifiziert. Stattdessen korrespondieren sie mit Objekten, die im Okusys-System (siehe Kapitel 3.3) als Hypothesen auftreten. Die Objekte werden also in ihren Klassen anhand von Sakkadenzügen vom System gelernt. Werden also aus einem Bild eines Raumausschnitts die enthaltenen Objekte segmentiert, können diese einzelnen Objekte wie oben beschrieben vom System klassifiziert werden.

Wie in Kapitel 3.1 bereits erwähnt, tritt bei dieser Modellierung das Problem mit Teilen und ganzen Objekten auf. An dieser Stelle sollen die einzelnen Gegenstände, die auf der untersten Hierarchiestufe in Abbildung 5.2 liegen, als die einzelnen Objekte betrachtet werden. Diese einzelnen Objekte sind wie bereits erwähnt die Teile des Raumes, der sich aus diesen einzelnen Objekten zusammensetzt und als komplexes Objekt betrachtet wird.

### 5.4 Glaubensmaß für Raumkonzepte

In der Ontologie werden die Glaubensmaße so modelliert, wie sie in Kapitel 4.3.1 beschrieben sind. Das Konzept „niche\_evidence“ verknüpft somit einen physikalischen Raum mit seinem möglichen Raumkonzept (Niche) und belegt diese Relation mit einem Glaubensmaß der Dempster-Shafer-Theorie.

Das Konzept ist entsprechend in der Ontologie mit den drei Eigenschaften „refersToPhysical\_room“, „refersToNiche“ und „belief\_value“ umgesetzt. Die ersten beiden Eigenschaften sind owl:ObjectProperties mit dem Range „physical\_room“ bzw. „niche“, die dritte Eigenschaft ist eine DatatypeProperty und gibt das Glaubensmaß



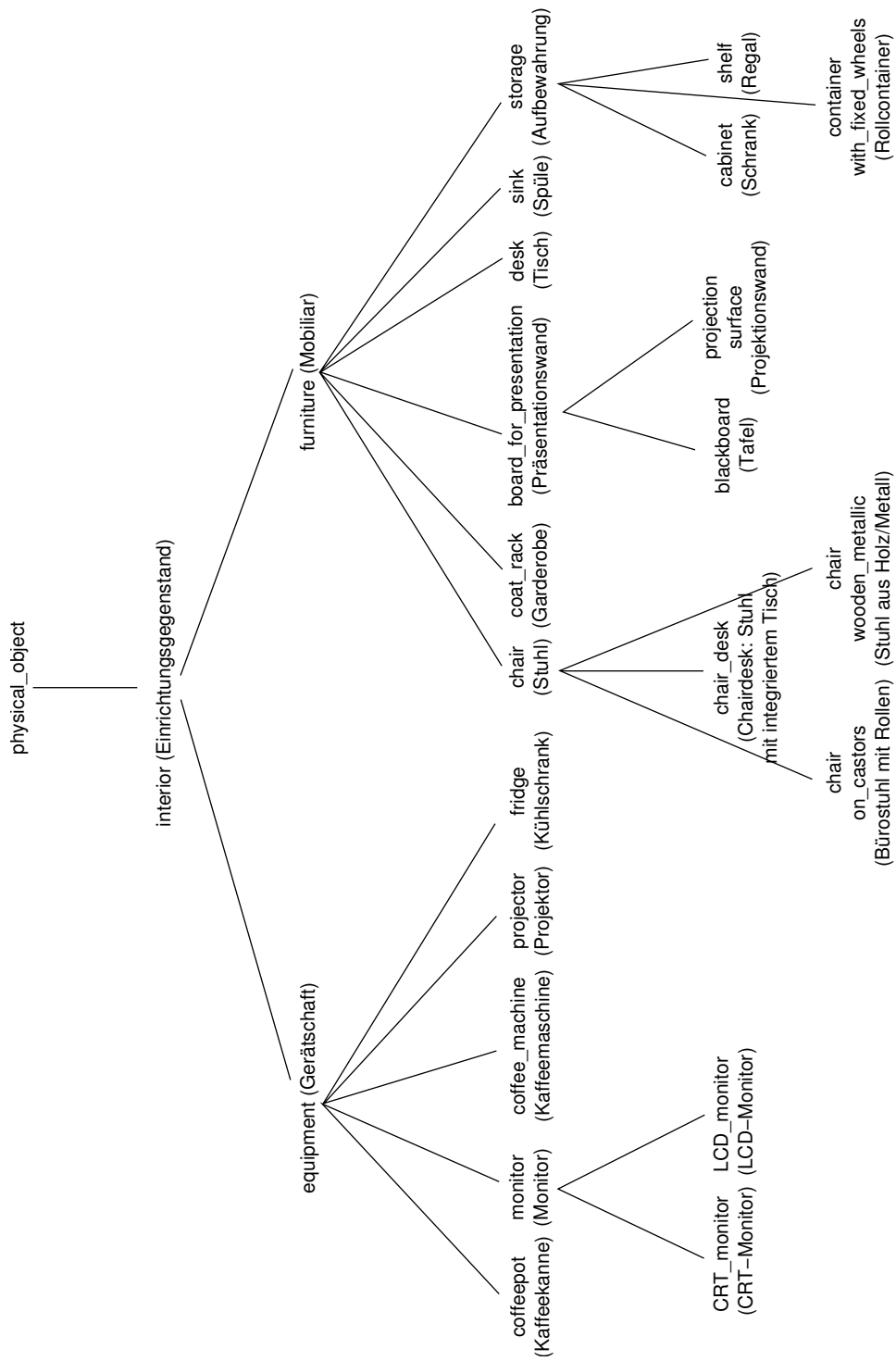


Abbildung 5.2: Einrichtungsgegenstände der Ontologie

an. Dabei muss jede Instanz von „niche\_evidence“ diese drei Properties jeweils genau ein Mal haben, weil gerade ein physikalischer Raum mit nur einem Raumkonzept über einen Belief verbunden sein kann. In OWL sieht die Modellierung für das Konzept „niche\_evidence“ und die Properties, die dieses Konzept besitzt, wie folgt aus:

```
<owl:Class rdf:ID="niche_evidence">
  <rdfs:subClassOf rdf:resource="http://www.w3.org/2002/07/owl#Thing"/>

  <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
    defines the belief-value of a physical_room to support a certain niche
  </rdfs:comment>

  <owl:disjointWith>
    <owl:Class rdf:about="#physical_object"/>
  </owl:disjointWith>
  <owl:disjointWith>
    <owl:Class rdf:about="#physical_room"/>
  </owl:disjointWith>
  <owl:disjointWith>
    <owl:Class rdf:about="#niche"/>
  </owl:disjointWith>

  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty>
        <owl:DatatypeProperty rdf:ID="has_belief_value"/>
      </owl:onProperty>
      <owl:cardinality rdf:datatype="http://www.w3.org/2001/XMLSchema#int">
        1
      </owl:cardinality>
    </owl:Restriction>
  </rdfs:subClassOf>

  <rdfs:subClassOf>
    <owl:Restriction>
      <owl:onProperty>
        <owl:ObjectProperty rdf:ID="refersToNiche"/>
      </owl:onProperty>
    </owl:Restriction>
  </rdfs:subClassOf>
</owl:Class>
```

```

        <owl:cardinality rdf:datatype="http://www.w3.org/2001/XMLSchema#int">
            1
        </owl:cardinality>
    </owl:Restriction>
</rdfs:subClassOf>

<rdfs:subClassOf>
    <owl:Restriction>
        <owl:onProperty>
            <owl:ObjectProperty rdf:ID="refersToPhysical_room"/>
        </owl:onProperty>
        <owl:cardinality rdf:datatype="http://www.w3.org/2001/XMLSchema#int">
            1
        </owl:cardinality>
    </owl:Restriction>
</rdfs:subClassOf>

</owl:Class>

<owl:DatatypeProperty rdf:about="#has_belief_value">
    <rdfs:comment rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
        belief_value of the Dempster-Shafer-Theory of evidence [0..1]
    </rdfs:comment>
    <rdfs:domain rdf:resource="#niche_evidence" />
    <rdfs:range rdf:resource="http://www.w3.org/2001/XMLSchema#string" />
</owl:DatatypeProperty>

<owl:ObjectProperty rdf:about="#refersToNiche">
    <rdfs:domain rdf:resource="#niche_evidence" />
    <rdfs:range rdf:resource="#niche" />
</owl:ObjectProperty>

<owl:ObjectProperty rdf:about="#refersToPhysical_room">
    <rdfs:domain rdf:resource="#niche_evidence" />
    <rdfs:range rdf:resource="#physical_room" />

```

</owl:ObjectProperty>

Eine graphische Darstellung der Relationen zwischen den modellierten Konzepten ist in Anhang G abgebildet.

Die Beschränkung darauf, dass „niche\_evidence“ auch genau jeweils eine der drei Eigenschaften besitzt, werden in OWL durch *Restrictions* ausgedrückt. Eine Restriction kann entsprechende Einschränkungen auf Klassen spezifizieren. Dabei sind Restrictions Konzepte in OWL, von denen die Klasse, die durch eine solche Restriction eingeschränkt wird, erbt. Entsprechend werden die drei Eigenschaften von „niche\_evidence“ durch Unterklassen-Beziehungen eingeleitet.

Bei der Modellierung von „belief\_value“ ist außerdem zu beachten, dass die Datatype-Property durch einen String repräsentiert wird. Die Eigenschaft soll allerdings das Glaubensmaß für ein Raumkonzept angeben. Insofern wäre es zwar sinnvoller eine Dezimalzahl zwischen 0 und 1 zu verwenden anstelle eines Strings. Allerdings wird bei einer Anfrage an die Ontologie mithilfe des Reasoners Pellet (siehe Kapitel 5.6.1) intern mit Strings gearbeitet, denn die Nutzung von Dezimalzahlen führte bei der Implementierung zu Fehlern (siehe Kapitel 6).

## 5.5 Taxonomie der Raumontologie

In Abbildung 5.3 sind die Konzepte der hier modellierten Ontologie dargestellt. Die bereits vorgestellten Konzepte, die einzelnen Objekte („physical\_object“), die physikalischen Räume („physical\_room“), die Glaubensmaße („niche\_evidence“) und die Raumkonzept-Niches („niche“), sind Unterklassen von „owl:Thing“, der Oberklasse aller Konzepte in OWL. Ausgelassen wurden die Unterklassen von „interior“ und „university\_room“. Diese sind in den vorangegangenen Abbildungen bereits dargestellt. Eine ausführliche Darstellung der Taxonomie ist in Anhang G abgebildet.

Mit dieser Taxonomie und den oben erläuterten partonomischen Beziehungen zwischen einzelnen Konzepten ist die grundlegende Modellierung der Raumkonzepte formuliert. Schließlich fehlen noch die Instanzen der Konzepte, um physikalische Räume analysieren und identifizieren zu können, da die Ontologieanfragen die Instanzen betreffen. Es müssen also Räume, die Gegenstände enthalten und Raumkonzepte in Höhe bestimmter Beliefwerte unterstützen, graphisch modelliert werden. Die Vorgehensweise der Instanziierungen wird im Folgenden beschrieben.

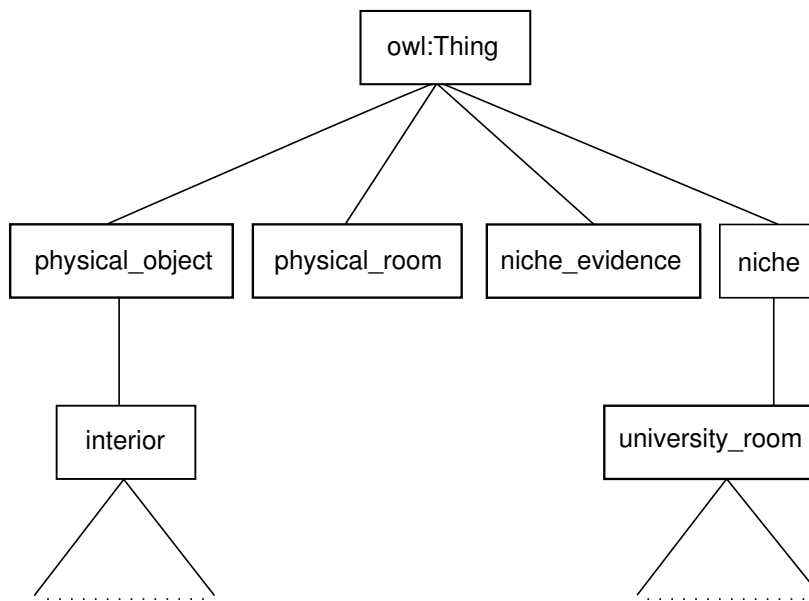


Abbildung 5.3: Taxonomie der Ontologie

### 5.5.1 Bildmaterial

Das Bildmaterial beruht auf Fotoaufnahmen von verschiedenen Räumen der Universität Bremen, hauptsächlich aus dem Mehrzweckhochhaus (MZH). Dabei sind solche Räume ausgewählt worden, deren Nutzungsweise in die Rubriken der fünf gewählten Raumkonzepte fallen. Insgesamt sind Aufnahmen von vier Büros, fünf Küchen, fünf Praktikumsräumen, fünf Seminarräumen und drei Hörsälen aus verschiedenen Blickwinkeln erstellt worden (siehe beifügte CD, Anhang D).

### 5.5.2 Graphische Bearbeitung des Bildmaterials

Da der Fokus dieser Arbeit nicht auf der Segmentierung von Objekten liegt, erfolgt die Extraktion der einzelnen Objekte anhand eines trivialen Algorithmus: Jedem Objekt im Gesamtbild wird ein eindeutiger Grauwert zugewiesen, anhand dessen das Objekt segmentiert wird. Infolgedessen kann nicht auf den Originaldaten gearbeitet werden. Dafür müssen Testbilder auf Basis dieser Daten erstellt werden.

Die Überschneidung von Objekten im Bild, so dass diese nur noch teilweise sichtbar sind, stellt ein weiteres Problem dar. Möglichkeiten zur Lösung dieser komplexen Fragestellung sollen an dieser Stelle nicht weiter betrachtet werden. Daher werden Überschneidungen von Objekten in den Testbildern vermieden.

Ausgehend von den realen Fotoaufnahmen werden zu jedem Raumkonzept jeweils fünf Instanzen erstellt. Somit ergeben sich 25 Raumausschnitte, die im Anhang E dargestellt sind. Dabei sind die Instanzen teilweise sehr nach Vorgabe des Bildmaterials entstanden (Seminarräume und Praktikumsräume), teilweise aber auch freier kreiert (Küchen und Büros). Beispielsweise liegt bei den Küchen die Ursache für die eher unabhängig vom Bildmaterial entworfenen Raumausschnitte darin, dass die Objekte in den Originalbildern nur schwer zu modellieren waren. Die Einbauschränke besitzen keine allzu charakteristischen Eigenschaften und sind hinderlich für die geforderte Überschneidungsfreiheit der Objekte, da sie viel Platz im Bild einnehmen.

Probleme bei der Konfiguration einzelner Raumszenen stellen aber auch komplexere Gegenstände dar. Dies ist insbesondere bei der graphischen Modellierung der Hörsäle der Fall: Hier setzen sich die einzelnen Sitzreihen aus den herunterklappbaren Sitzplätzen und den ebenfalls klappbaren Tischflächen der davor liegenden Reihe zusammen. Rein graphisch erweist es sich als nahezu unmöglich, solche Sitzreihen ohne Überschneidung und in vollem Umfang auf den Testbildern darzustellen. Ebenso problematisch ist die Abstraktionsebene eines solchen Mobiliars. Sie liegt aufgrund ihrer Zusammensetzung mehrerer einzelner, funktionaler Elemente bereits höher als die in den übrigen Fällen dargestellten Objekten, da sie eine größere Anzahl an Tischen und Sitzmöglichkeiten kombiniert, während in den übrigen Fällen Tische und Stühle lediglich einzeln vorkommen. Aufgrund dieser Problematik wurde darauf verzichtet, die Sitzreihen des Hörsaals „großer Hörsaal - HS 2010“ (siehe beigefügte CD, Anhang D) graphisch abstrakt in die Instanzen zu integrieren.

Bei der graphischen Bearbeitung der einzelnen Objekte wurden Objekte entweder direkt aus den Originalbildern extrahiert (Stühle, Monitore, Projektoren, Kaffeemaschinen/kannen), auf Basis der Originaldaten abstrahiert und erstellt (Schränke, Regale, Rollcontainer) oder es gab eine Kombination aus beiden Verfahren (Spülbecken, Tafeln). Dies lag zum einen am vorhandenen Bildmaterial, aus dem sich die entsprechenden Objekte wegen Überschneidungen oder ungeeigneten Perspektiven nicht problemlos extrahieren ließen, zum anderen daran, dass sich durch eine freiere Zusammenstellung der Raumausschnitte mehr Objekte anordnen ließen.

Zwei der Garderoben wurden zum Teil so modifiziert, dass sie ebenfalls besser im Bild zu integrieren sind. So wurden nur die Aufhängungsmöglichkeiten für die Garderobe aus dem Bildmaterial extrahiert, nicht aber die kompletten Garderoben mit fahrbarem Gestell. Dies hätte im Bild zuviel Platz eingenommen.

Die Auswahl der Objekte hängt in erster Linie davon ab, wie häufig sie in den Räumen enthalten sind: Je häufiger Objekte einer Kategorie in den Räumen auftritt, desto eher werden sie modelliert. Auch die Größe spielt eine Rolle: Sind die Objekte zu klein, werden sie nicht berücksichtigt. Dasselbe gilt, wenn die Objekte zu groß sind

und nicht mehr auf den Testbildern darstellbar sind. Auch Objekte, die nicht direkt im Raum „enthalten“ sind, beispielsweise Fenster, sind nicht berücksichtigt.

### 5.5.3 Glaubensmaße

Die Glaubenswerte der einzelnen Rauminstanzen in Bezug auf ihr mögliches Raumkonzept ergeben sich aus der Befragung von Experten. Dabei wurden zwei Experten befragt, indem sie die einzelnen Rauminstanzen individuell beurteilen sollten. Sie bekamen die Aufgabe anzugeben, mit welchem Glauben sie eine Instanz einem Raumkonzept zuordnen würden. Die Ergebnisse sind in Anhang F zu finden.

Den Experten wurden die fünf genannten Kategorien der Raumkonzepte als Klasseneinteilung vorgelegt. Sie wurden daran erinnert, dass es sich hierbei um eine funktionale Einteilung handelt. Sie sollten jede Rauminstanz individuell betrachten und unabhängig von den vorher betrachteten Instanzen ihre Glaubenswerte festhalten. Dabei vergebene Werte sollten zwischen 0 und 1 liegen. Die Glaubenswerte einer Instanz mussten weder in der Summe ihrer möglichen Raumkonzepte 1 ergeben - so wird den unterstützenden (confirming) Glaubenswerten Rechnung getragen - noch mussten Glaubenswerte für alle Raumkonzepte in Bezug auf eine Instanz vergeben werden.

Von den so ermittelten Werten wurde das arithmetische Mittel gebildet und als Eigenschaft in die Ontologie integriert. Die Mittelwertbildung ist unabhängig von der Dempster-Shafer-Theorie zu betrachten. Sie soll im Vorfeld einen Ausgleich zwischen zwei Expertenmeinungen schaffen. So können Extremwerte aus der Expertenbefragung in der Ontologie nicht zu unerwünscht hohen oder niedrigen Beliefs bei der Berechnung führen.

Bei der Mittelwertbildung ist Folgendes zu ergänzen: Hat nur einer der Experten bei einem Raumkonzept für eine Rauminstanz einen Wert vergeben, so wurde für den anderen Experten ein Glaubensmaß von 0 unterstellt. Beispielsweise lag das Glaubensmaß eines Hörsaals für Raum 01 bei Experte 1 bei 0.8, Experte 2 hat allerdings keinen Wert hier vergeben. Um keine Ungleichgewichtung gegenüber den anderen gemittelten Werten, bei denen beide Experten Glaubenswerte genannt haben, zu bekommen und um die eben bereits erwähnten, unerwünschten Extrema zu relativieren, wurde dieser Wert auf 0.4 gemittelt.

### 5.5.4 Instanzen der Ontologie

Anhand dieser Glaubenswerte und den graphisch modellierten Raumausschnitten lassen sich Instanzen für die Ontologie bilden. Ein Beispiel einer Instanz für den Raumausschnitt „office01“ (siehe Abbildung 5.4) ist in OWL wie folgt definiert:



Abbildung 5.4: Rauminstanz „office01“

```
<physical_room rdf:ID="office01">
  <contains rdf:resource="#container_with_fixed_wheels01" />
  <contains rdf:resource="#cabinet04" />
  <contains rdf:resource="#chair_on_castors21" />
  <contains rdf:resource="#desk66" />
  <contains rdf:resource="#chair_wooden_metallic56" />
  <contains rdf:resource="#CRT_monitor24"/ >
  <contains rdf:resource="#shelf01"/ >
  <contains rdf:resource="#chair_on_castors20"/ >
  <contains rdf:resource="#LCD_monitor01"/ >
</physical_room>

<niche_evidence rdf:ID="niche_evidence54_office01">
  <refersToPhysical_room rdf:resource="#office01" />
  <has_belief_value rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
    0.2
  </has_belief_value>
  <refersToNiche rdf:resource="#laboratory_room" />
</niche_evidence>
```



```
<niche_evidence rdf:ID="niche_evidence55_office01">
  <refersToPhysical_room rdf:resource="#office01" />
  <refersToNiche rdf:resource="#office_room" />
  <has_belief_value rdf:datatype="http://www.w3.org/2001/XMLSchema#string">
    0.825
  </has_belief_value>
</niche_evidence>
```

Die Instanz „office01“ ist demnach dadurch charakterisiert, dass sie die neun aufgeführten Objekte beinhaltet („contains“):

- „container\_with\_fixed\_wheels01“ - Instanz des Konzepts „Rollcontainer“
- „cabinet04“ - Instanz des Konzepts „Schrank“
- „chair\_on\_castors21“ - Instanz des Konzepts „Bürostuhl mit Rollen“
- „desk66“ - Instanz des Konzepts „Tisch“
- „chair\_wooden\_metallic56“ - Instanz des Konzepts „Stuhl aus Holz/Metall“
- „CRT\_monitor24“ - Instanz des Konzepts „CRT-Monitor“
- „shelf01“ - Instanz des Konzepts „Regal“
- „chair\_on\_castors20“ - Instanz des Konzepts „Bürostuhl mit Rollen“
- „LCD\_monitor01“ - Instanz des Konzepts „LCD-Monitor“

Die Bezeichnung der Objektinstanzen ergibt sich dabei aus dem Konzeptnamen mit einer angehängten Zahl. Diese Zahl ist ohne Belang; während der Ontologiemodellierung wurden die erstellten Instanzen lediglich nummeriert. Für jede erstellte Instanz eines „physical\_room“-Konzepts wurden die enthaltenen Objekte des Raumes instanziiert und in „contains“-Relation zur Instanz gesetzt. So ist auch sichergestellt, dass kein instanziiertes Raumobjekt in mehreren physikalischen Räumen auftritt, da ein Objekt sich in nur einem Raum zur Zeit befinden kann und die Ontologie keine Zeitangaben unterstützt, dass beispielsweise Objekte zu bestimmten Zeiten in unterschiedlichen Räumen enthalten sind. Die Instanzen der physikalischen Objekte, also der Einrichtungsgegenstände, modellieren entsprechend diejenigen Objekte, die im Bild dargestellt sind.

Die Instanzen vom Typ „niche\_evidence“ sind in der hier modellierten Ontologie ebenfalls nummeriert und enden auf den Namen des physikalischen Raums, mit dem sie das funktionale Niche-Raumkonzept in Beziehung setzen. So verknüpft die Instanz „niche\_evidence54\_office01“ den Raum zu einem Glaubenswert von 0.2 mit der Instanz des Raumkonzepts Praktikumsraum („laboratory\_room“). Die Instanz „niche\_evidence55\_office01“ verknüpft „office01“ mit der Raumkonzeptinstanz „office\_room“ mit einem Glaubenswert von 0.825. Die jeweiligen Werte sind auch in Anhang F zu sehen.

Erklärend ist hinzuzufügen, dass in OWL Instanzen nicht mit Konzepten in Beziehung stehen können, die Instanz „office01“ also nicht direkt mit dem Konzept „office“ sondern nur mit einer Instanz des Konzepts verknüpft werden kann. Diese Instanz „office\_room“ kann dann aber von allen anderen „physical\_room“-Instanzen als funktionale Eigenschaft genutzt werden.

Insgesamt sind sämtliche in Anhang E abgebildeten Raumausschnitte als Instanz von „physical\_room“ in der Ontologie dargestellt. Ebenso die enthaltenen Objekte in Form der „contains“-Relation. Abhängig von den Expertenbefragungen können die Instanzen Verbindungen zu den Raumkonzepten „classroom“, „lecture\_hall“, „kitchen“, „office“ und „laboratory“ mit einem gewissen Glaubensmaß haben. Von jedem der fünf Raumkonzepte wird einmalig eine Instanz in der Ontologie modelliert. Sie steht semantisch für den funktionalen Charakter, den eine Instanz von „physical\_room“ einnehmen kann. Ihre Namen ergeben sich dabei aus dem Namen des Konzepts und einem angehängten „\_room“.

## 5.6 Reasoning

Die Ontologie mit ihren modellierten Instanzen kann nun durch einen Reasoner bezüglich ihrer Konsistenz geprüft werden. Außerdem sollen Anfragen an die Ontologie gestellt werden, die Aufschluss über Raumkonzepte geben. Wie bereits erwähnt wurde, liegt die Aufgabe darin, anhand identifizierter Objekte im Bild zu ermitteln, welche Raumkonzepte vom Raumausschnitt mit welchem Belief unterstützt werden.

In dieser Arbeit wird der Reasoner Pellet benutzt, der im Folgenden kurz vorgestellt werden soll.

### 5.6.1 Reasoner: Pellet

Pellet<sup>1</sup> ist ein Reasoner für OWL DL. Die Wahl, diesen Reasoner in der vorliegenden Arbeit einzusetzen, begründet sich darin, dass es sich um eine Open-Source-Implementierung in Java handelt. Auch die Erweiterung des Okusys-Systems ist in Java umgesetzt. Näheres zur Implementierung ist in Kapitel 6 angegeben.

Wie in Kapitel 5.4 bereits erwähnt, kann Pellet allerdings nicht mit Dezimalzahlen für owl:DatatypeProperties umgehen, so dass in dieser Modellierung Strings für die Angabe der Beliefwerte verwendet werden.

Bei der Konsistenzprüfung der Ontologie und dem Schlussfolgern aus den taxonomischen Daten nutzt Pellet das Tableau-Kalkül für die Beschreibungslogik OWL DL. Daneben bietet der Reasoner die Möglichkeit, Anfragen an die Ontologie zu stellen und auszuwerten. Die Anfragen dürfen sich aber nur auf Instanzen beziehen. Dabei kommt die Ontologieanfragesprache RDQL zum Einsatz. Das weiteren wird eine Online-Demo<sup>2</sup> angeboten.

### 5.6.2 Anfragesprache: RDQL

Die Resource Description Framework Query Language (RDQL)<sup>3</sup> ist eine datenorientierte Anfragesprache für Ontologien. Sie baut auf Jena<sup>4</sup>, einem Framework für das Semantic Web, auf und ermöglicht es, gezielt nach Instanzen oder ihren Eigenschaften aufgrund bestimmter Muster zu suchen. Ihre interne Repräsentation der Ontologie ist graphbasiert und spiegelt sich in der Art möglicher Anfragen wieder. Die Syntax einer Anfrage ähnelt SQL-Anfragen und besteht aus folgenden Bausteinen:

```
SELECT <gesuchte Variable[n]>
  [FROM <URI der Ontologie, in der gesucht werden soll>]
  WHERE <Selektion der Variablen durch Triple des Graphschemas>
  [AND <zusätzliche Einschränkungen durch Boolesche Ausdrücke>]
  [USING <verwendete Abkürzungen der URI>]
```

Auf diese Weise können nun die oben beschriebenen Raumkonzepte anhand der enthaltenen Objekte eines Raumausschnitts erfragt werden. Angenommen, in dem obigen Beispiel von „office01“ seien schon die Objekte Tisch, Stuhl, Schrank extrahiert

---

<sup>1</sup><http://www.mindswap.org/2003/pellet/index.shtml>

<sup>2</sup><http://www.mindswap.org/2003/pellet/demo.shtml>

<sup>3</sup><http://www.w3.org/Submission/2004/SUBM-RDQL-20040109/>

<sup>4</sup><http://jena.sourceforge.net/>

worden. So würde die Ontologieanfrage nach den physikalischen Räumen mit den gleichen Objekten wie folgt lauten:

```
SELECT ?physical_roomInstance
  WHERE (?physical_roomInstance, <base:contains>, ?physical_object01)
        (?physical_object01, rdf:type, <base:desk>)
        (?physical_roomInstance, <base:contains>, ?physical_object02)
        (?physical_object02, rdf:type, <base:chair>)
        (?physical_roomInstance, <base:contains>, ?physical_object03)
        (?physical_object03, rdf:type, <base:cabinet>)
USING base FOR <http://www.informatik.uni-bremen.de/~joana/roomOntology.owl#>
```

Es werden mit dieser Anfrage all diejenigen Instanzen („?physical\_roomInstance“) erfragt, die jeweils ein Objekt („physical\_object01“, „physical\_object02“ und „physical\_object03“) beinhalten („contains“), die vom Typ („rdf:type“) Tisch, Stuhl und Schrank („desk“, „chair“ und „cabinet“) sind. Die URI der Ontologie lautet dabei „http://www.informatik.uni-bremen.de/~joana/roomOntology.owl“.

Reicht man diese Anfrage mit der hier modellierten Ontologie (beigefügt auf der CD, siehe auch Anhang D) an Pellet weiter, so werden die in Tabelle 5.1 dargestellten Angaben ausgegeben.

Tabelle 5.1: Ergebnis der Anfrage bezüglich der (physikalischen) Rauminstanzen

physical_roomInstance
http://www.informatik.uni-bremen.de/~joana/roomOntology.owl#office04
http://www.informatik.uni-bremen.de/~joana/roomOntology.owl#kitchen01
http://www.informatik.uni-bremen.de/~joana/roomOntology.owl#kitchen05
http://www.informatik.uni-bremen.de/~joana/roomOntology.owl#office02
http://www.informatik.uni-bremen.de/~joana/roomOntology.owl#office01

Diese Rauminstanzen müssen nun im nächsten Schritt dahingehend untersucht werden, in welchen „niche\_evidence“-Relationen sie enthalten sind und welche Evidenzen damit auf welche Raumkonzepte hindeuten. Im nächsten Schritt müssen daher Anfragen für jede Rauminstanz gestellt werden. Beispielhaft würde eine solche Anfrage für die Instanz „office04“ wie folgt lauten:

```
SELECT ?niche_evidenceInstance, ?beliefValue, ?nicheInstance
  WHERE (?niche_evidenceInstance, rdf:type, <base:niche_evidence>)
```

```
(?niche_evidenceInstance, <base:refersToPhysical_room>, <base:office04>)
(?niche_evidenceInstance, <base:has_belief_value>, ?beliefValue)
(?niche_evidenceInstance, <base:refersToNiche>, ?nicheInstance)
USING base FOR <http://www.informatik.uni-bremen.de/~joana/roomOntology.owl#>
```

Die Anfrage an Pellet liefert das in Tabelle 5.2 dargestellte Ergebnis. (In Pellet wird zusätzlich bei der Angabe der Evidenzen der Typ der DatatypeProperty, also „String“, angegeben.)

Tabelle 5.2: Ergebnis der Anfrage bezüglich der Raumkonzeptinstanzen und Glaubensmaße von „office04“

niche_evidenceInstance	beliefValue	nicheInstance
http://...#niche_evidence61_office04	0.775	http://...#office_room
http://...#niche_evidence60_office04	0.075	http://...#laboratory_room

Auf diese Weise können mit allen Ergebnissen aus Tabelle 5.1, in der die Instanzen erfragt wurden, die Objekte vom Typ Tisch, Stuhl und Schrank beinhalten, Anfragen gestellt werden bezüglich ihrer Glaubensmaße von Raumkonzepten. Über eine zusätzliche Anfrage können die Typen der Raumkonzepte („nicheInstance“) in Erfahrung gebracht werden.

Schließlich ergibt sich so eine Liste mit Raumkonzepten und Glaubensmaßen, die in Tabelle 5.3 dargestellt sind. Wie sich zeigt, sind es die modellierten Glaubensmaße, wie sie in Tabelle F.3 aus Anhang F dargestellt sind.

Basierend auf diesen Angaben lassen sich die Beliefs für die einzelnen Raumkonzepte berechnen. Die Berechnung wird im nächsten Abschnitt näher betrachtet. An dieser Stelle soll noch eine Anmerkung zu den so gestellten Anfragen an die Ontologie gemacht werden.

Es wurde erwähnt, dass mögliche Raumkonzepte eines Raumausschnitts anhand der identifizierten Objekte analysiert werden sollen. Nun darf nicht der Schluss gezogen werden, dass in der Ontologie Informationen darüber stehen würden, ob ein Objekt  $x$  zu gewissen Glaubensmaßen die Raumkonzepte  $y$  und  $z$  unterstützt. Bei der hier beschriebenen Modellierung wäre eine solche Aussage nicht sinnig. Nicht ein einzelnes Objekt, sondern ein Objekt, das im Kontext steht mit einem Raumausschnitt, der wiederum auf ein Raumkonzept hindeutet, kann Informationen bezüglich aktueller Analysen bieten. Daher werden bereits bekannte Raumkonstellationen, die ebenfalls das Objekt  $x$  besitzen, auf ihre unterstützten Raumkonzepte hin geprüft.

Tabelle 5.3: Raumkonzepte und Glaubensmaße für Raumausschnitte mit den enthaltenen Objekten Tisch, Stuhl und Schrank

Raumkonzept	Glaubensmaß	aus Rauminstanz
office	0.775	„office04“
laboratory	0.075	„office04“
classroom	0.1	„kitchen01“
laboratory	0.05	„kitchen01“
kitchen	0.775	„kitchen01“
office	0.2	„kitchen01“
office	0.05	„kitchen05“
classroom	0.025	„kitchen05“
kitchen	0.65	„kitchen05“
laboratory	0.025	„kitchen05“
office	0.9	„office02“
laboratory	0.025	„office02“
office	0.825	„office01“
laboratory	0.2	„office01“

So lassen sich auch Kombinationen von Objekten besser auswerten. Je mehr Objekte im Bild analysiert werden, desto mehr schränkt es die Auswahl möglicher Raumkonzepte ein, dadurch dass weniger in der Ontologie enthaltene Raumausschnitte existieren, die die gleichen Objekte beinhalten.

### 5.6.3 Belief-Berechnung: Dempster’s Rule of Combination

Die Berechnung der Beliefs für obige Anfragen erfolgt im Okusys-System und damit außerhalb der Ontologie. Die Glaubensmaße für Raumkonzepte werden gemäß Dempster’s Rule of Combination (siehe Kapitel 2.4.3) verknüpft. Wie oben bereits erwähnt wurde, werden hier nur die unterstützenden Hypothesen kombiniert. Das heißt, es werden jeweils die Glaubensmaße, die für dasselbe Raumkonzept sprechen, miteinander verknüpft.

Die Beschränkung auf die unterstützenden Hypothesen ist deshalb möglich, da in diesem Anwendungsfall zwei Raumkonzepte nicht in Konflikt miteinander stehen. Ein Raum kann durchaus mehr als eine funktionale Eigenschaft unterstützen, auch wenn die Modellierung der Raumkonzeptinstanzen sich nur anhand einer Funktion orientiert hat (siehe Kapitel 5.5.2). Insbesondere bei Seminarraum und Hörsaal bzw. Praktikumsraum

und Büro sind die Grenzen fließend und beide Hypothesen können auf den aktuellen Raum zutreffen. Auch die Ergebnisse der Befragung (siehe Anhang F) zeigen, dass kaum ein Raum nur ein Raumkonzept unterstützt, die Instanz „kitchen05“ besitzt beispielsweise gleich für vier Raumkonzepte Glaubensmaße.

Insofern ergibt auch die Summe aller Beliefs der verschiedenen Raumkonzepte eines Raumausschnitts nicht zwingend 1. Wenn ein Glaubensmaß beispielsweise aussagt, der vorliegende Raum sei Raumkonzept A, kollidiert das nicht mit einem zweiten Belief, der sich auf das Raumkonzept B bezieht. Ist der Glauben für A bei 0.8, heißt das also nicht, dass B weniger als 0.2 sein muss.

Die Aufgabe ist es nun, die jeweiligen „Blätter“ in der Hierarchie der Raumkonzepte auszugeben. Der aktuelle Raumausschnitt soll also den Konzepten Seminarraum, Hörsaal, Küche, Praktikumsraum oder Büro entsprechen. Das bedeutet nicht, dass andere Ergebnisse nicht möglich wären. Man bräuchte lediglich den Instanzen der physikalischen Räume in der Ontologie Glaubensmaße für entsprechende Raumkonzepte zuzuweisen. Beispielsweise könnten Instanzen über die Relation „has\_belief\_value“ auf die Raumkonzepte Veranstaltungsraum oder Arbeitsraum hindeuten. Doch soll hier die Aufgabe darin bestehen, die Raumkonzepte so konkret wie möglich anzugeben. Daher wird Dempster’s Rule of Combination separat auf die *singleton hypotheses* angewandt.

Für die obige Anfrage, deren Ergebnisse in Tabelle 5.3 dargestellt sind, würden also die einzelnen Glaubensmaße jeweils für „office“, „laboratory“, „classroom“ und „kitchen“ nach der Kombinationsregel verknüpft werden. Die daraus resultierenden Ergebnisse sind in Tabelle 5.4 angegeben.

Tabelle 5.4: Belief der Raumkonzepte für einen Raumausschnitt mit Tisch, Stuhl und Schrank

Raumkonzept	einzelne Glaubensmaße	kombinierter Belief
office	$0.775 \oplus 0.2 \oplus 0.05 \oplus 0.9 \oplus 0.825$	0.997
kitchen	$0.775 \oplus 0.65$	0.921
laboratory	$0.075 \oplus 0.05 \oplus 0.025 \oplus 0.025 \oplus 0.2$	0.332
classroom	$0.1 \oplus 0.025$	0.123

Wurden in einem Raumausschnitt also die drei Objekte Tisch, Stuhl und Schrank extrahiert und identifiziert, so ergeben sich die Beliefs von 0.997 für ein Büro, 0.921 für eine Küche, 0.332 für einen Praktikumsraum und 0.123 für einen Seminarraum.

## 5.7 Okusys-Erweiterung

Die in den vorigen Abschnitten erläuterten Verfahren sollen nun auch in das Okusys-System integriert werden, um es um die Analyse von Räumen zu erweitern. Im Folgenden soll der Programmablauf erklärt und anhand von Beispielen erläutert werden. Implementierungsdetails sind in Kapitel 6 aufgeführt.

Zuerst müssen die einzelnen Objekte gelernt werden, damit sie vom System im gesamten Raumausschnitt erkannt werden können. Dieser Teil wird im folgenden Abschnitt dargestellt. Hier wird auch die Oberfläche des Okusys-Systems, das in Kapitel 3.3 bereits beschrieben wurde, dargestellt und erläutert. Anschließend an diesen Abschnitt wird die Erweiterung des Systems vorgestellt.

### 5.7.1 Lernen der Objekte

Um anfangs die einzelnen Objekte anhand von Sakkadenzügen zu lernen, werden die Funktionen des bestehenden Okusys-Systems genutzt. In Kapitel 3.3 wurde die Funktionsweise des Systems bereits vorgestellt. An dieser Stelle soll kurz die graphische Oberfläche von Okusys vorgestellt werden. In Abbildung 5.5 ist als Beispiel ein bereits gelerntes Objekt („shelf01“) durch das Okusys-System als „Regal“ analysiert worden.

Das Bild auf der linken Seite des Fensters ist das Originalbild, das analysiert werden soll. Die Features sind dabei farbig gekennzeichnet. Auf der rechten Seite ist das gefilterte Bild dargestellt. Hier werden auch die Augenbewegungen durch Pfeile dargestellt.

Die sich darunter anschließende Tabelle beinhaltet folgende Informationen: Die Spalte „Act. Feat“ gibt die aktuelle Anzahl der Fixationspunkte an und an welchen Koordinaten im Bild sich der aktuelle Fixationspunkt befindet. Die dabei fixierbaren Punkte sind diejenigen, die als Features während der Vorverarbeitung extrahiert wurden.

In der Spalte „Proposal of IBIG“ wird darüber informiert, was für eine Sakkade als nächstes ausgeführt werden sollte. In Kapitel 3.3 wurde bereits beschrieben, dass jeweils die Augenbewegung ausgeführt wird, die den größten Informationszuwachs im nächsten Schritt verspricht. Dabei gibt „I“ in der Spalte den jeweiligen Informationszuwachs an. Es werden zusätzliche Information über die Richtung und die Länge der folgenden Sakkade gegeben. Wurden bisher noch keine Objekte gelernt und sind keine Augenbewegungen bisher ausgeführt worden, so dass auch kein Informationszuwachs berechnet werden kann, können in dieser Spalte entsprechend keine Resultate angezeigt werden.

In „New Eye Movement“ werden die ausgeführten Augenbewegung angezeigt, die zum nächsten Fixationspunkt führen. Dabei kann entweder die von IBIG vorgeschlagene Augenbewegung ausgeführt werden („IBIG“), oder es wird eine andere Augenbewegung



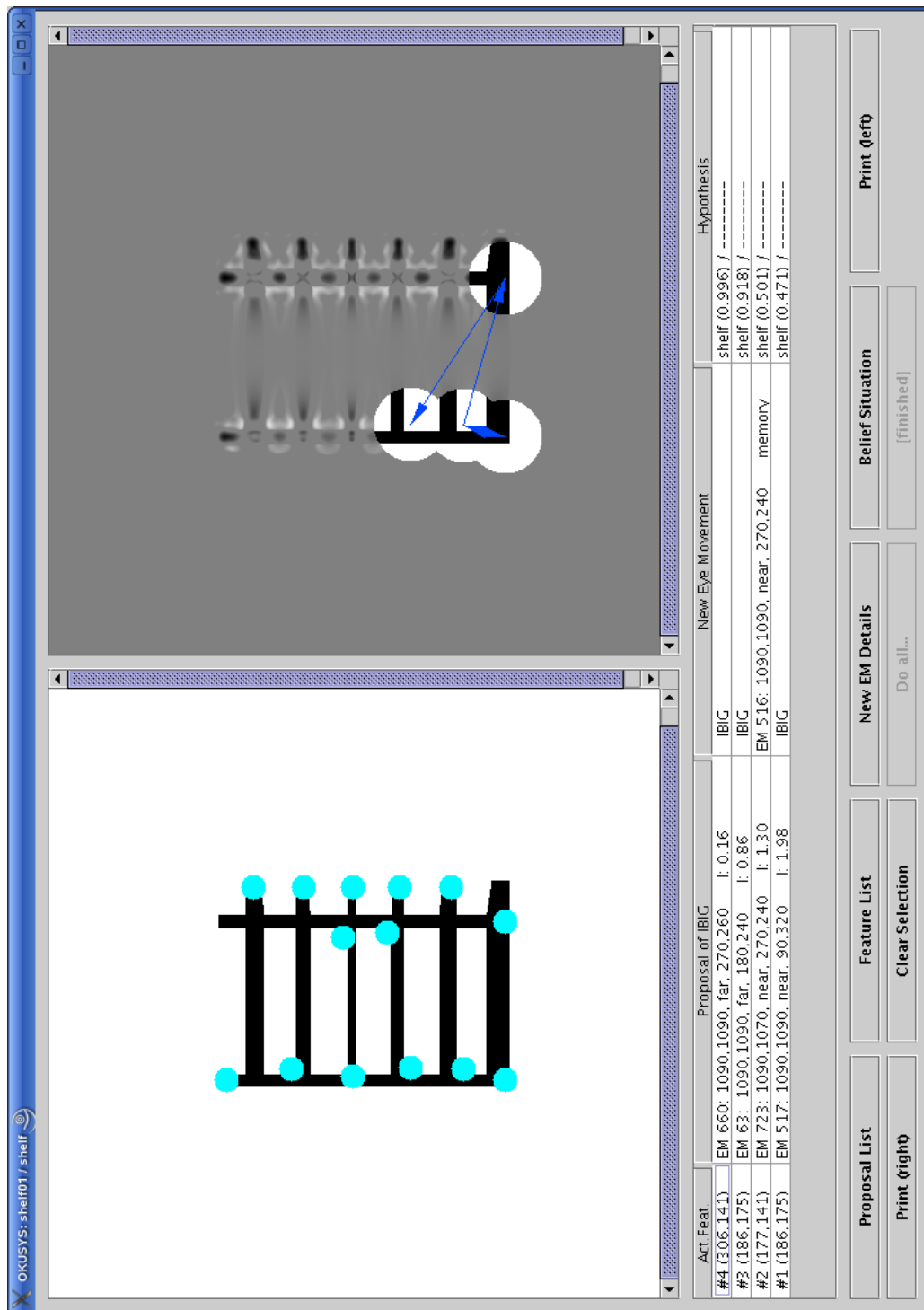


Abbildung 5.5: Objektidentifikation von „shelf01“ durch das Okusys-System

in Abhängigkeit der überhaupt fixierbaren Features ausgeführt. Diese Augenbewegung kann entweder neu („new“) oder schon einmal ausgeführt worden sein („memory“). In der Spalte „Hypothese“ sind die aktuellen Hypothesen mit dem jeweiligen Belief zur Szene angegeben, sofern schon eine Hypothese vorhanden ist. Die Buttons an der unteren Seite des Fenster bieten noch zusätzliche Informationen über den aktuellen Analysevorgang, doch diese sind hier nicht weiter relevant.

In Abbildung 5.5 wurden demnach vier Augenbewegungen ausgeführt, die schließlich zu dem Ergebnis führten, dass es sich bei dem Objekt um ein Regal zu einem Belief von 0.996 handelt. Wurden nun die einzelnen Objekte vom System gelernt, kann ein Raumausschnitt analysiert werden. Dieser Vorgang wird im folgenden Abschnitt beschrieben.

### 5.7.2 Raumanalyse

Das in Kapitel 2.2.2 beschriebene Konzept des Bildverstehens findet sich insgesamt auch in der vorliegenden Arbeit wieder: Die Prozesskette des Systems besteht aus der Vorverarbeitung und Extraktion von signifikanten Eigenschaften, dem zusätzlichen Wissen, an welchen Stellen die einzelnen Objekte, die anhand von Sakkadenzügen gelernt werden können, zu finden sind und schließlich können anhand der Ontologie mögliche Raumkonzepte erschlossen werden.

Die hier umgesetzte Erweiterung des bestehenden Okusys-Systems bietet die Möglichkeit, Bilder zu analysieren, indem die einzelnen Objekte des Bildes analysiert werden und mittels Informationen aus einer Ontologie Rückschlüsse auf das gesamte Bild gezogen werden können. Dafür werden Schritt für Schritt die einzelnen Objekte aus dem Bild segmentiert, analysiert und eine Anfrage auf der Basis dieser Analyse an die Ontologie geleitet. Im Folgenden soll dieser Vorgang beispielhaft an einem Raumausschnitt eines Büros („office05“) vorgestellt werden.

Die Oberfläche des bestehenden Okusys-Systems wurde für die erweiterten Funktionen geändert: Wie in Abbildung 5.6 dargestellt, befindet sich das zu analysierende Bild verkleinert auf der linken, oberen Seite des Fensters. Hier werden im Folgenden auch die sogenannten *Intersakkaden* abgebildet (siehe Abbildung 5.7), die zwischen zwei Objekten ausgeführt werden. Dabei handelt es sich bei dem Begriff Intersakkaden nicht um eine offizielle Bezeichnung. Es sollen in der vorliegenden Arbeit lediglich die Intersakkaden von den sogenannten *Intrasakkaden* unterschieden werden, die durch das bestehende Okusys-System auf einzelnen Objekten ausgeführt werden. Im rechten Bereich des Fensters befindet sich die Analyse der einzelnen Objekte. Die beiden Teilbereiche des Fensters, in denen das Objekt mit seinen Features bzw. das gefilterte Bild mit den Augenbewegungen angezeigt werden, sowie die zugehörige Tabelle, in

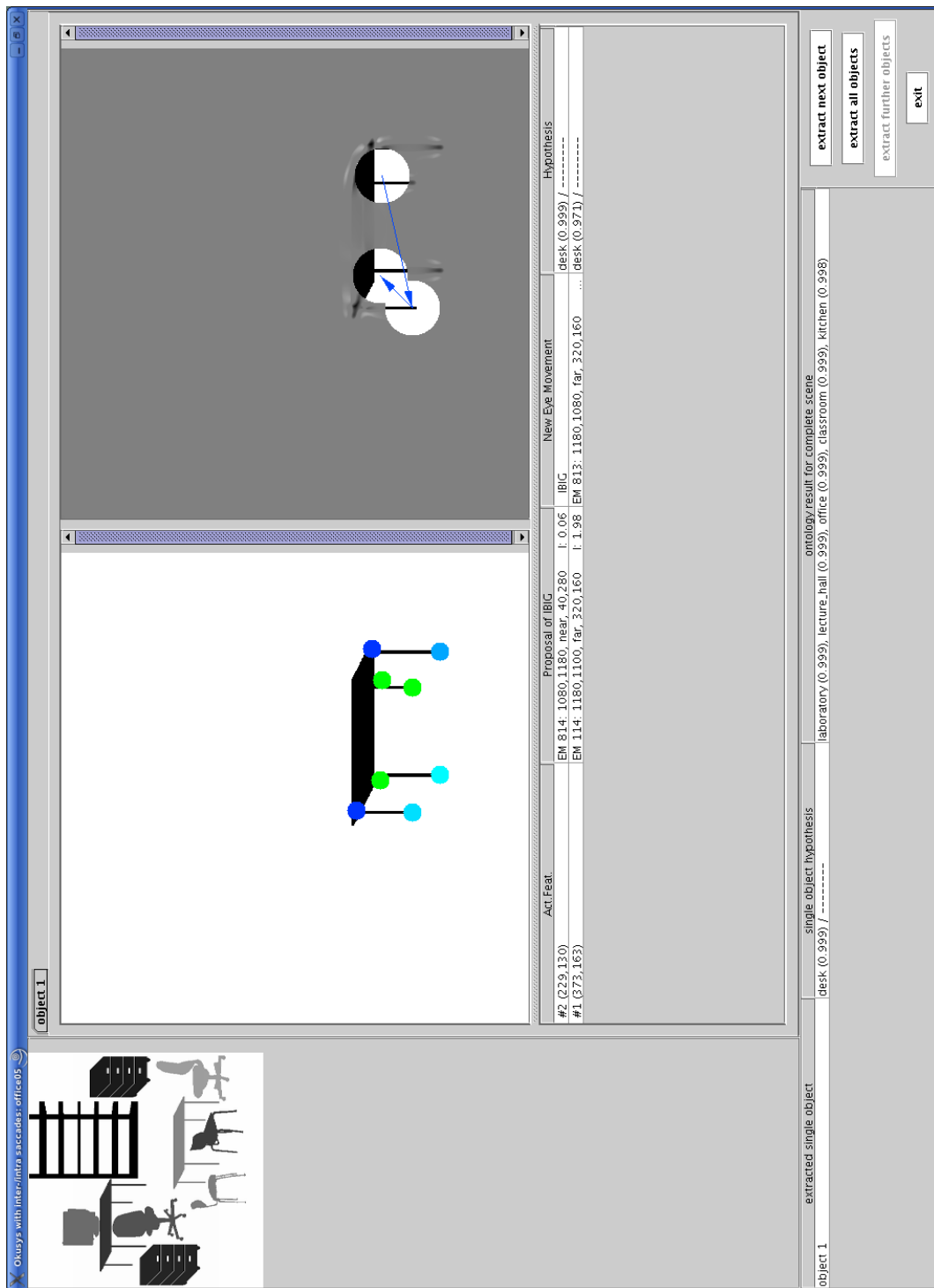


Abbildung 5.6: Raumanalyse von „office05“ - erstes Objekt

der Informationen über die Augenbewegungen dargestellt werden, aus dem im vorigen Abschnitt beschriebenen Okusys-System sind in der rechten Hälfte des Fensters angeordnet.

Am unteren Ende des Fensters schließt sich eine Tabelle an, in der die Anfragen an die Ontologie protokolliert werden. In der Spalte „extracted single object“ werden die einzelnen, analysierten Objekte aus dem Bild nummeriert. Die Nummerierung korrespondiert mit den gleichnamigen Reitern der Analyse im darüberliegenden, rechten Fenster. In der Spalte „single object hypothesis“ werden die Ergebnisse, also die Hypothesen über die jeweilige Szene mit dem zugehörigen Belief der Analyse der einzelnen Objekte notiert. Durch die Spalte „ontology result for complete scene“ wird darüber informiert, welche möglichen Raumkonzepte aufgrund der bisherigen Analyse möglich sind, und wie hoch ihre Beliefs sind.

Im unteren, rechten Teil des Fensters schließlich befinden sich Buttons, die die Analyse des nächsten einzelnen Objekts starten. Insofern kann ein Benutzer entweder nur ein Objekt zur Zeit analysieren („extract next object“) oder er kann die verbleibenden Objekte analysieren („extract all objects“). Einer zusätzlichen Erklärung bedarf der Button „extract further objects“.

Zu Beginn der Analyse eines Raumausschnitts werden auf diesem Bild gemäß dem in Kapitel 3.3 beschriebenen Vorgang vorverarbeitende Schritte ausgeführt. Dies dient dazu, um Fixationspunkte der Intersakkaden zu finden. Anhand dieser Fixationspunkte erfolgt ebenso wie bei der Analyse einzelner Objekte eine Sakkadenfolge. Diese Augenbewegungen, die auf dem kompletten Raumausschnitt getätigt werden, repräsentieren die Intersakkaden. An jedem Fixationspunkt dieser Intersakkade wird der Raumausschnitt an den aktuellen Koordinaten untersucht: Befindet sich dieser Punkt auf einem Objekt, ist an dieser Stelle im Bild also ein Grauwert vorhanden, so wird dieses Objekt anhand seines Grauwerts aus dem gesamten Bild extrahiert, analysiert und, sofern möglich, anhand seiner Intrasakkade identifiziert. Ist ein Objekt mit diesem Grauwert bereits analysiert worden, so findet keine weitere Analyse statt und es wird der nächste Fixationspunkt der Intersakkade untersucht.

Auf diese Weise werden nach und nach die einzelnen Objekte analysiert. Nach jeder Analyse wird ein Ontologieaufruf nach dem obigen Schema (siehe Kapitel 5.6.2) ausgeführt und der Belief berechnet (siehe Kapitel 5.6.3). Da es vorkommen kann, dass die Fixationspunkte der Intersakkade nicht alle einzelnen Objekte erreichen, werden die bis zu diesem Zeitpunkt nicht analysierten Bilder trotzdem als nächste Analyseschritte angeboten. Hierfür dient der Button „extract further objects“. Er ermöglicht die Analyse der übrigen Objekte, bis schließlich alle einzelnen Objekte des Raumausschnitts analysiert sind.

Bei der Abbildung 5.6 wird als erstes ein Objekt extrahiert, das als Tisch analysiert

Okusys with inter-intra saccades: office05

object 1    object 2

Act. Feat.	Proposal of IBIG	New Eye Movement	Hypothesis
#6 (138,209)	EM 266: 1180,1070, near, 125,200    f: 0.04	IBIG	chair_on_castors (1.000) / -----
#5 (161,316)	EM 260: 1120,1180, far, 160,240    f: 1.34	IBIG	chair_on_castors (0.978) / -----
#4 (157,172)	no result	EM 1075:1100,1120, far, 160,240	CRT_monitor (0.688) / -----
#3 (138,209)	EM 515: 1180,1100, near, 40,0    f: 1.31	IBIG	CRT_monitor (0.688) / -----
#2 (157,172)	EM 52: 1100,1180, near, 320,220    f: 0.09	IBIG	chair_on_castors (0.953) / -----
#1 (191,191)	EM 267: 1070,1100, near, 275,200    f: 1.97	IBIG	chair_on_castors (0.986) / -----

extracted single object

single object hypothesis

chair\_on\_castors (1.000) / -----  
 desk (0.999) / -----

ontology/result for complete scene

laboratory (0.999), lecture\_hall (0.025), office (0.999), classroom (0.546)  
 laboratory (0.999), lecture\_hall (0.999), office (0.999), classroom (0.999), kitchen (0.998)

object 2

extract next object  
 extract all objects  
 extract further objects

exit

Abbildung 5.7: Raumanalyse von „office05“ - zweites Objekt

wird. Wie am Resultat der Ontologieanfrage zu sehen ist, deutet dieses Objekt zu einem sehr hohen Belief auf alle fünf möglichen Raumkonzepte hin. Das nächste extrahierte Objekt wird als Bürostuhl mit Rollen analysiert (siehe Abbildung 5.7). Diese Analyse schränkt die Ergebnisse dafür, um welches Raumkonzept es sich handeln könnte, ein: Die Ontologieanfrage liefert noch sehr hohe Werte für die Raumkonzepte Praktikumsraum und Büro, das Raumkonzept Seminarraum liegt nur noch bei einem Belief von 0.546, der Hörsaal wird mit einem Belief von 0.025 kaum noch unterstützt und das Raumkonzept der Küche ist nicht mehr aufgeführt, kommt also nicht mehr in Frage. Als drittes wird ein weiteres Objekt als Bürostuhl mit Rollen analysiert. An den Ergebnissen ändert sich somit nichts. Bei dem vierten Fixationspunkt der Intersakkade wird ein Objekt extrahiert, das als Stuhl aus Holz/Metall analysiert wird (siehe Abbildung 5.8). Die nun folgende Ontologieanfrage bezüglich möglicher Raumkonzepte für Raumausschnitte mit den Objekten Tisch, Bürostuhl mit Rollen und Stuhl aus Holz/Metall liefert den Belief von 0.99 für das Konzept Büro und den Belief 0.297 für das Konzept Praktikumsraum.

Im darauffolgenden Schritt wird wieder ein Stuhl aus Holz/Metall analysiert, das Ontologieanfrageergebnis bleibt unverändert. Im sechsten Analyseschritt wird ein Objekt extrahiert, das als Regal klassifiziert wird (siehe Abbildung 5.9). Hierauf ändern sich die Beliefs der Raumkonzepte auf jeweils 0.956 und 0.24 für Büro bzw. Praktikumsraum. Die folgenden Analyseschritte extrahieren jeweils Objekte vom Typ CRT\_Monitor, Tisch und zwei Mal Rollcontainer. Das Ergebnis der Ontologieanfrage bleibt bis zum Ende der Analyse unverändert, wie in Abbildung 5.10, in der der letzte Analyseschritt abgebildet ist, zu sehen ist.

Dabei gibt es keinen Mechanismus, der die Raumanalyse abbricht. Man kann das Bild so lange analysieren, wie es noch nicht analysierte einzelne Objekte im Bild gibt. Der Benutzer entscheidet also, wann eine Analyse abgebrochen werden soll. Eine Auswertung der eben beschriebenen Analyse folgt in Kapitel 7.1. Dort werden auch weitere Testfälle dargestellt und evaluiert. Vorher wird im nächsten Kapitel die Implementierung dokumentiert.

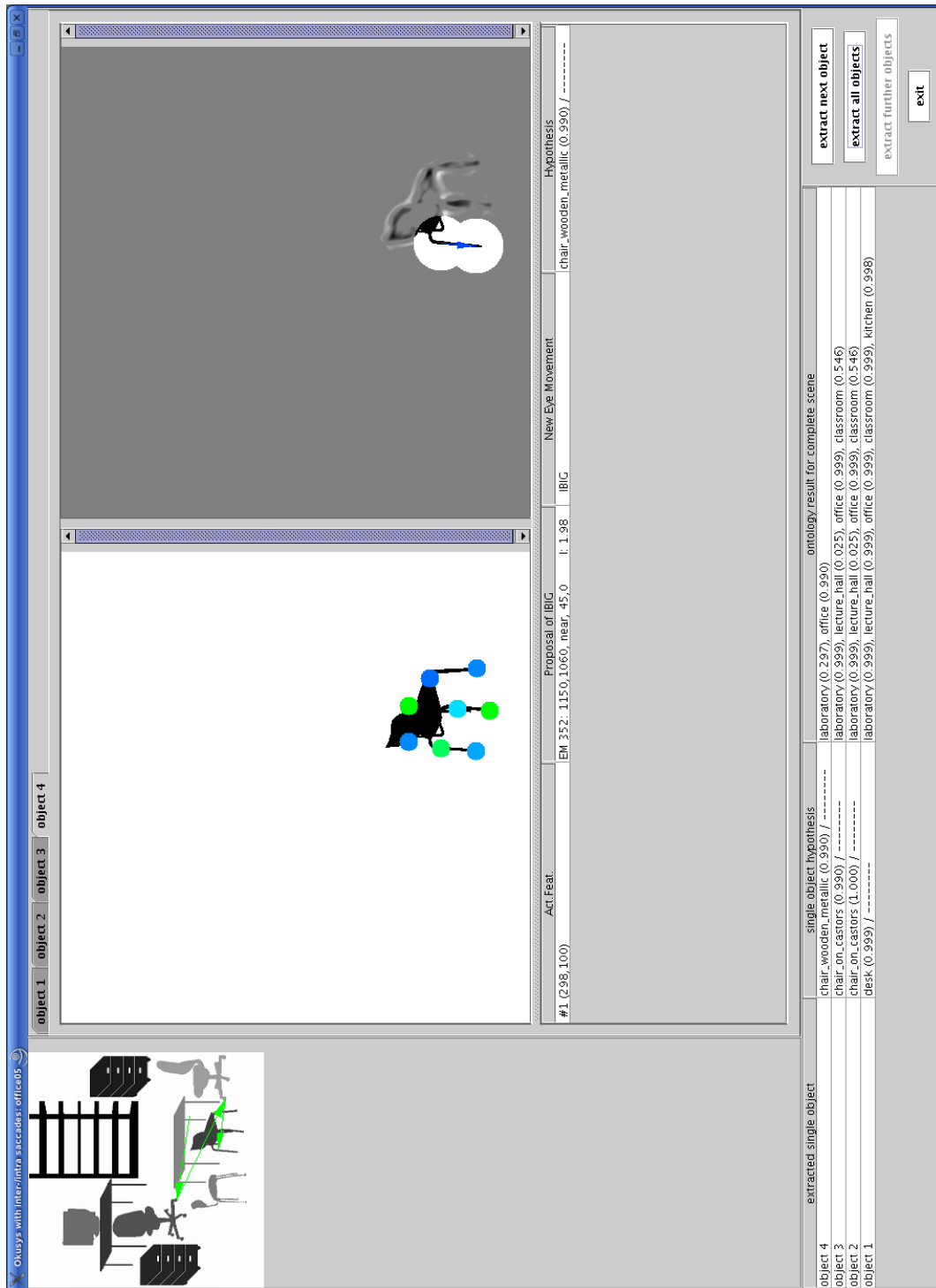


Abbildung 5.8: Raumanalyse von „office05“ - viertes Objekt

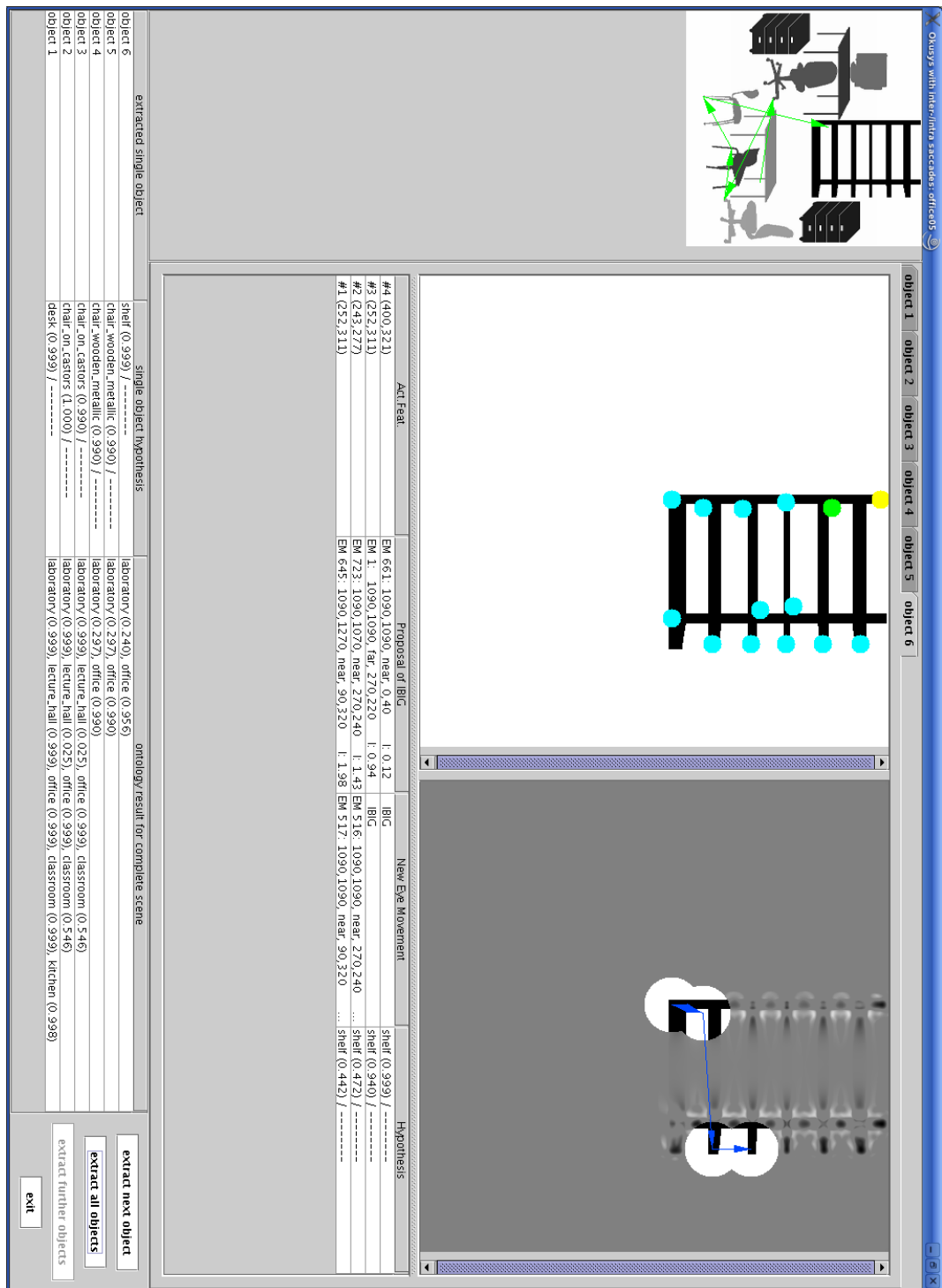


Abbildung 5.9: Raumanalyse von „office05“ - sechstes Objekt



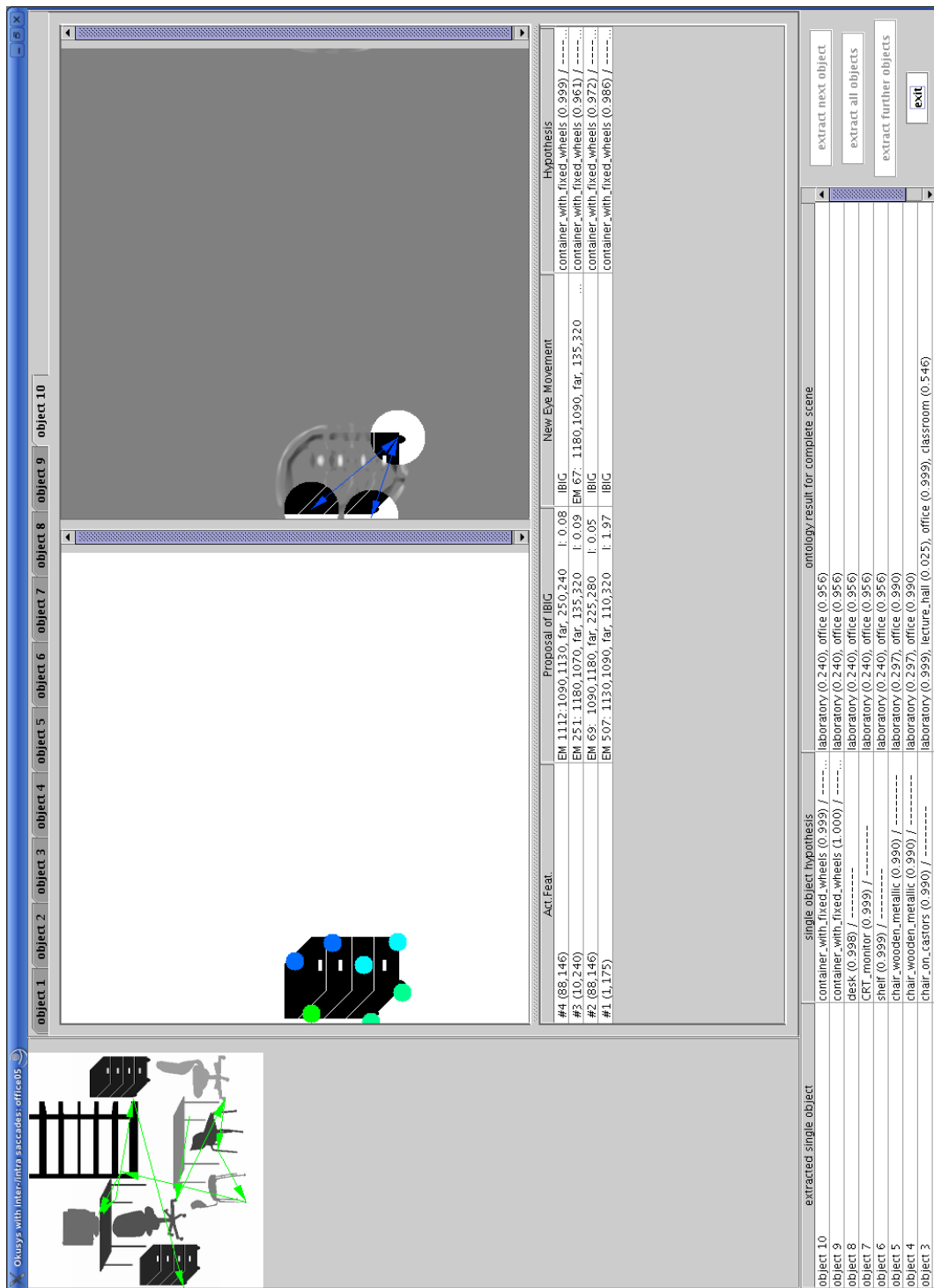


Abbildung 5.10: Raumanalyse von „office05“ - zehntes (letztes) Objekt



## 6 Implementierung

In diesem Kapitel soll die Implementierung der im vorherigen Kapitel beschriebenen Funktionen dokumentiert werden. Die Umsetzung erfolgte mit Java, da der bereits existierende Teil der graphischen Oberfläche von Okusys in Java umgesetzt wurde und die benutzten Bibliotheken JIU und Pellet (siehe unten) ebenfalls in Java vorliegen.

### 6.1 Klassenstruktur

Da das bestehende System keine Package-Struktur besitzt, wird eine neue eingeführt, die aus den Packages `okusys.control`, `okusys.model`, `okusys.view` und `okusys.util` besteht. Orientiert am MVC-Pattern ergibt sich folgende Struktur:

- `control`: Klassen, die den Kontrollfluß steuern
- `model`: Klassen, die für die Datenhaltung zuständig sind
- `view`: Klassen für die graphische Ausgabe
- `util`: Klassen für sonstige Operationen

Um das bestehende System weitestgehend unangetastet zu lassen, wurden die zugehörigen Klassen im Package `okusys` abgelegt. In der dortigen Klasse `DisplayServer` wurde die zusätzliche Funktionsmöglichkeit der Analyse von Inter- und Intrasakkaden eingefügt. Wird der Menüpunkt für diese Analyse im Okusys-Auswahlfenster (siehe Abbildung 6.1) betätigt, werden die aktuellen Okusys-Fenster ausgeblendet und der Vorgang gestartet:

```
if (what.equals(Constants.INTER_INTRA_SACCADES_MENU_ITEM_NAME)) {  
    if (parameterOk()) {  
        mainWin.setNullMenu();  
        sceneSelectionDialog.hideTmp();  
        hypothesisSelectionDialog.hideTmp();  
        controlThreadWithInterIntraSaccades =
```

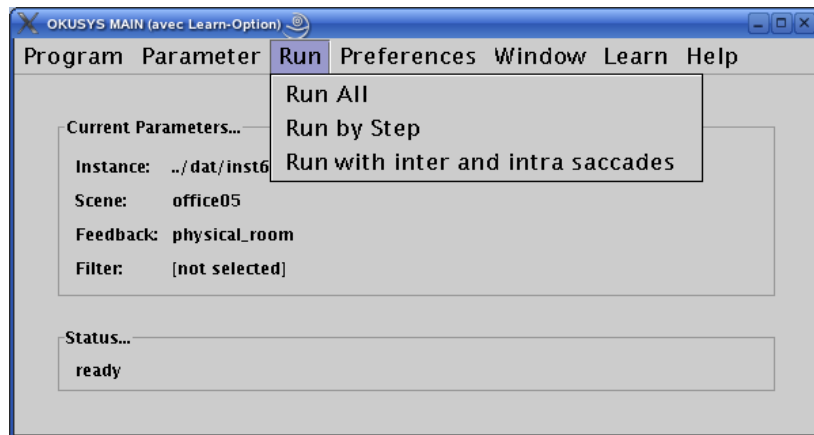


Abbildung 6.1: Okusys-Auswahlfenster

```

new ControlThreadWithInterIntraSaccades(this);
threadRunning=true;
controlThreadWithInterIntraSaccades.start();
}
else JOptionPane.showMessageDialog(mainWin, "Parameter Error",
    "Error...", JOptionPane.INFORMATION_MESSAGE);
}

```

Hier wird eine neue Instanz der Klasse `ControlThreadWithInterIntraSaccades` erzeugt und aufgerufen. Diese Klasse liegt im Package `okusys.control`. Sie kontrolliert den Ablauf der Analyse durch Inter- und Intrasakkaden.

```

public ControlThreadWithInterIntraSaccades(DisplayServer displayServer) {
//...
initializeScene(imageName);
initializeCompleteFrame();
initializeFeaturesOfCompleteScene();
createExtractedImagesList();
//...
}

```

Bei der Initialisierung dieser Klasse wird das komplette Eingabebild einmalig analysiert und ein Sakkadenzug, der die Intersakkaden darstellt, ausgeführt (`initializeScene(imageName);`). Die graphische Oberfläche wird erzeugt

(`initializeCompleteFrame()`;) mit dem kompletten Eingabebild, der Anzeige für die einzelnen, zu extrahierenden Objekte, der Tabelle für die Ergebnisse der Ontologieanfrage und den Buttons für weitere Objektanalysen (siehe auch Abbildung 5.6). Mit `initializeFeaturesOfCompleteScene()`; werden die Features, die bei der initialen Analyse des gesamten Bildes erstellt wurden, ausgelesen, um im nächsten Schritt die Reihenfolge der Extraktion der einzelnen Objekte (`createExtractedImagesList()`;) vorzugeben.

Für die graphische Oberfläche, wie sie auch in Kapitel 5.7.2 erklärt ist, ist die Klasse `ImageExtractionWindow` im Package `okusys.view` zuständig. Hier werden die entsprechenden graphischen Elemente festgelegt. Das im linken, oberen Bereich der Oberfläche angebrachte Ausgangsbild wird in der Klasse `JImagePanel` behandelt. Diese Klasse existierte bereits im vorhandenen System, wurde aber in das Package `okusys.view` verschoben und modifiziert, um die Augenbewegungen auf dem kompletten Bild anzeigen zu können und weitere intern benötigte Methoden zur Verfügung zu stellen. Für die Darstellung der Analyse der einzelnen Objekte im oberen, rechten Fenster der Oberfläche dient die Klasse `InternalSaccadeWindow`. Ihre Funktionen sind zum Teil von der bestehenden Okusys-Klasse `SakkadenWindow` übernommen worden, gemeinsame Elemente wurden in das neu erstellte Interface `I_SakkadenWindow`, das ebenfalls im Package `okusys.view` befindet, zusammengefasst. Dies führt dazu, dass die bestehende Klasse `FixComponent`, die für die Darstellung des gefilterten Bildes und der Sakkaden zuständig ist, nun nicht mehr direkt mit der Klasse `SakkadenWindow`, sondern mit der Schnittstelle kommuniziert.

Die Klasse `ImageExtractionWindow` benutzt zudem für die Definition der Buttonfunktionen die Klasse `ImageExtractionActionListener` aus dem Package `okusys.model`. Hier wird festgelegt, welche Operationen bei Betätigung der jeweiligen Buttons ausgeführt werden sollen. Für die Datenhaltung der sich am unteren Rand der Oberfläche befindenden Tabelle, in der die Ontologieanfragen protokolliert werden, ist die Klasse `CompleteSceneTableModel` im Package `okusys.model` zuständig. Dabei werden neue Reihen am oberen Tabellenende angefügt, um ein gleiches Verhalten zu zeigen wie bei der Tabelle der Sakkadenzüge des bestehenden Okusys-Systems.

Für die Extraktion der einzelnen Objekte im Bild werden, wie oben beschrieben, die einzelnen Fixationspunkte der Intersakkade verfolgt. Wird an einem Punkt ein Grauwert analysiert, der bisher nicht extrahiert wurde, so erfolgt eine Extraktion anhand dieses Grauwerts. Für die Untersuchung des Eingabebilds und die Extraktion ist die Klasse `ImageUtilities` des Packages `okusys.util` zuständig. Diese Klasse bietet zusätzlich die Methode `public static Image halveImageSize(String imagePath)` zur Verkleinerung des gesamten Eingabebilds bei der Darstellung im linken, oberen Fensterbereich (siehe Abb. 5.6) und die Methode `private static void convertPpmToByte`

(final String ppmFilename, String singleObjectImage) zur Umwandlung der extrahierten Einzelbilder in das für die Analyse mit Okusys notwendige Byte-Format. Für das Bearbeiten des Eingabebilds, bei dem das jeweilige einzelne Objekt im Bild anhand der Pixel des zugehörigen Grauwerts extrahiert wird, und für weitere graphische Bearbeitungen, wie das Verkleinern eines Bilds und das Erstellen von GIF-Dateien, wird die Open Source-Bibliothek *Java Images Utilities (JIU)*<sup>1</sup> in der Version 0.12.0 verwendet. Für die Umwandlung in Byte-Dateien wird die entsprechende Funktion ppm2byte des Systems *IPRS, Image Processing and Pattern Recognition Systems* (Caeli u. a., 1997), in der Version 2.0.0 verwendet, von dem das bestehende Okusys-System bereits Funktionen nutzt.

Wurden die einzelnen Objekte extrahiert und für die spätere Analyse gespeichert, wird in der Klasse `ControlThreadWithInterIntraSaccades` die entsprechende Routine für die Analyse einzelner Objekte aufgerufen. Die für die Analyse einzelner Bilder notwendigen Methoden sind in der Klasse `ConnectionToCprogram` aus dem Package `okusys.control` enthalten, in dem der C-Programmteil von Okusys aufgerufen wird. Teile dieser Klasse sind aus der ursprünglichen Okusys-Analyse übernommen worden. Hier werden die einzelnen Routinen für die Sakkadenausführung der extrahierten Objekte aufgerufen. Das Hypothesenergebnis der Analyse wird in der Klasse `ControlThreadWithInterIntraSaccades` weiter verarbeitet. Von hier aus wird die Klasse `OntologyQuery` im Package `okusys.util` aufgerufen, in der die Ontologieanfragen auf Basis der extrahierten Objekte, wie in Kapitel 5.6.2 beschrieben, ausgeführt werden.

`OntologyQuery` nutzt die Bibliothek von Pellet in der Version 1.3-beta, in der die Jena-Bibliothek eingebettet ist (siehe hierzu auch Kapitel 5.6.2). Die nach jeder Analyse einzelner Objekte identifizierten Objekte werden hier gespeichert und bei jeder Anfrage an die Ontologie berücksichtigt:

```
String queryBody = "WHERE ";
// hypothesisList stores the identified single objects
for (int index = 0; index < hypothesisList.size(); index++) {
    if (isValidExpression((String)hypothesisList.get(index))) {
        final String propertyType = getProperty((String)hypothesisList.get(index));
        // for individual variable name every cycle
        final String instanceName = propertyType + i;
// base = URI of ontology
queryBody += "    (?x, <" + base + "contains>, ?" + instanceName + ") " +
```

---

<sup>1</sup><http://sourceforge.net/projects/jiu/>

```

        "    (?\" + instanceName + "\", rdf:type, <\" + base + propertyType + ">");
    }
}
if (queryBody.equals("WHERE ")){
    return null;
}
String queryStr = "SELECT ?x " + queryBody;
Query query = new Query(queryStr);

```

Hier wird die Anfrage in RDQL-Notation gestellt, die alle Instanzen, die in der Relation „contains“ jeweils alle analysierten, einzelnen Objekte besitzen, zurückliefert. Die so ermittelten Instanzen werden bezüglich ihrer Unterstützung bestimmter Raumkonzepte untersucht. So ergeben sich die Raumkonzepte und ihre Glaubensmaße, welche nur noch bezüglich ihres *confirming Beliefs* kombiniert werden müssen (siehe hierzu auch Kapitel 5.6.3). Diese Berechnung erfolgt in der folgenden Methode:

```

public float getCombinedEvidences(final String newEvidence,
Float currentBeliefValue){
    return currentBeliefValue.floatValue()*new Float(newEvidence).floatValue()
        + (1-currentBeliefValue.floatValue()*new Float(newEvidence).floatValue()
        + currentBeliefValue.floatValue()*(1-new Float(newEvidence).floatValue()));
}

```

Hier taucht auch die bereits beschriebene Problematik auf, dass die interne Repräsentation von Pellet für „owl:Datatypes“ durch Strings dargestellt ist, die eine entsprechende Umwandlung in Gleitkommazahlen notwendig macht. Die Repräsentation von Dezimalzahlen in OWL unterstützt der Reasoner nicht. In Abbildung 6.2 sind zusammenfassend die hier beschriebenen Klassen in einem UML-Klassendiagramm dargestellt.

**Voraussetzungen** Notwendig für einen ordnungsgemäßen Ablauf des hier beschriebenen Programmabschnitts ist das Vorhandensein des Eingabebilds im Byte- und GIF-Format. Letzteres ist notwendig, da die Umwandlung von einer Byte-datei in eine GIF-Datei durch IRPS verlustbehaftet ist. So entstehen Störpixel in der GIF-Datei, die dazu führen, dass die Extraktion der Objekte anhand ihrer Grauwerte nicht mehr funktioniert: So können vereinzelt Pixel eines Grauwertes auftreten, die kein Objekt repräsentieren. Da diese Grauwerte dennoch extrahiert und für eine Analyse an das Okusys-System weitergereicht werden, kann dies zu einem Absturz des Systems führen,

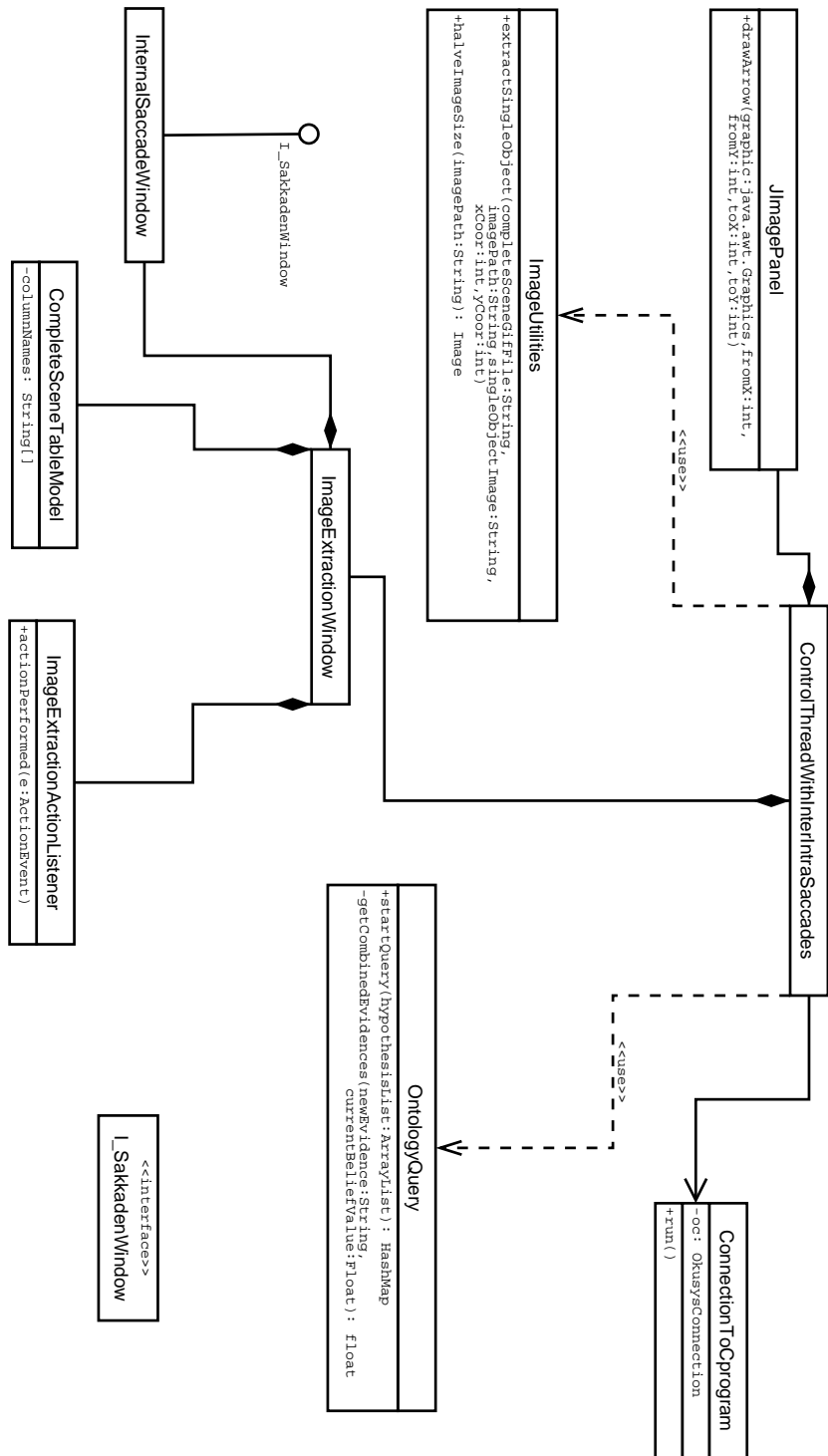


Abbildung 6.2: Klassendiagramm



da keine signifikanten Merkmale auf Basis dieser verinselten Pixel im Bild gefunden werden können. Des Weiteren wird vorausgesetzt, dass der Hintergrund des Bildes weiß ist. Auch ist nur eine Raumanalyse zur Zeit möglich.

## 6.2 Unittests

In den Test-Packages `test.model` und `test.util` liegen die zu den Klassen der entsprechenden Pakete gehörigen JUnit-Testklassen<sup>2</sup>. Im Einzelnen befinden sich hier die Klassen `TestCompleteSceneTableModel`, `TestImageUtilities` und `TestOntologyQuery`. In der Klasse `TestCompleteSceneTableModel` wird das korrekte Verhalten beim Speichern und Auslesen der Daten in die bzw. aus der Tabelle getestet. `TestImageUtilities` prüft die Korrektheit der Extraktion einzelner Objekte aus Eingabebildern und `TestOntologyQuery` prüft die Ergebnisse von Ontologieanfragen bezüglich ihrer Korrektheit. Die Klassen des `okusys.control`- und `okusys.view`-Packages werden dabei nicht getestet. Das Verhalten der Klassen des `okusys-Control`-Packages sind durch die Testfälle in Kapitel 7.1 abgedeckt, graphisches Verhalten wird nicht weiter überprüft.

## 6.3 Okusys-spezifische Daten

Für die hier zu untersuchenden Graphiken muss der Ordner „physicalroom“ im Verzeichnis „home/<user>/okusys/scene\_base“ angelegt werden, in dem die in Anhang E abgebildeten Instanzen, die von Okusys untersucht werden sollen, und Graphiken der einzelnen Objekte, die in den Rauminstanzen auftreten und von Okusys gelernt werden sollen, abgelegt werden. Um auf diese Daten von Okusys aus zugreifen zu können, werden entsprechende Verzeichnisstrukturen und Dateien nach Vorgabe im Okusys-Ordner bereit gestellt. Die konkreten Daten sind auf der beigefügten CD (siehe Anhang D) im Verzeichnis „inst6“ zu finden. Geringfügige Modifikationen wurden in dem in C geschriebenen Programmteil vorgenommen: Der Aufruf der Javasourcen erfolgt nun mit erhöhtem Speicher für den Stack und Heap, die Anzahl der analysierbaren Graphikdateien wurde erhöht.

---

<sup>2</sup>JUnit: Framework für automatisiertes Testen, <http://www.junit.org/>



## 7 Evaluation der Okusys-Erweiterung

In Kapitel 5.7.2 wurde bereits ein Testfall beispielhaft vorgestellt, um das Verhalten der Systemerweiterung zu demonstrieren. Die Bewertung der Analyseergebnisse ist aber nicht unproblematisch, da die Ergebnisse nicht klar in korrekt/inkorrekt zu gliedern sind. Da die Ergebnisse mit einem Belief gekennzeichnet sind und insbesondere die Expertenbefragungen (siehe Kapitel 5.5.3) auch teilweise keine eindeutige Klassifizierung der Raumkonzepte aufweisen, sind entsprechend die Ergebnisse auch nicht klar in richtig und falsch zu trennen. Dennoch lassen sich die Ergebnisse bewerten. Insbesondere kann überprüft werden, mit welchen Beliefs diejenigen Raumkonzepte gekennzeichnet sind, die bei der ursprünglichen Modellierung den einzelnen Raumausschnitten (siehe Kapitel 5.5.2) zugeordnet wurden. Auch das Verhalten des Okusys-Systems bei der Analyse kann untersucht werden.

Abschließend soll in diesem Kapitel die Modellierung der Ontologie, der Rauminstanzen und die Erfassung der Glaubenswerte bewertet werden.

### 7.1 Ergebnisse aus Testfällen und deren Bewertung

Die folgenden Testfälle repräsentieren beispielhaft Analyseszenarien des Systems, welches Antwortverhalten typischerweise auftritt und welche Erkenntnisse daraus gewonnen werden können. Zuerst wird der Testfall beschrieben, welche Ergebnisse er liefert, wie dieses Ergebnis zustande kommt und wie dies zu bewerten ist. Generelle Eigenschaften des Systems treten dabei in Erscheinung. Grundlage für die Erkennung der einzelnen Objekte, soweit bei den Testfällen nicht anders angegeben, ist eine achtmalige Durchführung der Lernphase der einzelnen Objekte. Dabei wurde jedes Objekt insgesamt acht Mal mit seinem korrekten Feedback vom Okusys-System gelernt.

#### 7.1.1 Testfall: „office05“

Der in Kapitel 5.7.2 dargestellte Testfall der Instanz „office05“ verlief erwartungsgemäß. Nacheinander wurden die einzelnen Objektklassen „desk“, „chair\_on\_castors“, „chair\_on\_castors“, „chair\_wooden\_metallic“, „chair\_wooden\_metallic“, „shelf“, „CRT\_monitor“, „desk“, „container\_with\_fixed\_wheels“ und

„container\_with\_fixed\_wheels“ identifiziert. Die Identifikation der einzelnen Objekte verlief korrekt und die Anfrage an die Ontologie ergab schließlich die Beliefs von 0.956 für „office“ und 0.24 für „laboratory“. Verglichen mit den Glaubensmaßen aus den Expertenbefragungen, die für die Instanz „office05“ bei 0.05 für „laboratory“ und bei 0.75 für „office“ lag, ist das Ergebnis also durchaus als korrekt einzustufen.

Der sehr viel höhere Beliefwert für die Raumkonzepte aus der Okusys-Analyse gegenüber den Werten der Expertenbefragung lässt sich durch die Nutzung von Dempster's Rule of Combination erklären. Weil es unterschiedliche Instanzen gibt, die auf die Raumkonzepte „office“ und „laboratory“ hindeuten, und deren Glaubensmaße, wie in Kapitel 5.6.3 beschrieben, kombiniert werden, ergibt sich der jeweils höhere Beliefwert für die Raumkonzepte.

Dabei ist auffällig, dass bereits nach der Analyse des vierten Objekts eindeutige Hinweise auf die Konzepte „office“ (0.99) und „laboratory“ (0.297) vorliegen - weitere Extraktion und Analyse von Objekten bringen kaum weiteren Informationsgewinn. Die Kombination der Objekte „desk“, „chair\_on\_castors“ und „chair\_wooden\_metallic“ schränken den möglichen Instanzbereich bereits so sehr ein, dass nur noch zwei Konzepte in Frage kommen.

### 7.1.2 Testfall: „laboratory04“ (1)

Bei der Analyse der Instanz „laboratory04“ kommt es zu einer falschen Analyse des ersten, einzelnen Objekts. Wie in Abbildung 7.1 zu sehen ist, wurde das erste Objekt als Instanz eines „chair\_wooden\_metallic“ identifiziert, obwohl es zur Klasse „projection\_surface“ gehört. Diese Fehlanalyse führt nach den beiden im Folgenden korrekt klassifizierten Objekten „sink“ und „projector“ dazu, dass die Anfrage der Ontologie keine Ergebnisse mehr liefert. Physikalische Räume, in denen die Objekte „chair\_wooden\_metallic“, „sink“ und „projector“ auftreten, liefern keine Unterstützung für mögliche Raumkonzepte. Insbesondere existieren in der Ontologie keine Räume, die die drei genannten Objekte überhaupt beinhalten. Unabhängig von der Realität, ob Räume existieren, in denen diese drei Objekte gleichzeitig auftreten, wird hieran deutlich, wie groß der Einfluss einer Fehlinterpretation eines einzelnen Objekts ist. Das System reagiert hier nicht sehr fehlertolerant, obwohl schon durch Integration der Modellierung von unsicherem Wissen eine Fehlertoleranz gegeben sein sollte.

Zur Lösung dieses Problems könnte sich anbieten, den Belief bezüglich der Objektklassifikation zu berücksichtigen. Geringe Beliefs oder häufig wechselnde Hypothesen bei der Analyse eines einzelnen Objekts könnten bei der Ontologieanfrage, insbesondere wenn sie keine Ergebnisse liefert, in Betracht gezogen werden. Doch lässt sich an dem genannten Beispiel absehen, dass die als „chair\_wooden\_metallic“ identifizierte

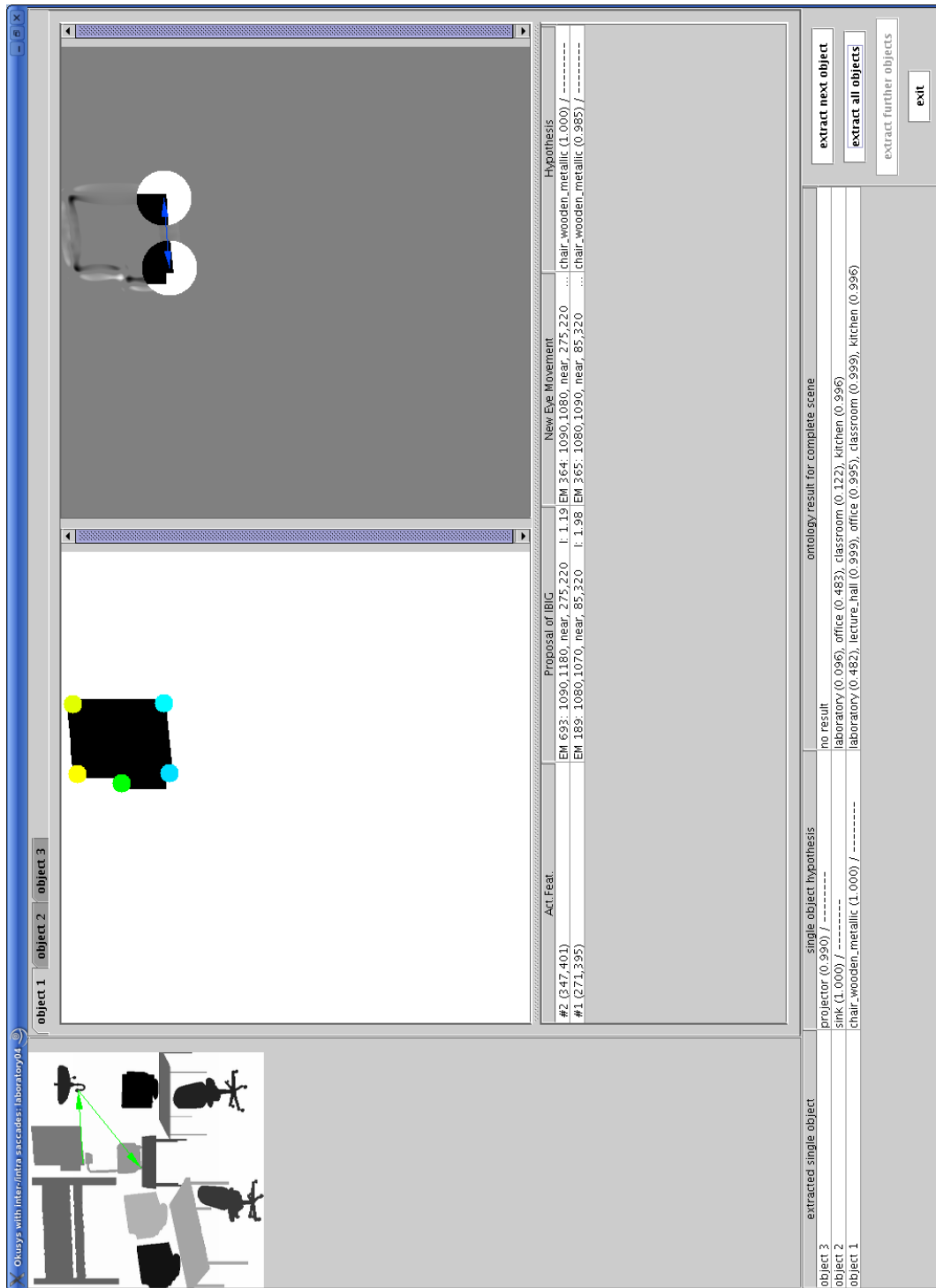


Abbildung 7.1: Analyse der Instanz „laboratory04“ nach drei analysierten Objekten (angezeigt wird das erste analysierte Objekt)

„projection\_surface“ mit einem Belief von 1.0 in nur zwei Sakkadenzügen erkannt wurde. Insofern müssten grundsätzlich alle erkannten Objekte bezüglich ihrer korrekten Klassifizierung in Zweifel gezogen werden.

Eine naheliegende Möglichkeit ist es hingegen, die Objekte erneut vom System lernen zu lassen (siehe nächsten Testfall).

An dieser Stelle sei noch festzuhalten, dass das Ergebnis dieses Testfalls generalisierbar ist, für solche Testfälle, in denen Raumausschnitte zu analysieren sind, in denen tatsächlich „unerwartete“ Objekte auftreten. Für den Fall, dass ein Raumausschnitt eines Praktikumsraums analysiert werden soll, der in der Tat die Objekte „chair\_wooden\_metallic“, „sink“ und „projector“ beinhaltet und diese einzelnen Objekte auch vom System korrekt klassifiziert werden, kann das System keine Ergebnisse liefern.

Hieraus ist auch ersichtlich, welchen großen Einfluss die Modellierung der Instanzen der Ontologie hat. Abhängig davon, was für Raumausschnitte dort modelliert sind, können auch nur solche Ausschnitte wieder erkannt werden. In der hier vorgestellten Modellierung werden allerdings nur die in der Ontologie bereits modellierten Instanzen vom System analysiert. Liefert die Ontologieanfrage keine Ergebnisse, ist folgerbar, dass die Identifikation eines der Objekte fehlgeschlagen sein muss, da alle Instanzen der Ontologie Relationen zu Raumkonzepten aufweisen. Aus dieser Erkenntnis könnte das System entsprechend die Objekte erneut analysieren oder solche Objekte mit nur geringem Belief übergehen. Allerdings sollte das System auch allgemeine Fälle lösen können und nicht nur für die hier modellierten 25 Raumausschnitte korrekte Ergebnisse liefern. Es wäre stattdessen denkbar, neu zu analysierende Raumkonzepte durch überwachtes Lernen mit einem Feedback zu versehen und so neues Wissen, das sich aus der Analyse ergibt, in die Ontologie zu integrieren.

### 7.1.3 Testfall „laboratory04“ (2)

Das falsch erkannte Objekt im vorherigen Testfall wird einer erneuten Lernphase unterzogen, welche vier Mal wiederholt wird. Bei einer erneuten Analyse des Raumausschnitts „laboratory04“ wird dieses Objekt nun korrekt identifiziert. Wie in Abbildung 7.2 wird das Objekt vom Typ „projection\_surface“ als zweites Objekt erkannt. So kann bereits nach der Klassifikation des dritten Objekts ein Belief von 0.8 für Praktikumsraum und von 0.3 für Seminarraum ermittelt werden, der bei den daraufhin überprüften Objekten auch bestehen bleibt.

Ein erneutes Lernen konnte das Problem des vorherigen Testfalls beheben. Die Analyse der Raumkonzepte ist demnach auch abhängig von der Anzahl der Lernphasen einzelner Objekte und der korrekten Klassifizierung. Allerdings sind häufige Lernphasen

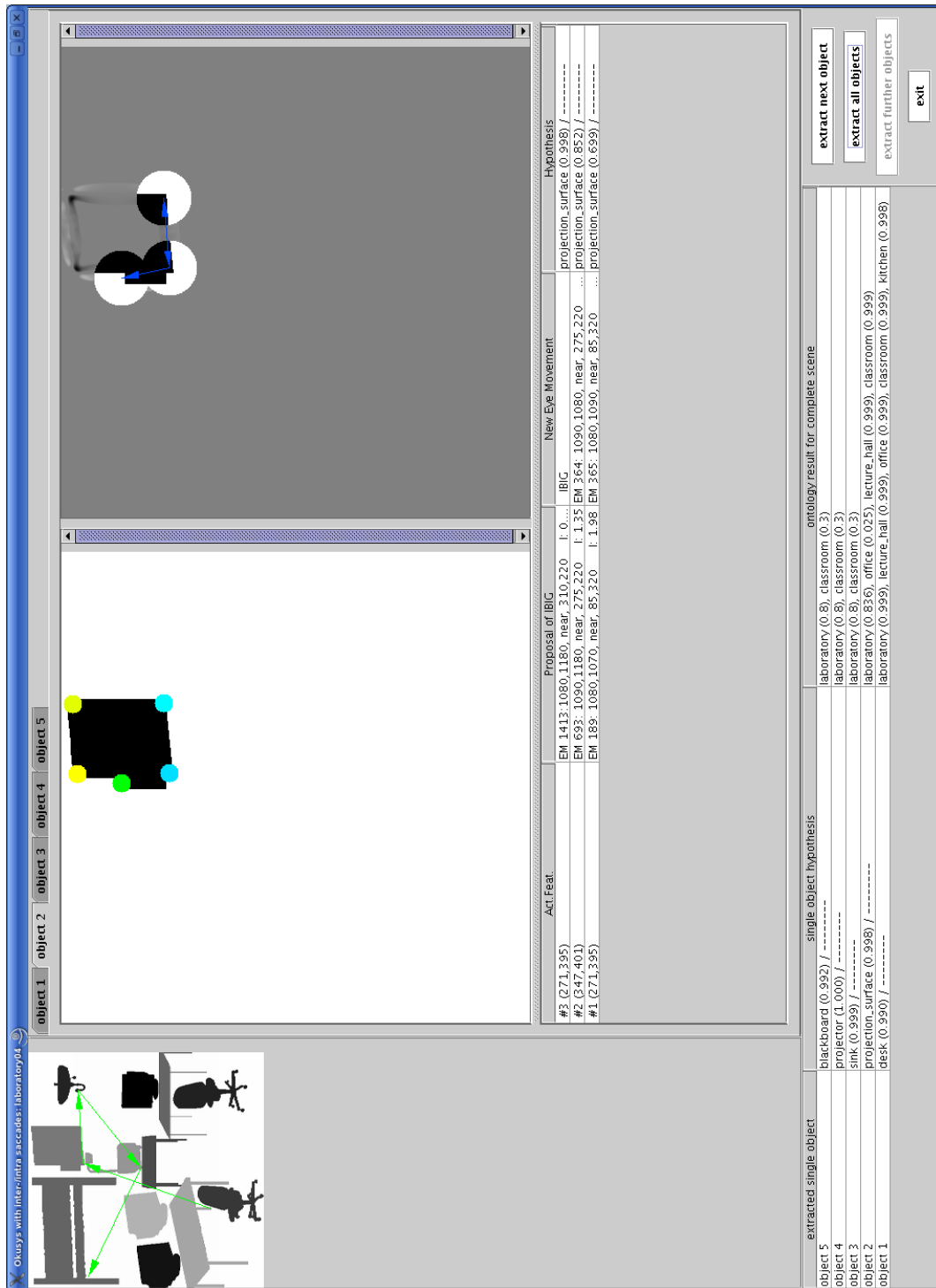


Abbildung 7.2: Analyse der Instanz „laboratory04“ nach fünf analysierten Objekten (angezeigt wird das zweite analysierte Objekt)

bei der praktischen Anwendung des Systems nicht wünschenswert. Andere Lösungen (siehe oben) müssten daher in Betracht gezogen werden.

Erschwerend kommt hinzu, dass die 90 modellierten Instanzen einzelner Objekte der insgesamt 17 Konzepte (siehe Abbildung 5.2) durch die triviale Segmentierung nur auf einen Grauwert beschränkt sind und insofern neben den Umrissen keine weiteren Features bei der Vorverarbeitung berücksichtigt werden können. Eine detailliertere Darstellung der Objekte in Kombination mit einem erweiterten Segmentieralgorithmus könnte insofern auch zu besseren Ergebnissen bei der Objektidentifikation führen. Auch wäre denkbar, eine Erwartungshaltung bei der Analyse der Objekte zu integrieren: Wenn bereits Objekte der Klasse A und B identifiziert wurden, würde bei der Analyse des nächsten Objekts vielleicht eher mit einem Objekt der Klasse D als C zu rechnen sein, obwohl C erkannt wurde. Rückschlüsse auf eine eventuelle Fehlidentifikation aufgrund fehlender Resultate der Ontologieanfrage könnten sich so ergeben.

### 7.1.4 Testfall „lecture\_hall02“

Der in Abbildung 7.3 dargestellte Testfall soll demonstrieren, dass häufig Ergebnisse bereits nach wenigen Analyseschritten einzelner Objekte vorliegen, die sich in der weiteren Analyse nur noch geringfügig ändern. Bereits nach der Identifikation des fünften Objekts „chair\_wooden\_metallic“ werden die Raumkonzepte „lecture\_hall“ mit einem Belief von 0.973 und „classroom“ mit 0.807 unterstützt. Bis zur Analyse des letzten und elften Objekts bleibt dieses Ergebnis unverändert. Das Ergebnis ergibt sich insbesondere daraus, dass die meisten der analysierten Objekte derselben Klasse angehören und somit keinen neuen Informationsgewinn bringen. Ein ähnliches Szenario war auch schon im Testfall „office05“ zu beobachten.

Hierbei liegen die Ergebnisse der unterstützten Raumkonzepte also schon frühzeitig vor, dennoch werden weiter Ressourcen für weitere Analyseschritte einzelner Objekte eingesetzt. Bei dem derzeitigen System entscheidet zwar der Benutzer, wie viele Objekte noch analysiert werden und wann die Analyse beendet werden soll, doch ist bereits in der Einleitung angesprochen worden, dass die hier vorgestellte Analyse Teil eines anderen technischen Systems sein könnte, das die Ergebnisse unabhängig von Benutzern verarbeitet.

Eine Lösung wäre der automatische Abbruch der Analyse, wenn sich nach einer bestimmten Anzahl identifizierter Objekte die Anfrageantworten der Ontologie nicht verändern. Der nächste Testfall zeigt aber, dass auch nach vielen Analyseschritten noch neue Erkenntnisse gewonnen werden können. Zudem kann gerade durch die Kombination mehrerer Objekte ein erfolgreiches Resultat erreicht werden. Eine Idee wäre es, die Intersakkade gezielter auszuführen, als dies bisher der Fall ist. Über diese



The screenshot shows a software interface for analyzing eye-tracking data. The interface is divided into several sections:

- Top Left:** A 3D scene view showing a lecture hall with chairs and desks. Green arrows indicate the gaze path of the user.
- Top Right:** A 2D eye movement visualization showing a black shape representing the eye and a blue arrow indicating the direction of gaze.
- Center:** A table of eye movement data with columns for 'Act. Feat.', 'Proposal of IBIG', and 'Hypothesis'. The data includes coordinates and confidence scores for various eye movements.
- Bottom Right:** A list of extracted single objects and their corresponding ontology results for the complete scene.

Act. Feat.	Proposal of IBIG	Hypothesis
#2 (397,93)	EM 635: 1180,1020, near, 80,0 f: 0,35 IBIG	desk (0,995) /
#2 (454,99)	EM 631: 1020,1180, near, 280,260 f: 1,96 IBIG	desk (0,818) /
#1 (454,61)	EM 579: 1070,1180, near, 45,300 f: 1,98 EM 1023:1070,1020, near, 45,300	..... /

extracted single object	single object hypothesis	ontology result for complete scene
object 11	desk (0,995) /	lecture_hall (0,973), classroom (0,807)
object 10	chair_desk (0,989) /	lecture_hall (0,973), classroom (0,807)
object 9	chair_desk (1,000) /	lecture_hall (0,973), classroom (0,807)
object 8	projector (0,990) /	lecture_hall (0,973), classroom (0,807)
object 7	chair_desk (0,999) /	lecture_hall (0,973), classroom (0,807)
object 6	chair_desk (0,999) /	lecture_hall (0,973), classroom (0,807)
object 5	chair_wooden_metallic (1,000) /	lecture_hall (0,973), classroom (0,807)
object 4	chair_desk (0,990) /	lecture_hall (0,973), classroom (0,959)

Abbildung 7.3: Analyse der Instanz „lecture\_hall02“ nach elf analysierten Objekten

Möglichkeit soll aber in Kapitel 8 noch eingegangen werden.

### 7.1.5 Testfall „classroom03“

Der hier getestete Raumausschnitt verdeutlicht nochmals die Problematik des vorherigen Testfalls. Wie in Abbildung 7.4 zu sehen, verändert sich das Resultat der unterstützten Raumkonzepte während der Analyse von sechs einzelnen Objekten nicht. Erst bei der Analyse des siebten Objekts ergibt sich ein deutlicher Informationsgewinn: Während von der ersten bis zur sechsten Objektidentifikation praktisch alle fünf Raumkonzepte für den aktuellen Raumausschnitt noch in Frage kommen, schränken sich erst nach der Klassifizierung des siebten Objekts die Möglichkeiten ein: „lecture\_hall“ und „classroom“ werden jeweils noch zu einem Belief von 0.999 unterstützt, während „laboratory“ nur noch einen Belief von 0.184 und „office“ 0.025 aufweisen und das Konzept „kitchen“ nicht weiter unterstützt wird.

Bei diesem Test ergibt sich die Problematik, dass über eine Reihe von Objektidentifikationen keine Hinweise auf mögliche Raumkonzepte vorliegen. Diese Beobachtung kann generell für solche Testfälle stehen, bei denen der zu analysierende Raumausschnitt nur „uninteressante“ Objekte, also solche, die wenig Informationsgewinn für Ontologieanfragen bieten, beinhaltet. Befinden sich in dem aktuellen Raumausschnitt also beispielsweise nur Tische und Stühle, bleibt völlig offen, welche Raumkonzepte unterstützt werden können.

Ein möglicher Lösungsansatz könnte darin bestehen, die Anzahl der auftretenden Objekte einer Klasse zu berücksichtigen. Doch ist auch dieses Vorgehen nicht unproblematisch. Die Anzahl der gleichen Objekte in einem Bildausschnitt würde sehr von dem Blickwinkel und Ausschnitt bzw. den modellierten Instanzen der Ontologie abhängen.

Es könnte zwar Raumkonzepte anhand einer gewissen Anzahl spezieller Objekte differenzieren, beispielsweise sind Hörsäle mit mehr Stuhl- und Tisch-Objekten ausgestattet als Seminarräume, doch wird hier stets ein Bildausschnitt betrachtet, der wiederum nur einen Teil der wirklich vorhandenen Objekte im gesamten Raum repräsentiert. Würde man das System insoweit erweitern, dass von dem aktuellen Blickwinkel aus mehrere Perspektiven als Bildeingaben vorhanden sind, könnten diese kombiniert und dann unter Umständen auch die Anzahl der Objekte berücksichtigt werden.

Da zudem keine scharfen Grenzen gesetzt werden können (beispielsweise bei mehr als 8 Stühlen hohe Unterstützung für „lecture\_hall“, geringe Unterstützung für „kitchen“), da dies zu restriktiv wäre, kann die Raumkonzeptunterstützung nicht direkt aus der Ontologie erschlossen werden. Entweder müssten qualitative Eigenschaften integriert werden oder die Informationen über die jeweilige Objektanzahl müsste in Dempster’s Rule of Combination berücksichtigt werden. Eventuell könnten auch hier (gelernte)

The screenshot shows a software interface for analyzing a 3D scene. The interface is divided into several sections:

- Top Left:** A 3D view of a classroom scene with a red bounding box around a table.
- Top Right:** A 3D view of a table with colored dots (red, green, blue, yellow) at its corners.
- Middle:** A table of analysis results with columns for 'Act. Feat.', 'Proposal of IBIG', 'New Eye Movement', and 'Hypothesis'.
- Bottom:** A list of objects and their corresponding analysis results.

Act. Feat.	Proposal of IBIG	New Eye Movement	Hypothesis
#9 (305,386)	EM 717: 1080,1110, near, 275,220	f: 0.67	projection_surface (0.995) /
#8 (308,495)	no result		projection_surface (0.660) /
#7 (244,462)	EM 949: 1280,1080, far, 0,140	new	projection_surface (0.660) /
#6 (254,505)	EM 261: 1180,1100, near, 130,240	f: 0.99	projection_surface (0.660) /
#5 (258,386)	no result		projection_surface (0.660) /
#4 (244,462)	EM 947: 1260,1180, near, 320,60	m...	projection_surface (0.660) /
#3 (308,495)	EM 480: 1180,1090, near, 45,0	f: 1.98	projection_surface (0.660) /
#2 (254,505)	EM 945: 1280,1180, near, 230,200	m...	projection_surface (0.660) /
#1 (308,495)	EM 33: 1260,1080, near, 90,320	f: 1.98	projection_surface (0.660) /
	EM 944: 1260,1280, near, 90,320	m...	projection_surface (0.660) /
	EM 943: 1280,1260, near, 270,240	m...	projection_surface (0.660) /

object	single object hypothesis	ontology result for complete scene
object 7	projection_surface (0.995) /	lecture_hall (0.999), classroom (0.999)
object 6	chair_wooden_metallic (0.990) /	laboratory (0.184), office (0.025), lecture_hall (0.999), classroom (0.999), kitchen (0.996)
object 5	chair_wooden_metallic (0.990) /	laboratory (0.482), lecture_hall (0.999), office (0.995), classroom (0.999), kitchen (0.996)
object 4	chair_wooden_metallic (0.990) /	laboratory (0.482), lecture_hall (0.999), office (0.995), classroom (0.999), kitchen (0.996)
object 3	desk (0.998) /	laboratory (0.482), lecture_hall (0.999), office (0.995), classroom (0.999), kitchen (0.996)
object 2	chair_wooden_metallic (0.990) /	laboratory (0.482), lecture_hall (0.999), office (0.995), classroom (0.999), kitchen (0.996)
object 1	chair_wooden_metallic (0.990) /	laboratory (0.482), lecture_hall (0.999), office (0.995), classroom (0.999), kitchen (0.996)

Abbildung 7.4: Analyse der Instanz „classroom03“ nach sieben analysierten Objekten

Erfahrungswerte Einfluss auf die Beliefs nehmen.

### 7.1.6 Testfall „kitchen01“

Bei der Untersuchung des Raumausschnitts „kitchen01“ wird im ersten Schritt ein Objekt der Klasse „desk“ korrekt erkannt. Die Ontologieanfrage bezüglich der Raumkonzepte ergibt jeweils sehr hohe Beliefs für alle fünf Konzepte und ist insofern noch nicht sehr aussagekräftig. Im nächsten Schritt wird ein Objekt der Klasse „sink“ identifiziert. Aufgrund dieser beiden Ergebnisse ergeben sich Beliefs von 0.998 für „kitchen“, 0.974 für „laboratory“, 0.591 für „classroom“ und 0.534 für „office“. Wie in Abbildung 7.5 dargestellt, wird im dritten Analyseschritt ein Objekt der Klasse „fridge“ erkannt. Durch diese Identifikation ergeben sich die Beliefs 0.998 für „kitchen“, 0.534 für „office“, 0.166 für „classroom“ und 0.142 für „laboratory“.

Während Spülen bei den Instanzen durchaus noch in anderen Räumen auftreten als nur in Küchen, und die entsprechenden Beliefs nach der Analyse des zweiten Objekts noch nachvollziehbar sind, so mag das Ergebnis überraschen, welches sich nach der dritten Objektanalyse, die einen Kühlschrank identifizierte, ergibt: Weiterhin sind die vier Konzepte Küche, Büro, Seminarraum und Praktikumsraum für den aktuellen Raumausschnitt möglich, wenngleich mit stark abnehmendem Belief in der angegebenen Reihenfolge. Dieses Ergebnis kann unerwartet erscheinen, da ja die Kombination von Spüle und Kühlschrank in der realen Welt sehr häufig in Küchen auftritt und hierdurch beispielsweise Seminarräume und Hörsäle auszuschließen sind.

Das Ergebnis spiegelt aber die Tatsache wieder, dass in der hier modellierten Ontologie keine „allgemeinen Prinzipien der Domäne“ (siehe Kapitel 4.1.1), wie es in Domain Ontologies üblich ist, dargestellt werden, da die verschiedenen Ontologiedefinitionen kombiniert werden. Solche Prinzipien könnten beispielsweise festlegen, dass bei Raumausschnitten, in denen Spülen und Kühlschränke auftreten, bestimmte Raumkonzepte nicht mehr unterstützt werden können. Insofern hat die hier genutzte, allgemeinere Modellierung der Ontologie den Vorteil, dass in ihr die Beziehungen zwischen physikalischen Räumen und Raumkonzepten nicht so restriktiv sind, dass unerwartete Konzepte nicht mehr unterstützt werden. So kann man sich beispielsweise ein Labor vorstellen, in dem ebenfalls Spülen und Küchen auftreten. In Domain Ontologies müssten diese Beziehungen wieder als allgemeines Prinzip modelliert werden, in der vorliegenden Ontologie bräuchten bei einer Erweiterung um das Raumkonzept Labor lediglich die entsprechenden Instanzen angepasst werden.

Act. Feat.	Proposal of IBIG	New Eye Movement	Hypothesis
#6 (358,329)	EM 619: 1180,1180, near, 0,180 f: 0,19	IBIG	fridge (0,993) /
#5 (387,492)	no result	EM 862: 1280,1180, far, 140,240	fridge (0,901) /
#4 (358,329)	EM 48: 1180,1260, far, 220,200 f: 1,63	EM 861: 1180,1280, far, 220,200	fridge (0,901) /
#3 (365,276)	EM 491: 1090,1180, near, 315,220 f: 0,16	IBIG	fridge (0,901) /
#2 (467,271)	EM 555: 1180,1070, far, 315,180 f: 0,99	EM 64: 1180,1090, far, 315,180	fridge (0,981) /
#1 (358,405)	EM 573: 1180,1100, far, 90,320 f: 1,98	EM 685: 1180,1180, far, 90,320	fridge (0,880) /

extracted single object

object 3	fridge (0,993) /	ontology/result for complete scene
object 2	sink (0,990) /	laboratory (0,142), office (0,534), classroom (0,165), kitchen (0,998)
object 1	desk (0,990) /	laboratory (0,374), office (0,534), classroom (0,594), kitchen (0,998)
		laboratory (0,999), lecture_hall (0,999), office (0,999), classroom (0,999), kitchen (0,998)

extract next object  
extract all objects  
extract further objects  
exit

Abbildung 7.5: Analyse der Instanz „kitchen1“ nach dem dritten extrahierten Objekt

## 7.2 Bewertung der Ontologiemodellierung

Der Schwerpunkt bei der Ontologiemodellierung lag darin, dass hier das ontologische mit dem epistemologischen Wissen kombiniert werden sollte. Diese Kombination ist auch umgesetzt worden: Die Ontologiesprache OWL DL bietet Konzepte an, mit denen die Glaubensmaße der Dempster-Shafer-Theorie modelliert werden können. Allerdings wurde Dempster's Rule of Combination nicht in den Reasoningprozess der Ontologie integriert, sondern außerhalb der Ontologie umgesetzt. Dafür wurden die Anfragen entsprechend gestellt, um die Beliefberechnung vorzubereiten.

Wie in den Testfällen zu sehen ist, hat jedoch die Modellierung der Instanzen großen Einfluss auf die Analyseergebnisse. Die in der vorliegenden Arbeit modellierten 25 Instanzen bieten zwar keinen so hohen Informationsgehalt, dass neue Instanzen problemlos analysiert werden könnten (siehe Testfall aus Kapitel 7.1.2). Dennoch ist ersichtlich, wie eine solche Modellierung für weitere Instanzen strukturiert wäre. Auch die Idee, dass neue Instanzen gelernt und diese Informationen in die Ontologie integriert werden können, wurde bereits erwähnt.

Demgegenüber stehen aber die Ergebnisse aus dem Testfall des Kapitels 7.1.6: Da die Modellierung der Instanzen den Ergebnissen der Expertenbefragung (siehe Kapitel F) entnommen sind, ergeben sich gerade für „unerwartete“ Kombinationen einzelner Objekte Beliefs für Raumkonzepte, die in der realen Welt eher seltener solche Kombinationen aufweisen würden. Beziehungsweise richten sich die Belief ausschließlich nach den Begutachtungen der Experten. Durch diese Modellierung sind insbesondere die genannten „allgemeinen Prinzipien“ nicht weiter zu untersuchen. Diese Prinzipien können die Instanzen häufig an zu große Restriktionen binden oder zu allgemein sein, so dass sie den Reasoningprozess nicht unterstützen können.

## 7.3 Bewertung des Bildmaterials

Das Bildmaterial beruht hauptsächlich auf den Fotoaufnahmen in der Universität Bremen. Um wirklich generelle Konzepte der genannten Raumkategorien modellieren zu können, müssten noch mehr Universitätsräume bezüglich ihrer enthaltenen Objekte und Funktionen aufgezeichnet und untersucht werden. Hier ging es aber darum, anhand eines Beispiels die Modellierung von sicherem und unsicherem Wissen zu kombinieren und darzustellen. Insbesondere war auch zu überprüfen, ob das menschliche Wahrnehmen von Raumkonzepten anhand der einzelnen Bausteine in einem System umsetzbar ist und zu welchen Resultaten die Analysen führen.

Die Übertragung des originalen Bildmaterials auf die abstrakten Testbilder ist zudem

ein kritischer Punkt: Es wurde nur ein Teil der im Original vorhandenen Objekte überhaupt modelliert. Die zusätzliche Einschränkung, dass sich Objekte nicht überlagern und jedes Objekt sich durch seinen eigenen Grauwert auszeichnet, führt dazu, dass insgesamt nur wenige Objekte auf einem Testbild modelliert werden können. Dennoch ist festzuhalten, dass sich die Modellierung an den Originalaufnahmen orientiert und keine Testbilder vorkommen, die so nicht auch im Original zu finden wären. Es wurde also beispielsweise keine absichtliche Differenz zwischen den Konzepten modelliert. So bleibt auch die starke Überschneidung einzelner Raumkonzepte bestehen, wie es bei Hörsälen und Seminarräumen der Fall ist. Die Objekte wurden auch nicht nach ihrem Informationsgehalt für ein Raumkonzept ausgewählt, zum Beispiel treten im originalen Bildmaterial nur in Büros Objekte der Klasse Aktenordner auf, welches aber nicht als Konzept modelliert wurde. Doch selbst bei solchen Modellierungen würden auch hier in erster Linie die Glaubensmaße, die sich aus der Expertenbefragung ergeben, Hinweise auf Raumkonzepte liefern.

## 7.4 Bewertung der Glaubensmaße

Obwohl die Glaubensmaße aus den Expertenbefragungen nur einen kleinen Teil der Arbeit ausmachen, sind sie doch sehr entscheidend für die Analyse des Systems. Die Modellierung der Testbilder vollzog sich zwar anhand der Vorgaben aus dem Originalmaterial und damit verbunden war auch das zugehörige Raumkonzept, doch erst die Begutachtung durch die Experten und die Ermittlung der Glaubensmaße für einzelne Beliefs entschieden die möglichen Raumkonzeptunterstützungen des Raumausschnitts.

Dabei sind aber Tendenzen der Ergebnisse (siehe Kapitel F) ersichtlich: Nur mit Ausnahme der Instanz „classroom05“ lag der jeweils höchste, gemittelte Belief nicht bei dem Raumkonzept, von dem aus die Instanz modelliert wurde. Die Unterklassen-Kategorisierung von Veranstaltungsraum in Seminarraum und Hörsaal bzw. Arbeitsraum in Büro und Praktikumsraum spiegelt sich in den Glaubenswerten ebenfalls wider: Instanzen, die einen Belief bei Seminarraum bzw. Büro haben, weisen häufig auch einen bei Hörsaal bzw. Praktikumsraum auf und vice versa. Die funktionale Untergliederung der vorgestellten Raumkonzepte könnte im Nachhinein also ihre Berechtigung haben. Allerdings können die ähnlichen Ergebnisse auch durch die Modellierung der Graphiken beeinflusst sein. Doch, wie oben beschrieben, orientierte sich diese an dem vorhandenen Bildmaterial der originalen Universitätsräume.

Wie bereits erwähnt, haben die Glaubenswerte einen großen Einfluss auf die Analyse. Aufgrund der Ontologiemodellierung und -anfrage entscheiden sie die Zugehörigkeit zu Raumkonzepten. Da die Ergebnisse aus den Expertenbefragungen einen so großen

Anteil an der Analyse ausmachen, sollten diese auch besonders sorgfältig ausfallen. Es ist insofern zu untersuchen, ob die Analysen dadurch verbessert werden können, indem deutlich mehr als nur zwei Experten bezüglich der Glaubensmaße befragt werden.

Da aber die Raumkonzepte die funktionalen Eigenschaften eines physikalischen Raumes darstellen, kann es auch durchaus zu Konflikten bei den Glaubensmaßen kommen, da unterschiedliche Gutachter die Funktionen der Raumausschnitte subjektiv und dadurch unterschiedlich einschätzen. Doch solche Probleme können generell bei Expertenbefragungen auftreten.



## 8 Zusammenfassung

Dass Objekte anhand ihrer Sakkadenzüge, die Menschen auf ihnen ausführen, erkannt und klassifiziert werden können, wurde bereits in einem System umgesetzt (Schill u. a., 2001). Die vorliegende Arbeit erweitert diesen Ansatz, indem komplexere Objekte, die sich aus den bereits analysierbaren Teilobjekten zusammensetzen, anhand ihrer Teile erkannt und klassifiziert werden können. Die Grundlage hierfür sind experimentelle Erkenntnisse bezüglich der menschlichen Wahrnehmung von komplexen Szenen. Dabei wurde sich aber auf die Domäne der Universitätsräume beschränkt. Doch eine Erweiterung für andere Anwendungsbereiche ist nicht auszuschließen.

Dem eingangs erwähnten Zitat von Ivan Herman wird in der vorliegenden Arbeit Rechnung getragen: Es ist gezeigt worden, wie unsicheres Wissen in Ontologien integriert werden kann, insbesondere ohne dabei mit den verschiedenen Sichtweisen auf Ontologien zu brechen. Es wurde eine Möglichkeit vorgeschlagen, die Glaubensmaße der Dempster-Shafer-Theorie ontologisch zu modellieren und zu nutzen. Die Kombination der Glaubensmaße gemäß der Dempster's Rule of Combination erfolgte dabei außerhalb der Ontologie.

Bezüglich der Dempster-Shafer-Theorie ist festzuhalten, dass sich auf die Confirming Beliefs für Single Hypotheses beschränkt wurde. Diese Beschränkung reduziert die Rechenkomplexität, vereinfacht die Handhabung der Glaubensmaße und war für die beschriebene Modellierung völlig ausreichend. Es bliebe zu untersuchen, ob eine Aufhebung dieser Beschränkung die Analyseergebnisse verbessern würde.

Auch muss erwähnt werden, dass zwar die ursprünglich als sehr vorteilhaft erachtete hierarchische Struktur der Ontologie sich sehr gut für die modellierten Objekte eignet. Doch im Verlauf der Analyse wurde nur noch wenig auf mögliche Vorteile, die sich durch eine solche Struktur ergeben, zurückgegriffen. Die Analyse wurde nicht primär für hierarchisch angeordnete Daten entwickelt oder daraufhin optimiert, wenngleich die Nutzung einer Ontologie eine solche Hierarchie ohnehin vorgibt und die Anfragen auch Konzeptzugehörigkeiten erfassen. Die Unterstützung durch die baumartige Anordnung für das Reasoning wurde bzw. brauchte nicht genutzt zu werden.

Die vorliegende Arbeit hat aber auch Fragen aufgeworfen, die weiter zu untersuchen sind. Die Modellierung einer Erwartungshaltung gegenüber den Objekten aufgrund der analysierten Objekte wurde bereits vorgeschlagen. Ähnlich dem Verfahren, das

bei IBIG verwendet wird, könnten auch die Intersakkaden zielgerichtet ausgeführt werden. Die Ontologieanfrage im anschließenden Schritt würde bereits darüber Auskunft geben, wie groß der Informationsgewinn aufgrund der getätigten Augenbewegung ist. Bei unverändertem Antwortverhalten der Ontologie ist kein Informationsgewinn durch das analysierte Objekt festzustellen. Ein Lernen solcher Augenbewegungen könnte als zusätzliches Wissen für zukünftige Analysen berücksichtigt werden und den Analyseprozess optimieren und zeitlich reduzieren.

Ebenso könnte zusätzliches Weltwissen über Gruppierungen oder weitere Eigenschaften von Objekten im Kontext physikalischer Räume modelliert werden. Zwar wurde auf eine Modellierung allgemeiner Prinzipien bewusst verzichtet, doch könnten erkannte Relationen zwischen Objektanordnungen auf zusätzliche Informationen hinweisen. Geht man von dem eingangs erwähnten Anwendungsfall eines Roboters aus, der seine Umwelt auf der Basis des vorgestellten Systems wahrnimmt und analysiert, könnte er beispielsweise aufgrund der Anordnung von Tischen und Stühlen zueinander schließen, ob er sich im Fall eines Hörsaals im hinteren oder vorderen Bereich befindet. Ob sich die Stühle von ihm aus also tendenziell vor oder hinter den Tischen befinden.

In Anlehnung an das bestehende Okusys-System könnte in der Erweiterung auch ein Lernen der Raumkonzepte und ein Belief-Update innerhalb der Ontologie stattfinden. Anstelle der initial modellierten Instanzen in der Ontologie, könnten aufgrund von Testbildern, die ein entsprechendes Feedback für ihre Unterstützung von Raumkonzepten liefern, das System selbst Instanzen in der Ontologie modellieren. Dabei könnten am Anfang der Lernphase aber keine Ontologieanfragen gestellt werden, da diese noch keine Ergebnisse liefern könnten. Entsprechende Alternative müssten erarbeitet werden.

Wie und ob die Berechnung der Beliefs durch die Dempster's Rule of Combination in den Reasoningprozess der Ontologie integrierbar ist, bleibt auch noch zu untersuchen.

Bereits angesprochen wurde auch der sehr einfach gehaltene Segmentieralgorithmus. Eine Erweiterung des Systems um die Integration einer komplexeren Segmentierung könnte die Objektidentifikation begünstigen und schränkt die Objektmodellierung weniger ein. Überschneidungen von Objekten im Bild bleiben allerdings weiterhin ein Problem und müssten entsprechend gelöst werden.

Erwähnung fand auch die Analyse von einer Kombination von Raumausschnitten aus einer Perspektive, so dass die unterschiedlichen Analysen in ihrer Kombination fehlertoleranter ausfallen könnten, als für sich allein betrachtet. Man stelle sich wieder den Anwendungsfall eines Roboters vor: Dieser könnte seine 3D-Eingabedaten auf zweidimensionale Bildausschnitte reduzieren und mit dem System untersuchen. Neben den Augenbewegungen könnten so auch Kopfbewegungen berücksichtigt werden. Statt nur auf den Intra- und Intersakkaden zu arbeiten, würde die Analyse bildübergreifend stattfinden können.

---

Auch der eingangs angesprochenen Anwendungsfall eines „intelligenten Gebäudes“ könnte mit der hier beschriebenen Arbeit umgesetzt werden. Anfragen bezüglich des Verwendungszwecks für einen Raum könnten in Abhängigkeit der unterstützten Raumkonzepte beantwortet werden. Räume, die beispielsweise einen sehr hohen Belief für Küchen aber einen sehr niedrigen Belief für Hörsäle aufweisen, würden so nicht für eine Lehrveranstaltung in Frage kommen. Spezielle Anfragen, ob beispielsweise ein Projektor in einem Raum vorhanden ist, könnten auch integriert werden. Bezieht man zudem neben den klassifizierbaren Objekten eines Raumes auch Menschen ein, die sich dort aufhalten, können entsprechende Rückschlüsse über die Belegung von Räumen stattfinden. (Wie kritisch ein solches kameraüberwachtes Gebäude aus Sicht des Datenschutzes ist, sei hier einmal unberücksichtigt.)

Im ganzen wurden also Teilaspekte der menschlichen Wahrnehmung und Identifikation von Objekten und Räumen im System modelliert. Bezüglich der konkreten Modellierung wurden verschiedene Gebiete der Wissensrepräsentation und -verarbeitung kombiniert. Von diesem Punkt aus bieten sich viele weitere Untersuchungsmöglichkeiten an.



## A Literaturverzeichnis

- [Abmayr 1994] ABMAYR, Wolfgang: *Einführung in die digitale Bildverarbeitung*. Stuttgart : B. G. Teubner, 1994
- [Aliev u. a. 2000] ALIEV, Rafik A. ; BONFIG, Karl W. ; ALIEW, Fuad: *Soft Computing - Eine grundlegende Einführung*. Huss-Medien, Verlag Technik, 2000
- [Anderson 2001] ANDERSON, John R.: *Kognitive Psychologie*. Heidelberg : Spektrum der Wissenschaft, 2001
- [Antoniou und van Harmelen 2004] ANTONIOU, Grigoris ; HARMELEN, Frank van: Web Ontology Language: OWL. In: STAAB, Steffen (Hrsg.) ; STUDER, Rudi (Hrsg.): *Handbook on Ontologies*. Springer-Verlag, 2004, S. 67–92
- [Bajcsy 1988] BAJCSY, Ruzena: Active Perception. In: *Proceedings of the IEEE* Bd. 76, 1988, S. 996–1005
- [Barnett 1981] BARNETT, Jeffrey A.: Computational Methods for a Mathematical Theory of Evidence. In: *Proceedings of the 7th International Joint Conference on Artificial Intelligence*. Vancouver, Canada, 1981, S. 868–875
- [Bateman 2005] BATEMAN, John: *Ontologie-Workshop des SFB/TR8 Spatial Cognition: Reasoning, Action, Interaction*. Juni 2005. – Tutzing
- [Bateman u. a. 1995] BATEMAN, John A. ; HENSCHEL, Renate ; RINALDI, Fabio: The Generalized Upper Model 2.0. 1995. – Forschungsbericht
- [Bhatnagar und Kanal 1986] BHATNAGAR, Raj K. ; KANAL, Laveen N.: Handling Uncertain Information: A Review of Numeric and Non-Numeric Methods. In: KANAL, Laveen N. (Hrsg.) ; LEMMER, John F. (Hrsg.): *Uncertainty in Artificial Intelligence*. Amsterdam : North-Holland, 1986, S. 3–26
- [Borgelt u. a. 2003] BORGELT, Christian ; TIMM, Heiko ; KRUSE, Rudolf: *Unsicheres und vages Wissen*. Kap. 9, S. 291–347. In: GÖRZ, Günther (Hrsg.) ; ROLLINGER, Claus-Rainer (Hrsg.) ; SCHNEEBERGER, Josef (Hrsg.): *Handbuch der künstlichen Intelligenz*, Oldenbourg Verlag München Wien, 2003

- [Brendel und Koppelberg 1996] BRENDEL, Elke ; KOPPELBERG, Dirk: Erkenntnistheorie, Epistemologie. In: STRUBE, Gerhard (Hrsg.) ; BECKER, Barbara (Hrsg.) ; FREKSA, Christian (Hrsg.) ; HAHN, Udo (Hrsg.) ; OPWIZ, Klaus (Hrsg.) ; PALM, Günther (Hrsg.): *Wörterbuch der Kognitionswissenschaft*. Stuttgart, Germany : Klett-Cotta, 1996, S. 151–157
- [Caelli u. a. 1997] CAELLI, Terry ; DILLON, Craig ; OSMAN, Erol ; KRIEGER, Gerhard: The IPRS image processing and pattern recognition system. In: *Spatial Vision* 11 (1) (1997), S. 107–116
- [Corcho und Gómez-Pérez 2000] CORCHO, Oscar ; GÓMEZ-PÉREZ, Asunción: A Roadmap to Ontology Specification Languages. In: DIENG, Rose (Hrsg.) ; CORBY, Olivier (Hrsg.): *Proceedings of the 12th International Conference on Knowledge Engineering and Knowledge Management*. Berlin : Springer, Oktober 2000, S. 80–96
- [Dempster 1968] DEMPSTER, Arthur P.: A generalization of Bayesian inference. In: *Journal of the Royal Statistical Society Series B* 30 (1968), S. 205–247
- [Deubel 1994] DEUBEL, Heiner: *Visuelle Verarbeitung und kognitive Faktoren bei der Generierung sakkadischer Augenbewegungen*. Kap. 5, S. 189–153. In: PRINZ, Wolfgang (Hrsg.) ; BRIDGEMAN, Bruce (Hrsg.): *Enzyklopädie der Psychologie, Serie II: Kognition, Band 1: Wahrnehmung*, Hogrefe-Verlag, Göttingen, 1994
- [Ding und Peng 2004] DING, Zhongli ; PENG, Yun: A Probabilistic Extension to Ontology Language OWL. In: *Proceedings of the 37th Hawaii International Conference On System Sciences (HICSS-37)*. Hawaii : IEEE, Januar 2004. – URL <http://ebiquity.umbc.edu/v2.1/paper/html/id/92/?EBS=f6fe71d1e04f5d4d205de61da41617ab>
- [Ding u. a. 2004] DING, Zhongli ; PENG, Yun ; PAN, Rong: A Bayesian Approach to Uncertainty Modeling in OWL Ontology. In: *Proceedings of the International Conference on Advances in Intelligent Systems - Theory and Applications*. Luxemburg : IEEE, November 2004, S. 9. – URL <http://ebiquity.umbc.edu/paper/html/id/204/?EBS=287197e7c86f76bb9da8c108a5bd1b0f>
- [Emde 1996] EMDE, Werner: Maschinelles Lernen. In: STRUBE, Gerhard (Hrsg.) ; BECKER, Barbara (Hrsg.) ; FREKSA, Christian (Hrsg.) ; HAHN, Udo (Hrsg.) ; OPWIZ, Klaus (Hrsg.) ; PALM, Günther (Hrsg.): *Wörterbuch der Kognitionswissenschaft*. Stuttgart, Germany : Klett-Cotta, 1996, S. 362–376
- [Engelkamp 1990] ENGELKAMP, Johannes: *Das menschliche Gedächtnis*. Göttingen : Verlag für Psychologie, 1990

- [Ferber 1999] FERBER, Rafael: *Philosophische Grundbegriffe*. München : C. H. Beck, 1999
- [Ferber 2003] FERBER, Reginald: *Information Retrieval - Suchmodelle und Data-Mining-Verfahren für Textsammlungen und das Web*. Heidelberg : dpunkt.verlag, 2003. – URL <http://information-retrieval.de/irb/irb.html>
- [Gómez-Pérez u. a. 2004] GÓMEZ-PÉREZ, Asunción ; FERNÁNDEZ-LÓPEZ, Mariano ; CORCHO, Oscar: *Ontological Engineering with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web*. London : Springer-Verlag, 2004
- [Goldenberg 1998] GOLDENBERG, Georg: *Neuropsychologie: Grundlagen, Klinik, Rehabilitation*. Stuttgart : Fischer Verlag, 1998
- [Gordon und Shortliffe 1985] GORDON, Jean ; SHORTLIFFE, Edward H.: A Method for Managing Evidential Reasoning in a Hierarchical Hypothesis Space. In: *Artificial Intelligence* 26 (1985), S. 323–357
- [Gruber 1993] GRUBER, Thomas R.: A translation approach to portable ontology specifications. In: *Knowledge Acquisition* 5 (1993), Juni, S. 199–220
- [Herman 2005] HERMAN, Ivan: Automobilindustrie setzt am stärksten auf Semantic Web. In: *Computer Zeitung* 28 (2005), Juli, S. 15
- [Hermes 2005] HERMES, Thorsten: *Digitale Bildverarbeitung*. München Wien : Carl Hanser Verlag, 2005
- [Herrmann 1996] HERRMANN, Theo: Blickpunkte und Blickpunktsequenzen. In: *Sprache und Kognition* 15 (1996), S. 159–177
- [Hubel 1989] HUBEL, David H.: *Auge und Gehirn: Neurobiologie des Sehens*. Heidelberg : Spektrum Verlag, 1989
- [Knauff 1997] KNAUFF, Markus: *Räumliches Wissen und Gedächtnis*. Wiesbaden : Deutscher Universitätsverlag, 1997
- [Kodratoff und Michalski 1990] KODRATOFF, Yves ; MICHALSKI, Ryszard S.: Research in Machine Learning: Recent Progress, Classification of Methods, and Future Directions. In: KODRATOFF, Yves (Hrsg.) ; MICHALSKI, Ryszard S. (Hrsg.): *Machine Learning: An Artificial Intelligence Approach (Volume III)*. San Mateo, CA : Kaufmann, 1990, S. 3–30

- [Kosslyn 1994] KOSSLYN, Stephen M.: *Image and brain: the resolution of the imagery debate*. Cambridge : MIT Press, 1994
- [Mallot 2000] MALLOT, Hanspeter A.: *Sehen und die Verarbeitung visueller Information - Eine Einführung*. Wiesbaden : Vieweg, 2000
- [Mallot 2004] MALLOT, Hanspeter A.: Kompass im Kopf. In: *Gehirn und Geist* 5 (2004), S. 18–26
- [Mallot u. a. 1992] MALLOT, Hanspeter A. ; KOPECZ, Jörg ; SEELEN, Werner von: Neuroinformatik als empirische Wissenschaft. In: *Kognitionswissenschaft* 3 (1992), S. 12–23
- [Marr 1982] MARR, David: *Vision*. W. H. Freeman and Company, 1982
- [Masolo u. a. 2003] MASOLO, Claudio ; BORGIO, Stefano ; GANGEMI, Aldo ; GUARINO, Nicola ; OLTRAMARI, Alessandro: WonderWeb Deliverable D18 - Ontology Library / Laboratory For Applied Ontology - ISTC-CNR. Trento, 2003. – final
- [Metzger 1975] METZGER, Wolfgang: *Gesetze des Sehens*. Frankfurt am Main : Kramer, 1975
- [Neumann 2003] NEUMANN, Bernd: *Bildverstehen - ein Überblick*. Kap. 21, S. 815–841. In: GÖRZ, Günther (Hrsg.) ; ROLLINGER, Claus-Rainer (Hrsg.) ; SCHNEEBERGER, Josef (Hrsg.): *Handbuch der künstlichen Intelligenz*, Oldenbourg Verlag München Wien, 2003
- [Niemann 1985] NIEMANN, Heinrich: Wissensbasierte Bildanalyse. In: *Informatik-Spektrum* 8 (1985), S. 201–214
- [Noton und Stark 1971a] NOTON, David ; STARK, Lawrence: Scanpaths in Eye Movements during Pattern Perception. In: *Science* 171 (1971), S. 308–311
- [Noton und Stark 1971b] NOTON, David ; STARK, Lawrence: Scanpaths in Saccadic Eye Movements while Viewing and Recognizing Patterns. In: *Vision* 11 (1971), S. 929–942
- [Owsnicki-Klewe u. a. 2003] OWSNICKI-KLEWE, Bernd ; LUCK, Kai v. ; NEBEL, Bernhard: *Wissensrepräsentation und Logik - Eine Einführung*. Kap. 5, S. 153–197. In: GÖRZ, Günther (Hrsg.) ; ROLLINGER, Claus-Rainer (Hrsg.) ; SCHNEEBERGER, Josef (Hrsg.): *Handbuch der künstlichen Intelligenz*, Oldenbourg Verlag München Wien, 2003



- [Pan u. a. 2005] PAN, Jeff Z. ; STAMOU, Giorgos ; TZOUVARAS, Vassilis ; HORROCKS, Ian: A Fuzzy Extension of SWRL. In: *Proceedings of the International Conference on Artificial Neural Networks (ICANN 2005), Special section on „Intelligent multimedia and semantics“*, 2005. – to appear
- [Plach 1998] PLACH, Marcus: *Prozesse der Urteilsrevision - Kognitive Modellierung der Vorverarbeitung unsicheren Wissens*. Wiesbaden : Deutscher Universitätsverlag, 1998
- [Puppe u. a. 2003] PUPPE, Frank ; STOYAN, Herbert ; STUDER, Rudi: *Knowledge Engineering*. Kap. 15, S. 599–641. In: GÖRZ, Günther (Hrsg.) ; ROLLINGER, Claus-Rainer (Hrsg.) ; SCHNEEBERGER, Josef (Hrsg.): *Handbuch der künstlichen Intelligenz*, Oldenbourg Verlag München Wien, 2003
- [Reiss u. a. 1992] REISS, Kristina ; REISS, Matthias ; SPANDL, Horst: *Maschinelles Lernen - Modellierung von Lernen mit Maschinen*. Berlin, Heidelberg : Springer-Verlag, 1992
- [Richards 2003] RICHARDS, John E.: The Development of Visual Attention and the Brain. In: HAAN, Michelle de (Hrsg.) ; JOHNSON, Mark H. (Hrsg.): *The Cognitive Neuroscience of Development*, Psychology Press, 2003, S. 73–98
- [Ritter und Gründer 1971ff] RITTER, Joachim (Hrsg.) ; GRÜNDER, Karlfried (Hrsg.): *Historisches Wörterbuch der Philosophie, Bd. 12*. Basel, 1971ff
- [Roth 1997] ROTH, Gerhard: *Das Gehirn und seine Wirklichkeit*. Frankfurt am Main : Suhrkamp Verlag, 1997
- [Russell und Norvig 2003] RUSSELL, Stuart J. ; NORVIG, Peter ; RUSSELL, Stuart J. (Hrsg.) ; NORVIG, Peter (Hrsg.): *Artificial Intelligence - A Modern Approach*. New Jersey : Prentice Hall, 2003
- [Schill 1990] SCHILL, Kerstin ; SCHNUPP, Peter (Hrsg.): *Medizinische Expertensysteme: Methoden und Techniken*. München, Wien : Oldenbourg Verlag, 1990
- [Schill 1997] SCHILL, Kerstin: Decision Support Systems with Adaptive Reasoning Strategies. In: FREKSA, Christian (Hrsg.) ; JANTZEN, Matthias (Hrsg.) ; VALK, Rüdiger (Hrsg.): *Foundations of Computer Science: Potential - Theory - Cognition* Bd. 1337, Springer, 1997, S. 417–427

- [Schill u. a. 2001] SCHILL, Kerstin ; UMKEHRER, Elisabeth ; BEINLICH, Stephan ; KRIEGER, Gerhard ; ZETZSCHE, Christoph: Scene analysis with saccadic eye movements: Top-down and bottom-up modeling. In: *Journal of electronic imaging* 10 (1) (2001), S. 152–160
- [Shafer 1976] SHAFER, Glenn: *A Mathematical Theory of Evidence*. Princeton : Princeton University Press, 1976
- [Shafer 1992] SHAFER, Glenn: The Dempster-Shafer theory. In: SHAPIRO, Stuart C. (Hrsg.): *Encyclopedia of Artificial Intelligence, second edition*. Wiley, 1992, S. 330–331. – URL <http://www.glennshafer.com/assets/downloads/article48.pdf>
- [Smith 2005] SMITH, Barry: *Ontology and Information Systems*. DRAFT. 2005. – URL <http://ontology.buffalo.edu/smith/articles/ontologies.htm>. – zuletzt gesichtet am 28.06.2005
- [Smith und Varzi 1999] SMITH, Barry ; VARZI, Achille: The Niche. In: *Noûs* 33 (1999), Nr. 2, S. 198–222
- [Sowa 2000] SOWA, John F.: *Knowledge Representation - Logical, Philosophical, and Computational Foundations*. Brooks/Cole, 2000
- [Studer u. a. 1998] STUDER, Rudi ; BENJAMINS, V. R. ; FENSEL, Dieter: Knowledge Engineering: Principles and Methods. In: *Data Knowledge Engineering* 25 (1998), Nr. 1-2, S. 161–197
- [Uschold und Grüninger 1996] USCHOLD, Mike ; GRÜNINGER, Michael: Ontologies: Principles, Methods and Applications. In: *Knowledge Engineering Review* 11 (1996), Februar, S. 93–155
- [Waltz 1975] WALTZ, David L.: Understanding line drawings of scenes with shadows. In: WINSTON, Patrick H. (Hrsg.): *The Psychology of Computer Vision*. New York : McGraw-Hill, 1975
- [Yarbus 1967] YARBUS, Alfred L.: *Eye Movements and Vision*. New York : Plenum Press, 1967. – Übersetzung: B. Haigh
- [Zadeh und Kacprzyk 1992] ZADEH, Lotfi A. ; KACPRZYK, Janusz ; ZADEH, Lotfi A. (Hrsg.) ; KACPRZYK, Janusz (Hrsg.): *Fuzzy logic for the management of uncertainty*. Wiley, 1992

## B Abbildungsverzeichnis

2.1	Visuelles Schema der Sehbahnen (nach Hubel (1989)) . . . . .	14
2.2	Ablauf visueller Objektwahrnehmung (nach Marr (1982)) . . . . .	19
2.3	Prozessablauf des Analysesystems (nach Niemann (1985)) . . . . .	22
2.4	Ausschnitt aus der Taxonomie von UMLS (nach Gómez-Pérez u. a. (2004))	26
3.1	Sakkadenzüge unterschiedlicher Probanden, aus Yarbus (1967) . . . . .	46
3.2	Sakkadenzüge desselben Probanden, aus Yarbus (1967) . . . . .	47
4.1	Auszug aus der Taxonomie der Basiskonzepte von DOLCE . . . . .	52
4.2	Niches als Eigenschaft der Konzepte . . . . .	53
4.3	ontologische (oben) und epistemologische (unten) Sichtweise . . . . .	59
5.1	Universitätsraumkonzepte . . . . .	62
5.2	Einrichtungsgegenstände der Ontologie . . . . .	65
5.3	Taxonomie der Ontologie . . . . .	69
5.4	Rauminstanz „office01“ . . . . .	72
5.5	Objektidentifikation von „shelf01“ durch das Okusys-System . . . . .	81
5.6	Raumanalyse von „office05“ - erstes Objekt . . . . .	83
5.7	Raumanalyse von „office05“ - zweites Objekt . . . . .	85
5.8	Raumanalyse von „office05“ - viertes Objekt . . . . .	87
5.9	Raumanalyse von „office05“ - sechstes Objekt . . . . .	88
5.10	Raumanalyse von „office05“ - zehntes (letztes) Objekt . . . . .	89
6.1	Okusys-Auswahlfenster . . . . .	92
6.2	Klassendiagramm . . . . .	96
7.1	Analyse der Instanz „laboratory04“ nach drei analysierten Objekten (angezeigt wird das erste analysierte Objekt) . . . . .	101
7.2	Analyse der Instanz „laboratory04“ nach fünf analysierten Objekten (angezeigt wird das zweite analysierte Objekt) . . . . .	103
7.3	Analyse der Instanz „lecture_hall02“ nach elf analysierten Objekten . .	105
7.4	Analyse der Instanz „classroom03“ nach sieben analysierten Objekten .	107

7.5	Analyse der Instanz „kitchen1“ nach dem dritten extrahierten Objekt .	109
G.1	Relationen der Beispielontologie . . . . .	139
G.2	Taxonomie der Beispielontologie . . . . .	140

## C Tabellenverzeichnis

2.1	Kombination der Beliefwerte: Glaubwürdigkeit der Quellen . . . . .	31
5.1	Ergebnis der Anfrage bezüglich der (physikalischen) Rauminstanzen . .	76
5.2	Ergebnis der Anfrage bezüglich der Raumkonzeptinstanzen und Glaubensmaße von „office04“ . . . . .	77
5.3	Raumkonzepte und Glaubensmaße für Raumausschnitte mit den enthaltenen Objekten Tisch, Stuhl und Schrank . . . . .	78
5.4	Belief der Raumkonzepte für einen Raumausschnitt mit Tisch, Stuhl und Schrank . . . . .	79
F.1	Expertenbefragung: Experte 1 . . . . .	135
F.2	Expertenbefragung: Experte 2 . . . . .	136
F.3	Arithmetisches Mittel der Expertenbefragung und Aufschlüsselung der Raumnummern . . . . .	137



## D CD-Datenträger

Auf der beigefügten CD befinden sich folgende Ordner und Dateien:

- Ordner „okusys“: Software des Okusys-Systems
  - „okusys/imp/version3“: Source Code des Okusys-Systems
  - „okusys/imp/version3/java“: Source Code der Java-Implementierung und weitere Ordner der Bibliotheken JIU und Pellet
  - „okusys/imp/version3/dat/inst6“: Dateien für die Wissensbasis der Analyse
    - \* „scenes“: Auflistung aller Graphikdateien, die in Okusys eingebunden werden sollen
    - \* „/knowledge“: Hierarchische Strukturierung der Graphikdateien
    - \* „/ontology“: Ontologie-Datei für die Analyse
    - \* „/results“: Ergebnisse der Bildanalyse
    - \* „/sequences“: Protokollierung der zuletzt gelernten Objekte
    - \* „/training“: Trainingsdaten der gelernten Objekte
  - „okusys/scene\_base/physicalroom“: Graphikdateien (Formate: byte und GIF) der Testbilder für die Analyse mit Okusys
- „iprs/iprs\_SuSE8.2/gerdiprs/bin/ppm2byte“: Programm für die Umwandlung in das IPRS-Byte-Format für Bilder
- Ordner „Ontologie“: Ontologie in OWL und Projektdatei für Protege
- Ordner „Fotoaufnahmen“: Fotoaufnahmen der Universitätsräume





# E Testbilder der Raumkonzeptinstanzen

---

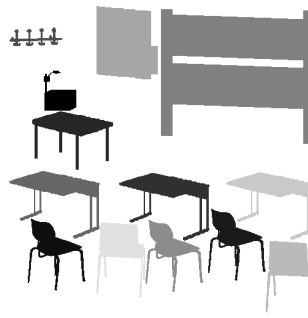
## Seminarräume

---

classroom01



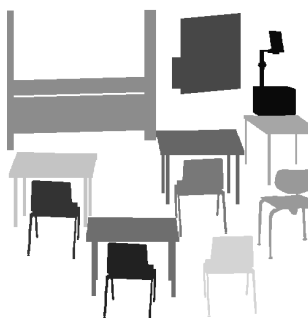
classroom02



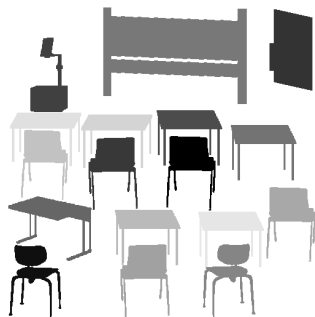
classroom03



classroom04



classroom05

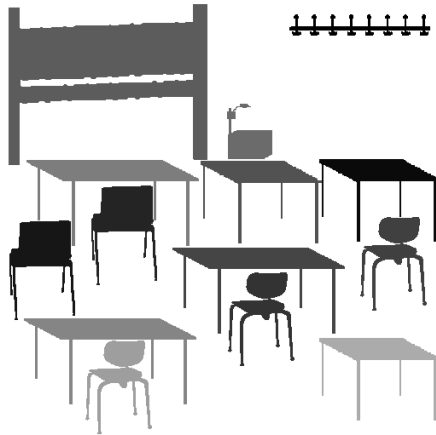


---

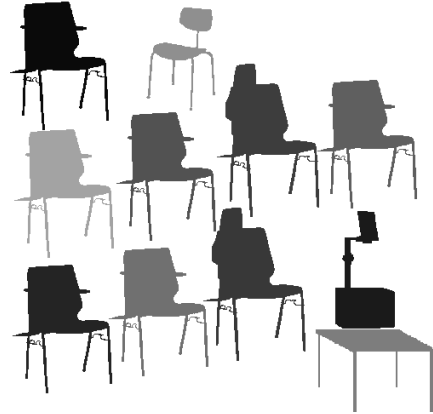
Hörsäle

---

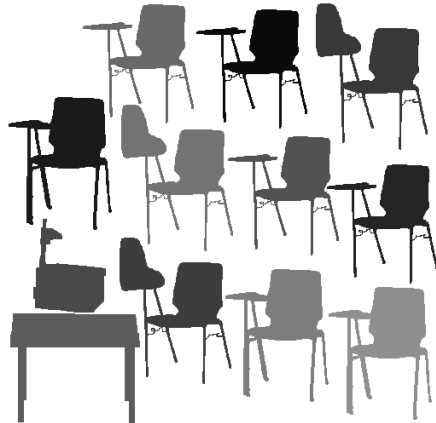
lecture\_hall01



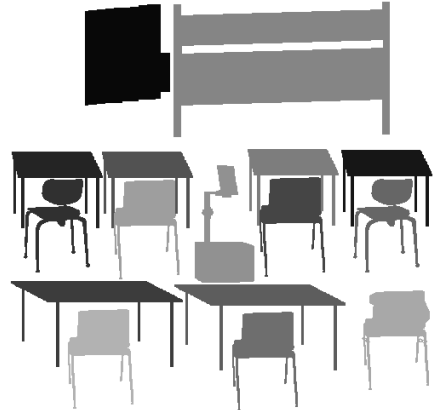
lecture\_hall02



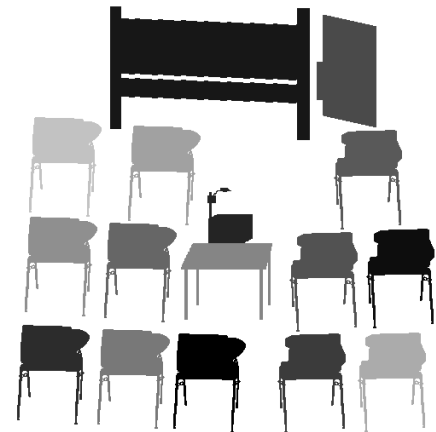
lecture\_hall03



lecture\_hall04



lecture\_hall05



---

Küchen

---

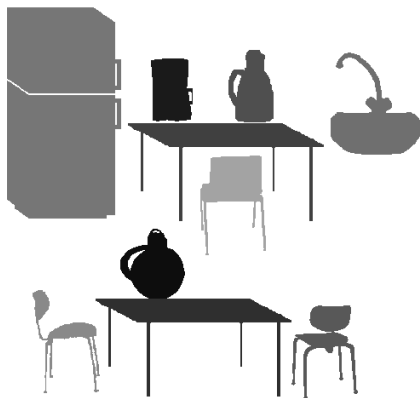
kitchen01



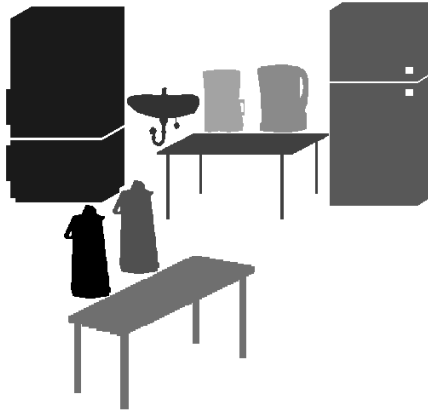
kitchen02



kitchen03



kitchen04



kitchen05

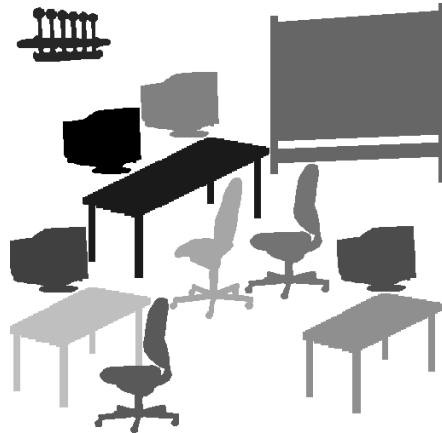


---

Praktikumsräume

---

laboratory01



laboratory02



laboratory03



laboratory04



laboratory05

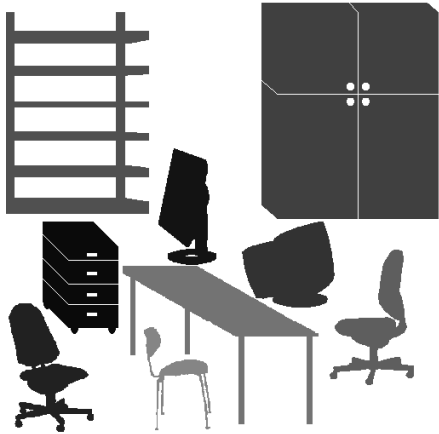


---

Büros

---

office01



office02



office03



office04



office05





## F Glaubenswerte von Raumkonzepten der Rauminstanzen (Expertenbefragung)

Tabelle F.1: Expertenbefragung: Experte 1

Raumnummer	Seminarraum	Vorlesungshalle	Küche	Praktikumsraum	Büro
01	0.9	0.8		0.1	0.05
02	0.8	0.9		0.1	
03	0.8	0.9		0.1	
04	0.75	0.95		0.05	
05	0.75	0.95		0.05	
06	0.2		0.7	0.1	0.4
07			0.8		0.3
08			0.5	0.05	0.4
09	0.1		0.5	0.1	0.2
10	0.05		0.4	0.05	0.1
11		0.05		0.8	0.1
12				0.9	
13				0.9	
14	0.3			0.8	
15	0.2			0.8	
16	0.85	0.9			
17	0.2	0.9			
18	0.15	0.9			
19	0.7	0.8			
20	0.3	0.9			
21				0.1	0.85
22				0.05	0.9
23				0.1	0.8
24				0.05	0.75
25					0.7

Tabelle F.2: Expertenbefragung: Experte 2

Raumnummer	Seminarraum	Vorlesungshalle	Küche	Praktikumsraum	Büro
01	0.9				
02	0.7	0.5			
03	0.7	0.5			
04	0.8				
05	0.7	0.8			
06			0.85		
07			0.9		
08			0.85		
09			0.8		
10			0.9		
11	0.15			0.8	0.1
12				0.9	
13				0.9	
14	0.3			0.8	
15	0.4			0.9	
16	0.7	0.9			
17	0.7	0.8			
18	0.7	0.8			
19	0.6	0.85			
20	0.6	0.85			
21				0.3	0.8
22					0.9
23					0.9
24				0.1	0.8
25				0.1	0.8



Tabelle F.3: Arithmetisches Mittel der Expertenbefragung und Aufschlüsselung der Raumnummern

Raumnummer (Rauminstanz)	Seminarraum	Vorlesungshalle	Küche	Praktikumsraum	Büro
01 (classroom01)	0.9	0.4		0.05	0.025
02 (classroom02)	0.75	0.7		0.05	
03 (classroom03)	0.75	0.7		0.05	
04 (classroom04)	0.775	0.475		0.025	
05 (classroom05)	0.725	0.875		0.025	
06 (kitchen01)	0.1		0.775	0.05	0.2
07 (kitchen02)			0.85		0.15
08 (kitchen03)			0.675	0.025	0.2
09 (kitchen04)	0.05		0.65	0.05	0.1
10 (kitchen05)	0.025		0.65	0.025	0.05
11 (laboratory01)	0.075	0.025		0.8	0.1
12 (laboratory02)				0.9	
13 (laboratory03)				0.9	
14 (laboratory04)	0.3			0.8	
15 (laboratory05)	0.3			0.85	
16 (lecture_hall01)	0.775	0.9			
17 (lecture_hall02)	0.45	0.85			
18 (lecture_hall03)	0.425	0.85			
19 (lecture_hall04)	0.65	0.825			
20 (lecture_hall05)	0.45	0.875			
21 (office01)				0.2	0.825
22 (office02)				0.025	0.9
23 (office03)				0.05	0.85
24 (office04)				0.075	0.775
25 (office05)				0.05	0.75



## G Relationen und Taxonomie der Beispielontologie

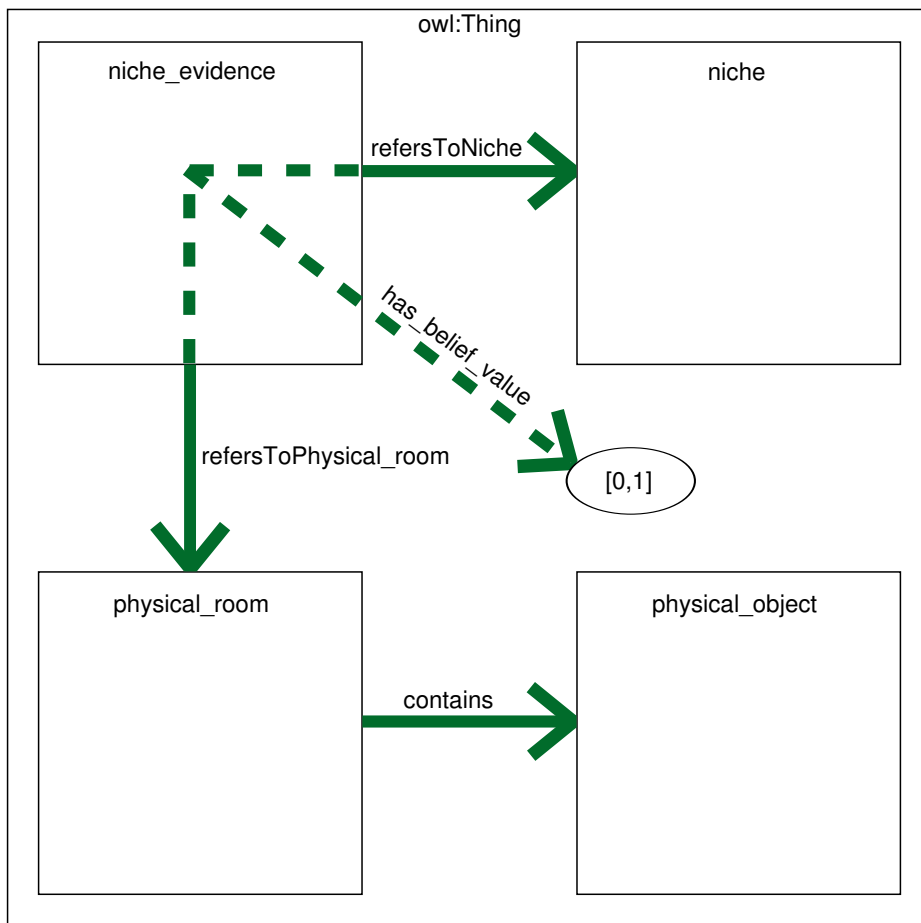


Abbildung G.1: Relationen der Beispielontologie

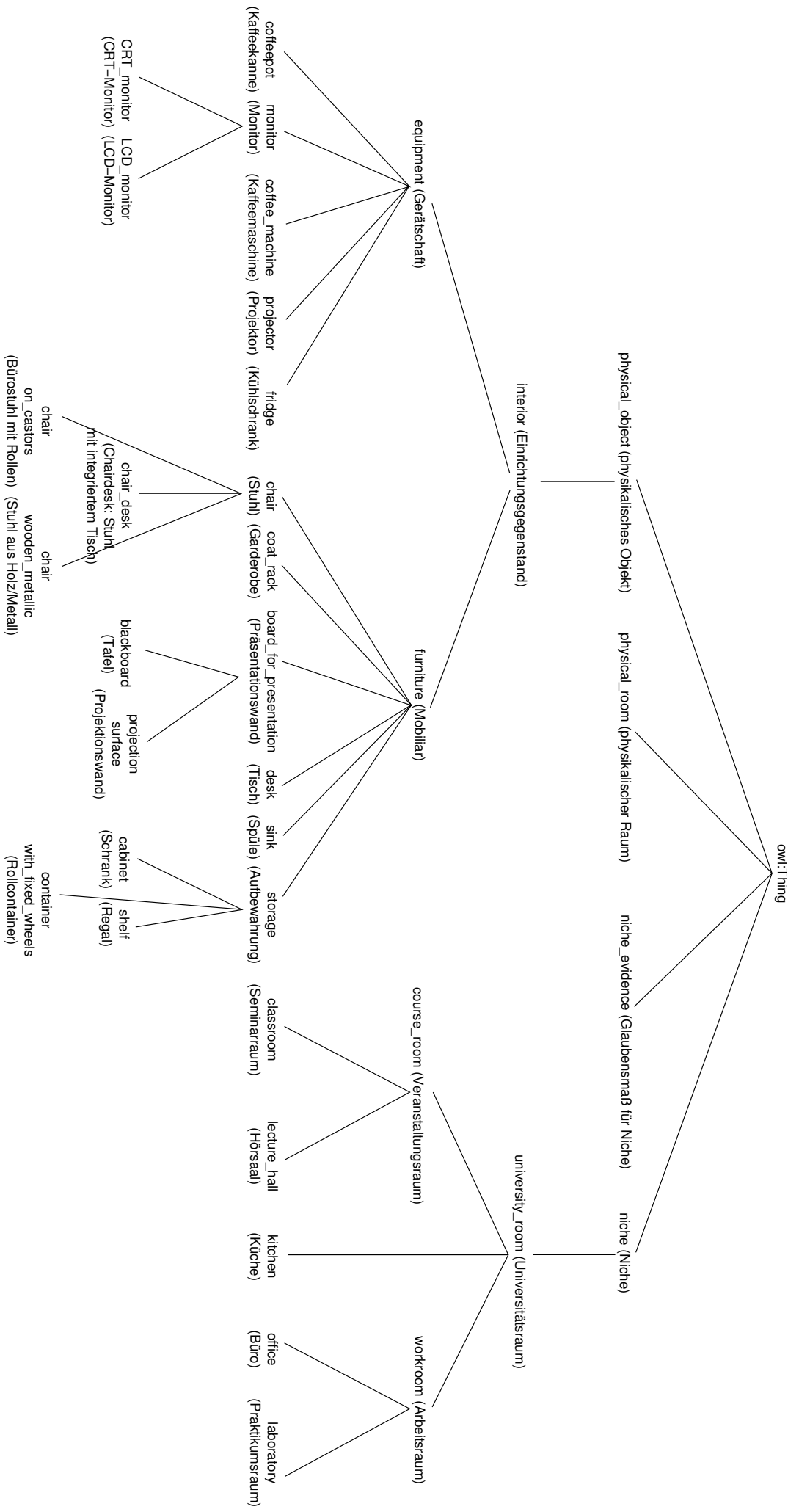


Abbildung G.2: Taxonomie der Beispielontologie